A Present-Value Approach to Variable Selection

Jhe Yun*

September 22, 2011†

Abstract

I propose a present-value approach to study which variables forecast returns and dividend growth rates, individually as well as jointly. This approach explicitly models time-varying expected returns and expected dividend growth rates, and uses information contained in additional predictors to filter them out from the price-dividend ratio within a present-value model. Using my approach, I can predict returns and dividend growth rates on the aggregate stock market with R-squared values of 17% and 23%, respectively. I find that Consumption-Wealth-Income Ratio, Equity Issuing Activity Ratio, and Long Term Rate of Returns significantly improve the return and dividend forecasts of the present-value model. The approach outperforms standard predictive regressions both in-sample and out-of-sample.

1 Introduction

Are returns predictable? Are dividend growth rates predictable? If so, which variables help forecast them, individually as well as jointly? Since Fama and French (1988) documented that the price-dividend ratio predicts realized returns, many papers followed and addressed these important questions in various contexts. Subsequent papers in the literature can be loosely categorized into two groups. On the one hand, researchers have introduced more sophisticated statistical tests and new estimation methods to complement the OLS regression results. Some argued that careful statistical analysis provides little support for the predictability of returns by the price-dividend ratio (Nelson and Kim, 1993; Stambaugh, 1999; Valkanov, 2003; Goyal and Welch, 2003; Ang and Bekaert, 2007). Others have introduced more powerful ways to test for return predictability, for example, by jointly looking at returns and dividends (Cochrane, 2008a), by using latent-variables approach (Pastor and Stambaugh,

*University of Chicago Booth School of Business; email: jhe.yun@chicagobooth.edu. I would like to thank the members of my dissertation committee John Cochrane, Ralph Koijen, Lubos Pastor, and Pietro Veronesi for their support and comments. I also thank Timothy Dore, Marina Nisssten, Savina Rizova, Seung Yae, and brownbag participants at the University of Chicago Booth School of Business for insightful discussions and suggestions.
†First version: May 25, 2010
2009; Binsbergen and Koijen, 2010; Rytchkov, 2008; Lacerda and Santa-Clara, 2010), by accounting for structural shifts in the mean of the price-dividend ratio (Lettau and Van Nieuwerburgh, 2008), and by imposing weak restrictions on the signs of coefficients and return forecasts (Campbell and Thompson, 2008). These papers suggest that it is important to go beyond simple predictive regressions as they can be misleading.

On the other hand, another group of researchers have studied predictive variables other than the price-dividend ratio and found that returns on the aggregate stock market are strongly predictable at various horizons. For example, Rozek (1984) and Campbell and Shiller (1988a, 1988b) studied whether various valuation ratios predict subsequent returns. Other papers reported that yields on short and long-term treasury and corporate bonds can forecast stock returns (Fama and Schwert, 1977; Keim and Stambaugh, 1986; Fama and French, 1989). More recently, several papers introduced new variables motivated from corporate payout and financing activity (Lamont, 1988; Baker and Wurgler, 2000) and the level of consumption relative to wealth (Lettau and Lubikson, 2001), which is shown to predict returns at “business cycle” frequencies. Kelly and Pruitt (2011a) use information extracted from the cross-section of price-dividend ratios. See Koijen and Van Nieuwerburgh (2011) for a literature review on return and dividend growth predictability focusing on recent work.

These empirical findings, taken together, have yet to provide conclusive answers to the previous questions. There are a large number of predictive variables that have been introduced in various papers. There is no clear guidance on which variables to include when forecasting returns, whether to add one predictor given another, and whether predictors are significant under various statistical tests. My paper attempts to bridge the gap between the two branches of work and to answer these questions with a present-value approach that accounts for additional predictive variables. I augment a present-value approach, which improves upon predictive regressions, by using information contained in various predictive variables, which can potentially help predict returns and dividend growth rates along with the price-dividend ratio. I treat conditional expected returns and expected dividend growth rates as latent variables within a present-value model (Campbell and Shiller, 1988) and use a Kalman filter to filter them out, which are shown to be strong predictors of realized returns and realized dividend growth rates, respectively.

I use the approach to test for return and dividend growth rate predictability, and to test for statistical significance of each predictive variable within a present-value model, given the whole history of the price-dividend ratio and dividend growth rates. The present-value relationship implies

\[ p_{dt} = \frac{\kappa}{1 - \rho} + E_t \left[ \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \right] - E_t \left[ \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \right], \]
i.e. the price-dividend ratio varies due to variations in either expected returns or expected dividend growth rates, or both\(^1\). My approach explicitly models time-varying, and possibly correlated, expected returns and expected dividend growth rates, and uses information contained in additional predictors to filter them out from the price-dividend ratio. If expected returns and expected dividend growth rates are positively correlated, shocks to expected returns and expected dividend growth rates largely offset each other. Hence, the price-dividend ratio alone is not sufficient and additional predictors can provide useful information. The approach presents more powerful tests for both return and dividend growth rate predictability, and statistical significance of each predictive variable.\(^2\)

I find that both returns and dividend growth rates are strongly predictable using my present-value model. Compared to the standard OLS regressions with the price-dividend ratio, the return R-squared value increases from 9.3% to 16.5% and the dividend growth rate R-squared value increases from 0.9% to 23.4%. The following predictive variables introduced in the literature significantly improve the return and dividend growth rate forecasts of the present-value model: *Book-to-Market Ratio*, *Consumption-Wealth-Income Ratio*, and *Stock Variance*. Both expected returns and expected dividend growth rates are time-varying and correlated with each other, but expected returns are more persistent. In this case, predictable variations in returns and dividends are hard to detect statistically using the price-dividend ratio alone. The present-value model can predict “business cycle” frequency variations in both expected returns and expected dividend growth rates using information contained in additional predictors. Hence, the model delivers substantially higher R-squared values compared to the present-value model without an additional predictor, which produce R-squared values of 8.8% for returns and 13.6% for dividend growth rates. (Binsbergen and Koijen, 2010).

My paper is closely related to recent papers in the return predictability literature. Recently, many papers used the present-value relationship to jointly study return and dividend growth predictability. For example, Cochrane (2008a) introduces a joint test of return and dividend growth predictability to argue that the lack of dividend growth predictability gives stronger evidence than the presence of return predictability. Similarly, Lewellen (2004) improves the test of return predictability by using knowledge of the price-dividend ratio’s autocorrelation. A latent-variables approach within a present-value framework has been used successfully in Binsbergen and Koijen (2010) to estimate the expected returns and expected dividend growth rates. The authors find that both returns and dividend growth rates are predictable and persistent. My present-value model uses information contained in additional predictors to better filter out expected returns and expected dividend growth rates, both of which are

\(^1\)See Appendix A for derivation.

\(^2\)Another benefit of the present-value approach is that it is more robust to structural breaks in the means of expected returns and expected dividend growth rates than standard predictive regressions. See Rytchkov (2008) for more discussion.
allowed to be time-varying, from the price-dividend ratio\textsuperscript{3}. Lettau and Ludvigson (2005) use a related consumption-based present-value relation to show that changing forecasts of dividend growth makes it hard for the price-dividend ratio to uncover such variation. They find that dividend forecasts covary with changing forecasts of excess stock returns over business cycle frequencies. My work confirms their findings with a broader set of predictive variables. This paper can also be understood relative to Goyal and Welch (2003). Instead of standard predictive regressions, I use a present-value approach to better test the statistical significance of each predictive variable given the price-dividend ratio, and show that returns are predictable both in-sample and out-of-sample and that some variables indeed help predict realized returns.

I proceed as follows: In Section 2, I introduce the linearized present-value model. In section 3, I explain the data used in this paper, the state-space representation of the model, and the estimation procedure. In section 4, I present the main empirical results on return and dividend predictability. I compare the results to predictive regressions, report hypothesis testing results, and study long-run predictability. I also extend the present-value model to simultaneously account for multiple instruments. I conclude in Section 5.

2 Present-Value Model

I assume that both expected returns and expected dividend growth rates are latent variables. Let \( r_{t+1} \) denote the log return on the aggregate stock market:

\[
r_{t+1} \equiv \log \left( \frac{P_{t+1} + D_{t+1}}{P_t} \right)
\]

Let \( pd_t \) denote the log price-dividend ratio of the aggregate stock market:

\[
pd_t = \log(PD_t) \equiv \log \left( \frac{P_t}{D_t} \right)
\]

and let \( \Delta d_{t+1} \) denote the aggregate log dividend growth rate:

\[
\Delta d_{t+1} \equiv \log \left( \frac{D_{t+1}}{D_t} \right)
\]

I use annual variables to avoid seasonality issues in dividends. In addition to these variables, there is a predictive variable \( (z_t) \) that is potentially correlated

\textsuperscript{3} Others noted that the price-dividend ratio is a noisy proxy for expected returns when expected dividend growth rates also move the price-dividend ratio, and vice versa. Hence, additional predictive variables can help better filter out expected returns along with the price-dividend ratio, see Fama and French (1988), Menzly, Santos and Veronesi (2004), and Goetzmann and Jorion (1995). Many documented evidence of time-varying expected dividend growth rates, see Binsbergen and Koijen (2010), Lettau and Ludvigson (2005), Lacerda and Santa-Clara (2010).
with either or both expected returns and expected dividend growth rates. I model expected returns ($\mu_t$), expected dividend growth rates ($g_t$), and a predictive variable ($z_t$) as a first-order vector autoregressive (VAR) process:

\[
\begin{align*}
\mu_{t+1} &= \delta_0 + \delta_1(\mu_t - \delta_0) + \delta_2(z_t - \xi_0) + \epsilon_{t+1}^\mu \\
g_{t+1} &= \gamma_0 + \gamma_1(g_t - \gamma_0) + \gamma_2(z_t - \xi_0) + \epsilon_{t+1}^g \\
z_{t+1} &= \xi_0 + \xi_1(z_t - \xi_0) + \zeta_{t+1}
\end{align*}
\]  

(1)

where

\[
\begin{align*}
\mu_t &\equiv E_t[r_{t+1}] \\
g_t &\equiv E_t[\Delta d_{t+1}]
\end{align*}
\]

and $\delta_0$, $\gamma_0$, and $\xi_0$ denote the unconditional means of $\mu_t$, $g_t$, and $z_t$, respectively. The realized dividend growth rate is modeled as:

$$\Delta d_{t+1} = g_t + \epsilon_{t+1}^D$$

Note that both expected returns ($\mu_t$) and expected dividend growth ($g_t$) rates respond to the lag of a predictive variable ($z_t$). I assume a first-order autoregressive (AR) process for a predictive variable. Now I can express the log-linearized return as:

$$r_{t+1} \approx \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t$$

(2)

with $\tilde{pd} = E[\text{pd}_t]$, $\kappa = \log(1 + \exp(\tilde{pd})) - \rho \tilde{pd}$ and $\rho = \frac{\exp(\tilde{pd})}{1 + \exp(\tilde{pd})}$, as in Campbell and Shiller (1988). See Binsbergen and Koijen (2010) for a similar approach without an additional predictive variable.\(^4\) Iterating the above equations, we get:

$$pd_t \approx \frac{\kappa}{1 - \rho} + \sum_{j=1}^\infty \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^\infty \rho^{j-1} r_{t+j}$$

Taking the expectation conditional on time-$t$ information set and then using the VAR(1) assumptions, it follows that:

\[
\begin{align*}
pd_t &= \frac{\kappa}{1 - \rho} + E_t \left[ \sum_{j=1}^\infty \rho^{j-1} \Delta d_{t+j} \right] - E_t \left[ \sum_{j=1}^\infty \rho^{j-1} r_{t+j} \right] \\
&= A_0 + [A_1(g_t - \gamma_0) + A_2(z_t - \xi_0)] + [A_3(\mu_t - \delta_0) + A_4(z_t - \xi_0)]
\end{align*}
\]

\(^4\)On the other hand, Binsbergen and Koijen (2011) develop a tractable exactly solved present-value model and estimate it without approximation error. They show that the results are robust to non-linearities.
with,

\[
A_0 = \frac{\kappa + \gamma_0 - \delta_0}{1 - \rho},
\]

\[
A_1 = \frac{1}{1 - \rho\gamma_1}.
\]

\[
A_2 = \frac{\gamma_2}{\gamma_1 - \xi_1} \left( \frac{1}{1 - \rho\gamma_1} - \frac{1}{1 - \rho\xi_1} \right),
\]

\[
A_3 = -\frac{1}{1 - \rho\delta_1},
\]

\[
A_4 = -\frac{\delta_2}{\delta_1 - \xi_1} \left( \frac{1}{1 - \rho\delta_1} - \frac{1}{1 - \rho\xi_1} \right),
\]

see Appendix A. The log price-dividend ratio \((p/d_t)\) is linear in the expected return \((\mu_1)\), the expected dividend growth rate \((g_t)\), and a predictive variable \((z_t)\). The above expression decomposes the price-dividend ratio into two components, one for expectations of future returns and another for expectations of future dividend growth rates. As the predictive variable drives both expected returns and expected dividend growth rates, its current value \((z_t)\) provides additional information on expected returns and expected dividend growth rates. The loadings of these terms depend on the relative persistence of these variables \((\delta_1, \gamma_1, \xi_1)\). The four shocks in the model, which are shocks to expected dividend growth rates \((\epsilon_{t+1}^g)\), shocks to expected returns \((\epsilon_{t+1}^\mu)\), realized dividend growth shocks \((\epsilon_{t+1}^D)\), and shocks to the predictive variable \((\epsilon_{t+1}^z)\), are mean-zero and have the following covariance matrix:

\[
\Sigma \equiv \text{var} \left( \begin{pmatrix} \epsilon_{t+1}^g \\ \epsilon_{t+1}^\mu \\ \epsilon_{t+1}^D \\ \epsilon_{t+1}^z \end{pmatrix} \right) = \begin{bmatrix} \sigma_g^2 & \sigma_{g\mu} & \sigma_{gD} & \sigma_{gz} \\ \sigma_{g\mu} & \sigma_{\mu}^2 & \sigma_{\muD} & \sigma_{\muz} \\ \sigma_{gD} & \sigma_{\muD} & \sigma_{D}^2 & \sigma_{Dz} \\ \sigma_{gz} & \sigma_{\muz} & \sigma_{Dz} & \sigma_{z}^2 \end{bmatrix}
\]

The shocks are independent and identically distributed over time. In the maximum likelihood estimation procedure, I further assume that the shocks follow a multivariate normal distribution.

Compared to the predictive regressions that include only the current price-dividend ratio and a predictive variable to predict future returns and dividend growth rates, my approach aggregates the whole history of price-dividend ratios, dividend growth rates and a predictive variable to estimate expected returns and expected dividend growth rates. As shown in Cochrane (2008b), and Binsbergen and Koijen (2010), this introduces moving-average terms of price-dividend ratios, dividend growth rates and a predictive variable into the predictive regressions. I find that the moving-average terms are important in predicting future returns and, particularly, in predicting future dividend growth rates.

The present-value model introduced here extends present-value models used in the recent return predictability literature. As shown in the above expression
for the price-dividend ratio (3), my approach explicitly takes into account that
the price-dividend ratio moves due to both expected returns and expected divi-
dend growth rates and that there is an additional variable that is correlated with
either or both of them. This enables me to include various predictive variables
that have been introduced in the literature and test their significance relative to
the price-dividend ratio within a present-value framework. The framework al-
 lows me to account for each predictive variable’s power in predicting both future
returns and dividend growth rates and filter out noise from the price-dividend
ratio. By looking returns and dividend growth rates jointly, I construct more ef-
ficient estimators from information contained in each predictive variable. Later,
I expand the model to include multiple predictive variables and report its esti-
mation results. See Section 4.6.

3 Data and Estimation

3.1 Data

I use the with-dividend and without-dividend monthly returns on the value-
weighted portfolio of all NYSE, AMEX, and NASDAQ stocks for the period
1945–2010 from the Center for Research in Security Prices (CRSP). I use these
data to construct my annual data for aggregate dividends and prices. Following
Binsbergen and Koijen (2010), I use dividends reinvested in 30-day T-bills and
compute the corresponding series for dividend growth rates, the price-dividend
ratio, and returns. They compare the results with the dividends reinvested in
30-day T-bills to the ones with the market-reinvested dividends. Data on the
30-day T-bill rate is also obtained from CRSP. Annual dividend growth is much
less volatile than typically-used market-invested dividend growth with an an-
nual unconditional volatility of 6.9% versus a volatility of 13.2%.

For the predictive variables, I use the same data as in Goyal and Welch
(2007) updated through 2010. The risk-free rate (rfree) is the T-Bill rate. The
Stock Variance (svvar) is computed as sum of squared daily returns on the
S&P 500. The Book-to-Market Ratio (b/m) is the ratio of book value to
market value for the Dow Jones Industrial Average. Book values are from Value
Line’s Website. From March to December every year, the ratio is computed by
dividing book value at the end of the previous year by the price at the end of
the current month. For January and February, I divide book value at the end
of two years ago by the price at the end of the current month. The Net Equity
Expansion (ntis) is the ratio of 12-month moving sums of net issues by the
NYSE-listed stocks divided by the total end-of-year market capitalization of all
NYSE stocks. The Percent Equity Issuing (eqis) is the ratio of equity issuing
activity as a fraction of total issuing activity. This is the variable proposed
by Baker and Wurgler (2000). The Treasury-Bill rates (tbi) are the 3-month
Treasury Bill: Secondary Market Rate from the economic research database at
the Federal Reserve Bank at St. Louis. The Long Term Yield (lty), the Long
*Term Rate of Returns* (itr), and the *Corporate Bond Returns* (corpr) data are from Ibbotson’s *Stocks, Bonds, Bills and Inflation Yearbook*. The *Corporate Bond Yields* on AAA and BAA-rated bonds (AAA and BAA respectively) are from FRED. *Inflation* (infl) is the *Consumer Price Index* (All Urban Consumers) from the Bureau of Labor Statistics. The *Investment-to-Capital ratio* (ik) is the ratio of aggregate (private non-residential fixed) investment to aggregate capital for the whole economy. The *Consumption, Wealth, Income ratio* (cay) is from Lettau and Ludvigson (2001) where they estimate the following equation:

\[
e_t = \alpha + \beta_a a_t + \beta_y y_t + \sum_{i=-k}^{k} b_{a,i} \Delta a_{t-i} + \sum_{i=-k}^{k} b_{y,i} \Delta y_{t-i} + \epsilon_t
\]

where \( c \) is the aggregate consumption, \( a \) is the aggregate wealth, and \( y \) is the aggregate income. Using estimated coefficients, they form \( cay \) as deviations from the shared trend in consumption, labor income, and assets:

\[
cay \equiv \hat{c}ay_t = c_{n,t} - \hat{\beta}_a a_t - \hat{\beta}_y y_t
\]

where "hats" denote estimated parameters. I report the summary statistics in Table 1.

3.2 State-space representation

The model has three state variables: expected returns \((\mu_t)\), expected dividend growth rates \((g_t)\), and a predictive variable \((z_t)\). I assume that they follow a first-order VAR process. Define de-meaned state variables:

\[
\begin{align*}
\hat{\mu}_t &= \mu_t - \delta_0, \\
\hat{g}_t &= g_t - \gamma_0.
\end{align*}
\]

The model has two transition equations:

\[
\begin{align*}
\hat{g}_{t+1} &= \gamma_1 \hat{g}_t + \gamma_2 (z_t - \xi_0) + \epsilon_{\hat{g},t+1}, \\
\hat{\mu}_{t+1} &= \delta_1 \hat{\mu}_t + \delta_2 (z_t - \xi_0) + \epsilon_{\hat{\mu},t+1},
\end{align*}
\]

and three measurement equations:

\[
\begin{align*}
\Delta d_{t+1} &= \gamma_0 + \hat{g}_t + \epsilon_{\Delta d,t+1}, \\
pd_t &= A_0 + A_1 \hat{g}_t + A_3 \hat{\mu}_t + (A_2 + A_4)(z_t - \xi_0), \\
z_{t+1} &= \xi_0 + \xi_1 (z_t - \xi_0) + \epsilon_{z,t+1},
\end{align*}
\]

where \( A_0, A_1, A_2, \) and \( A_3 \) are previously defined. As the second measurement equation has no error term, I can substitute the equation for the log price-dividend ratio \((pd_t)\) into the transition equation for de-meaned expected returns \((\mu_t)\). This yields the following system with one transition and three
measurement equations:

\[
\begin{align*}
\hat{g}_{t+1} &= \gamma_1 \hat{g}_t + \gamma_2 (z_t - \xi_0) + \epsilon_t^g, \\
\Delta d_{t+1} &= \gamma_0 + \hat{g}_t + \epsilon_{t+1}^D, \\
pd_{t+1} &= (1 - \delta_1) A_0 + \delta_1 pd_t + A_1 (\gamma_1 - \delta_1) \hat{g}_t, \\
&+ (\gamma_2 A_1 + \delta_2 A_3 + (\xi_1 - \delta_1)(A_2 + A_4))(z_t - \xi_0), \\
&+ A_3 \epsilon_{t+1} + A_4 \epsilon_{t+1}^g + (A_2 + A_4) \epsilon_t^2, \\
z_{t+1} &= \xi_0 + \xi_1 (z_t - \xi_0) + \epsilon_{t+1}^z.
\end{align*}
\]

As the price-dividend ratio is linear in expected returns, expected dividend growth rates, and a predictive variable, I can attribute innovations in the price-dividend ratio to innovations in expected returns, expected dividend growth rates, or a predictive variable. Therefore, I can recover the full time-series of expected returns and expected dividend growth rates. Note that there is no measurement equation for returns. From the present-value relationship, the measurement equation for dividend growth rates and the price-dividend ratio implies the measurement equation for returns\(^5\). As all equations are linear, I can compute the likelihood of the model using a Kalman filter (Hamilton, 1994). I then use conditional maximum-likelihood estimation (MLE) to estimate the vector of model parameters:

\[
\Phi = \{\gamma_0, \delta_0, \xi_0, \gamma_1, \delta_1, \xi_1, \gamma_2, \delta_2, \sigma_g, \sigma_{\mu}, \sigma_D, \sigma_z, \sigma_{gD}, \sigma_{gD}, \sigma_{\muD}, \sigma_{\muZ}, \sigma_{DZ}\}
\]

I describe the Kalman Filter and the estimation procedure in Appendix B. I maximize the likelihood using simulated annealing. The simulated annealing minimization algorithm is designed to find the global minimum (Goffe, Ferrier and Rogers, 1994). The model is estimated using annual data.

In the state-space model, I compute the R-squared values for returns and dividend growth rates as:

\[
\begin{align*}
R^2_{Ret} &= 1 - \frac{\text{var}(r_{t+1} - \hat{\mu}_t)}{\text{var}(r_t)}, \\
R^2_{Div} &= 1 - \frac{\text{var}(\Delta d_{t+1} - \hat{\mu}_t)}{\text{var}(\Delta d_{t+1})}
\end{align*}
\]

where \(\text{var}\) is the sample variance, \(\hat{\mu}_t\) is the filtered series for expected returns (\(\mu_t\)), and \(\hat{g}_t\) is the filtered series for expected dividend growth rates (\(g_t\)). Using this definition, I can compare the results from the present-value approach to the standard predictive regression results.

---

\(^5\)The implied returns are very close to the actual returns from CRSP. As discussed in Bisigtenen and Koijen (2010), the difference between the two accounts for less than 1% of the total variation in returns. I can easily use the other two variables to estimate the model. The results are very similar.
In my model, all but one of the parameters in the covariance matrix are identified. There are different ways to impose an identifying assumption. Rytchkov (2008) works with a set of parameters that yields the same likelihood value. His analysis is based on a set of parameters that attain the maximum likelihood. Cochrane (2008b) works with the observable shocks. In his approach, “structural” shocks are given by the linear combinations of them. Following Binsbergen and Koijen (2010), I normalize by setting the correlation between realized dividend growth shocks ($\epsilon_{t+1}^D$) and expected dividend growth shocks ($\epsilon_{t+1}'$) to zero.

4 Empirical Results

4.1 Return and dividend growth rate predictability

I study return and dividend growth rate predictability using my present-value model with each predictive variable. I report the estimates of the model parameters ($\Phi$) in Table 2. The present-value approach is applied to each predictive variable. The estimated models deliver significantly higher R-squared values for both returns and dividend growth rates. The return R-squared value goes up to 16.5% with eqis and the dividend R-squared value goes up to 23.4% with ltr. As a benchmark, I run the following predictive OLS regressions:

\[
\begin{align*}
    r_{t+1} & = \alpha_r + \beta_r pd_t + \gamma_r z_t + \epsilon_{t+1} \\
    \Delta d_{t+1} & = \alpha_d + \beta_d pd_t + \gamma_d z_t + \epsilon_{t+1}
\end{align*}
\]

I report the predictive regression results in Table 3. Compared to a predictive regression with the log price-dividend ratio as a sole predictor, the present-value models with various predictive variables nearly double the return R-squared values. The dividend R-squared value increases even more drastically from 0.9% to 23.4%. The present-value models augmented with bm, cay, infi, eqis, ltr, and corpr strongly forecast both returns and dividend growth rates in terms of the R-squared values. The return R-squared values are similar to those from the corresponding predictive regressions. However, the dividend R-squared values are significantly higher when I account for the predictive variables using my present-value approach. Many predictive variables indeed affect the joint dynamics of expected returns and expected dividend growth rates. The benchmark present-value model without any predictive variable results in the R-squared values of 8.8% for returns and 13.6% for dividend growth rates. See Binsbergen and Koijen (2010). Hence, the present-value model can be substantially improved by adding information in the additional predictive variables. Additional predictive variables are useful, even after we account for the lags of the price-dividend ratio and dividend growth rates within a present-value framework.

I compare the persistence of expected returns and expected dividend growth rates, see Table 4. The predictive regressions of returns and dividend growth rates on the price-dividend ratio assume that they are all equally persistent. In
my present-value model, however, expected dividend growth rates are not as persistent as expected returns. Additional predictive variables generally do not affect the persistence of expected returns. Only with svar, the expected returns are more persistent. Expected dividend growth rates are much less persistent than expected returns, consistent withBinsbergen and Koijen (2010). Next, I look at parameter estimates more carefully. With most predictive variables, the present-value model implies more volatile expected returns ($\sigma_e$). The volatility more than doubles with cay, eqis, and ltr. See Figure 1 and 2 for the plots of the filtered expected returns. The predictive variables mostly help forecast “business cycle” frequency variations in returns while not affecting the general trend identified by the price-dividend ratio. This makes sense as these variables are less persistent than the price-dividend ratio. Table 5 shows that only bm is as persistent as the price-dividend ratio while cay, eqis, ltr, and corpr are all much less persistent. On the other hand, only ltr and corpr increase the volatility of expected dividend growth rates from 6.5% to about 10%. bm hardly affects the volatility of expected dividend growth rates yet helps forecast realized dividend growth rates.

The correlation between shocks to expected returns and expected dividend growth rates ($\rho_{gy}$) is similar across different predictive variables. With most additional predictors, the model implies shocks to expected returns and expected dividend growth rates that are positively correlated, consistent withBinsbergen and Koijen (2010), Menzly, Santos, and Veronesi (2004), Lettau and Ludvigson (2005), and Rytchkov (2008). The present-value models with ltr and corpr imply especially higher correlations. As noted in Lettau and Ludvigson (2005), positively correlated expected returns and expected dividend growth rates make it harder to filter them out from the price-dividend ratio alone, and hence, additional predictors can provide useful information. I also report the correlation between innovations to expected returns and unexpected returns ($\rho_{ur}$). This is the correlation extensively studied in Pastor and Stambaugh (2009). As argued in their paper, I find that with every predictive variable, the correlation is negative. That is, unexpected increases in expected future returns (or discount rates) is accompanied by unexpected negative returns. Note that the correlation between expected returns and realized dividend shocks ($\rho_{\mu_D}$) are very high for the predictive variables that successfully increase return R-squared values. Hence, additional predictability in returns comes from being able to extract valuable information from unexpected dividend growth rate shocks. Interestingly, with the two most helpful variables, cay and eqis, the correlations ($\rho_{\mu_D}$) have opposite signs, -74% and 87%. Hence, with cay, a positive unexpected dividend shock implies a lower expected return next year, while with eqis, it implies a lower expected return. The results are surprising as these two predictive variables imply substantially different expected return and expected dividend growth rate series but still generate high return R-squared values. They seem to provide useful information independent of each other. This motivates me to look at a present-value approach with multiple predictive variables. See Section 4.6.1.
Using the parameter estimates, I look at the implied price-dividend ratio expression (3):

\begin{align*}
  \text{pd}_t &= A_0 + [A_1(g_t - \gamma_0) + A_2(z_t - \xi_0)] \\
  &\quad + [A_3(\mu_t - \delta_0) + A_4(z_t - \xi_0)]
\end{align*}

where the first term in the bracket corresponds to expected dividend growth rate variation and the second term corresponds to expected return variation. This enables me to look at how additional predictive variables help decompose the price-dividend ratio into expected returns and expected dividend growth rates. Table 6 reports the implied model parameters: $A_1$, $A_2$, $A_3$ and $A_4$. First, note that for $\text{bm}$, $\text{ntis}$ and $\text{rfree}$, $A_2$ and $A_4$ are of same signs and similar in magnitudes. Hence, they average out in the price-dividend ratio. Hence, variations in these predictive variables mostly capture time variations in the compositions of the log price-dividend ratio. That is, these predictive variables help decompose variation in the price-dividend ratio into expected return and expected dividend growth rate variation. On the other hand, for $\text{cay}$, $\text{infl}$, $\text{eqis}$ and $\text{svar}$, either $A_2$ and $A_4$ have different signs or $A_2$ is significantly larger than $A_4$ in magnitude. Note that these cases correspond to the ones with high return R-squared values. Large $A_3$ values suggest that the predictive variables substantially help predict future returns, which naturally leads to higher return R-squared values. These variables, however, do not help much in forecasting future dividend growth rates and, hence, naturally affect the price-dividend ratio. I revisit the topic in Section 4.4 when I look at the long-run predictability of returns and dividend growth rates.

### 4.2 Hypothesis testing

I run the likelihood ratio (LR) tests to test statistical significance of my findings. My likelihood-based estimation leads to a straightforward hypothesis testing using the likelihood ratio (LR) tests. Let the log-likelihood from the unconstrained model by $L_u$. Let $L_c$ denotes the log-likelihood from the constrained model with the appropriate constraint for each null hypothesis. Then, the likelihood ratio test statistic is given by:

\[ LR = 2(L_u - L_c) \]  

which asymptotically follows the $\chi^2$-distribution with the degrees of freedom equal to the number of constrained parameters.

First, I test the significance of including each predictive variable. The null hypothesis is given by:

\[ H_0 : \delta_2 = \gamma_2 = \rho_{gz} = \rho_{uz} = \rho_{Dz} = 0 \]

whose LR statistic follows a $\chi^2$-distribution. Under the null hypothesis, an additional predictive variable does not help decompose variation in the price-
dividend ratio into expected return and expected dividend growth rate variation.

Second, I test for the lack of return predictability. The null hypothesis is:

\[ H_0 : \delta_1 = \delta_2 = \rho_{g_t} = \rho_{\mu D} = \rho_{\mu z} = 0 \]

whose LR statistic follows a \( \chi^2 \) distribution. Under the null hypothesis, all variation in the price-dividend ratio comes from expected dividend growth rate variation. In this case, I can uncover expected dividend growth rates through an OLS regression of dividend growth rates on the lagged price-dividend ratio. (Binsberger and Kojjen, 2010)

Last, I test for the lack of dividend growth rate predictability. The null hypothesis is:

\[ H_0 : \gamma_1 = \gamma_2 = \rho_{g_t} = \sigma_{g z} = 0 \]

whose LR statistic follows a \( \chi^2 \) distribution. As before, if dividend growth rates are unpredictable, I can uncover expected returns through an OLS regression of returns on the lagged price-dividend ratio, a standard predictive.

The likelihood ratio (LR) test results are reported in Table 7. The tests support including \( \text{bm, BAA, cay, and svar} \) in the present-value model among the list of return-predicting variables. \( \text{cay} \) increases return R-squared value and the LR test confirms the importance of having it in the present-value model. On the other hand, despite the high R-squared value, \( \text{eqis} \) does not show up statistically significant in the LR tests. \( \text{bm, BAA and svar} \) are statistically significant in joint predictability of returns and dividend growth rates, despite their mediocre return R-squared values. This suggests that high R-squared values should be taken with a grain of salt when evaluating how much a variable predicts returns and/or dividend growth rates. R-squared values do not necessarily coincide with the LR test results, as shown in Binsberger and Kojjen (2010), and Harvey (1989). The tests suggest that returns are predictable at a statistically significant level by every predictive variable when combined with the price-dividend ratio within the present-value framework. On the other hand, the following list of predictive variables within a present-value model help forecast dividend growth rates: \( \text{cay, ltr, corpr, svar} \).

### 4.3 Wold decomposition

I present the Wold decomposition of the state-space model. I plot the coefficients of the implied VAR-MA representation of my state-space model. That is, I compute the implied coefficients of the following expressions:

\[
\begin{align*}
  r_t &= a_0^r + \sum_{i=0}^{\infty} a_{1,r}^d d_{t-i-1} + \sum_{i=0}^{\infty} a_{2,r}^d \Delta d_{t-i-1} + \sum_{i=0}^{\infty} a_{3,r}^d z_{t-i-1} + \epsilon_t^d \\
  \Delta d_t &= a_0^d + \sum_{i=0}^{\infty} a_{1,d}^d d_{t-i-1} + \sum_{i=0}^{\infty} a_{2,d}^d \Delta d_{t-i-1} + \sum_{i=0}^{\infty} a_{3,d}^d z_{t-i-1} + \epsilon_t^d
\end{align*}
\]

13
See Appendix C for derivation. I plot the coefficients on the lags of each predictive variable (\(\{a_{it}^2\}_{j=1}^{\infty}\) and \(\{a_{it}^2\}_{j=1}^{\infty}\)). First, I look at the return expression with \textbf{cay} and \textbf{eqis} in Figure 5. Note that even though the return R-squared values are comparable, the present-value model and the predictive regressions impose substantially different dynamics on how \textbf{cay} predicts returns. The coefficient of \textbf{cay} on returns is 1.85 in the predictive regression and around -1.8 in the implied VAR-MA representation. The difference gets smaller when we include more lags in the predictive regression. Including further lags significantly affects the relationship between expected returns and \textbf{cay}. This also happens with \textbf{eqis}. Next, I look at \textbf{rfree}, \textbf{ltr}, and \textbf{corpr} in Figure 6. Note that I need to include more lags to appropriately capture the joint dynamics of expected returns and these predictive variables. Figure 7 shows that dividend growth rates also require many lags of these predictive variables. It is difficult to include as many lags in predictive regressions. With additional predictive variables, the number of parameters that have to be estimated increases substantially when we add more lags. The present-value approach provides a parsimonious way to incorporate the information contained in the long lags of price-dividend ratios, dividend growth rates, and a predictive variable, without proportionally increasing the number of parameters to be estimated.

### 4.4 Long-run predictability

I look at long-run predictability of returns and dividend growth rates, comparing the long-run forecasts from the present-value model with different predictive variables. That is, I compute the following conditional expectations implied by the model:

\[
E_t \left[ \sum_{j=1}^{\infty} \rho^{-1} r_{t+j} \right], E_t \left[ \sum_{j=1}^{\infty} \rho^{-1} \Delta d_{t+j} \right]
\]  

(9)

As a benchmark, I compute the long-run forecasts using a first-order VAR with the log price-dividend ratio as a sole predictor.

First, I look at long-run return predictability. Figure 8 shows that \textbf{cay} and \textbf{eqis} do not affect the long-run forecasts of the present-value model. They imply very similar long-run forecasts to those from the benchmark present-value model and the first-order VAR system. They strongly predict one-year returns but do not change long-run expected returns. This is possible if they have offsetting effects on longer-run returns \(r_{t+j}\). They instead signal changes in the term structure of risk premia \(E_t[r_{t+j}]\), see Cochrane (2011). On the other hand, \textbf{bm} and \textbf{corpr} do alter long-run expected returns, as shown in Figure 9. The long-run forecasts differ from the forecasts of the benchmark present-value model without any predictive variable. \textbf{corpr} implies more volatile long-run forecasts of returns and dividend growth rates, as discussed below.
Next, I look at long-run dividend growth rate predictability. Recall that the present-value identity (3) implies that long-run forecasts of returns and dividend growth rates add up to the price-dividend ratio. In contrast to long-run forecasts of returns, long-run forecasts of dividend growth rates implied by different predictive variables look significantly different. First, Figure 10 compares the long-run forecasts from the present-value models with cay and eqis to the benchmark model and the VAR system. The present-value models generally imply much less variation in the long-run expected dividend growth rates. Compared to huge swings in the price-dividend ratio, which is clearly reflected in the long-run forecasts of the first-order VAR, the long-run dividend growth rate forecasts of these present-value models do not vary as much and stay in a much tighter range. This shows that the present-value model not only dramatically increases the one-year ahead dividend R-squared value but also implies significantly different long-run forecasts. The significance of differences across various predictive variables arises from the fact that long-run forecasts of dividend growth rates do not vary as much as the long-run forecasts of returns. Hence, the differences show up more significantly. It can be seen that cay and eqis, which are very helpful in predicting one-year returns yet do not affect the long-run forecasts of returns, imply substantially different long-run forecasts of dividend growth rates. ltr and corpr imply much more volatile long-run forecasts (Figure 11). The general trend remains the same but corpr provides information on higher-frequency variations in the expectation of long-run dividend growth rates. tbl, AAA, lty and rfree affect the long-run forecasts in a similar way. In their cases, a factor related to credit markets seems to be generating a very long-run trend in dividend growth rates (Figure 12). The results suggest that long-run forecasts of dividend growth rates provide another useful diagnostic for variable selection. When evaluating a value of a predictive variable, we should look jointly at return and dividend predictability, and its implications on long-run forecasts. Simple measures, such as R-squared values on one-year returns and dividend growth rates, can miss important, and possibly long-run, interactions among returns, dividend growth rates, the price-dividend ratio and other predictive variables.

4.5 Variance decomposition

I supplement the previous section on long-run predictability by studying the variance decomposition of the log price-dividend ratio implied by the present-value model. Recall that in the model the log price-dividend ratio is given by:

$$pd_t = A_0 + [A_1(g_t - \gamma_0) + A_2(z_t - \xi_0)] + [A_3(\mu_t - \delta_0) + A_4(z_t - \xi_0)]$$

The first term in the bracket denotes expected return variation and the second term corresponds to variation from expected dividend growth rates. Then, the variance of the log price-dividend ratio is given by:

$$var(pd_t) = var_\mu + var_\delta + 2cov_{\mu,\delta}$$
That is, I decompose the variance of the price-dividend ratio into three components: the variance of expected returns, the variance of expected dividend growth rates, and the covariance between the two. I report the variance decomposition results in Table 8. Generally, even with the predictive variables within my present-value approach, most variation in the price-dividend ratio comes from expected return variation. Note that the variance of expected dividend growth rates never exceeds 6%. On the other hand, the variance of expected returns never falls below 90%. This is consistent with Cochrane (2011), who argues that additional predictive variables cannot shift variance attribution from returns to dividends as higher long-run dividend forecasts must be matched by a higher long-run return forecasts. There are still notable differences among predictive variables. With \textbf{AAA}, \textit{ly}, and \textit{rfree}, the covariances between expected returns and expected dividend growth rates are quite significant ranging from -12% to -29%. In such cases, variances of expected returns compensate by exceeding 100%.

4.6 Additional results

4.6.1 Multiple instrument system

I extend the present-value model to account for more than one instrument at a time. Previously, I noted that some predictive variables seem to add information not captured by the price-dividend ratio. I use the extended present-value model to value a predictive variable conditional on another variable. I look at the following system where \( z_t \) and \( w_t \) are two predictive variables.

\[
\begin{bmatrix}
\mu_{t+1} \\
g_{t+1} \\
z_{t+1} \\
w_{t+1}
\end{bmatrix} =
\begin{bmatrix}
\delta_0 \\
\gamma_0 \\
\xi_0 \\
\psi_0
\end{bmatrix} +
A
\begin{bmatrix}
\mu_t \\
g_t \\
z_t \\
w_t
\end{bmatrix} +
\epsilon_{\mu_{t+1}}
\]

I impose AR(1) processes for both predictive variables. As with the previous present-value model, I obtain the following transition equations:

\[
\begin{align*}
\hat{g}_{t+1} &= \gamma_1 \hat{g}_t + \gamma_2 (z_t - \xi_0) + \gamma_3 (w_t - \psi_0) + \epsilon_{\hat{g}_{t+1}} \\
\hat{\mu}_{t+1} &= \delta_1 \hat{\mu}_t + \delta_2 (z_t - \xi_0) + \delta_3 (w_t - \psi_0) + \epsilon_{\hat{\mu}_{t+1}}
\end{align*}
\]

and the following four measurement equations:

\[
\begin{align*}
\Delta d_{t+1} &= \gamma_0 + \hat{g}_t + \epsilon_d^D_{t+1} \\
pd_t &= A_0 + A_1 \hat{g}_t + A_3 \hat{\mu}_t + (A_2 + A_4)(z_t - \xi_0) + (B_1 + B_2)(w_t - \psi_0) \\
z_{t+1} &= \xi_0 + \xi_1 (z_t - \xi_0) + \epsilon_z^D_{t+1} \\
w_{t+1} &= \psi_0 + \psi_1 (w_t - \psi_0) + \epsilon_w^D_{t+1}
\end{align*}
\]

(10)
where $A_1$, $A_2$, $A_3$, and $A_4$ are same as before. See Equation (3). $B_1$ and $B_2$ are defined similarly as:

$$B_1 = \frac{\gamma_3}{\gamma_1 - \psi_1} \left( \frac{1}{1 - \rho \gamma_1} - \frac{1}{1 - \rho \psi_1} \right)$$

$$B_2 = -\frac{\delta_3}{\delta_1 - \psi_1} \left( \frac{1}{1 - \rho \delta_1} - \frac{1}{1 - \rho \psi_1} \right)$$

I again substitute out the equation for de-meaned expected returns ($\mu_t$) using the log price-dividend ratio ($pd_t$). I estimate the extended present-value model using the conditional maximum likelihood estimation.

Based on the above model, I run the following additional likelihood ratio (LR) tests to test statistical significance of predictive variables relative to each other. First, I test the significance of including both predictive variables in the present-value model. The null hypothesis is given by:

$$H_0 : \delta_2 = \gamma_2 = \rho_{gz} = \rho_{\mu z} = \rho_{D z} = \ldots$$

$$\delta_3 = \gamma_3 = \rho_{gw} = \rho_{\mu w} = \rho_{D w} = 0$$

whose LR statistic follows a $\chi^2_{10}$-distribution. Under the null hypothesis, additional predictors $z$ and $w$ do not help filter out expected returns and expected dividend growth rates.

Second, I test for the significance of one predictive variable given the other, i.e. significance of $z$ given $w$ and vice versa. The null hypothesis is given by:

$$H_0 : \delta_2 = \gamma_2 = \rho_{gz} = \rho_{\mu z} = \rho_{D z} = 0$$

or

$$H_0 : \delta_3 = \gamma_3 = \rho_{gw} = \rho_{\mu w} = \rho_{D w} = 0$$

whose LR statistic follows a $\chi^2_2$-distribution. Under each null hypothesis, a predictive variable does not help given the other predictive variable.

I estimate the present-value model with each pair of the predictive variables that show up significantly in a univariate setting. I report the R-squared values from the extended present-value model and the likelihood ratio (LR) tests in Table 9. I do not report the actual estimates for brevity. Note that R-squared values do not dramatically increase compared to the one-instrument present-value model. Generally, the present-value model with two predictors help better forecast dividend growth rates. For all pairs, they are jointly significant, which is expected from the univariate likelihood ratio tests. The relative significance results show that once we account for svar, including either bm or cay does not significantly improve the model. Other variables seem to provide information independent of each other, and hence, are significant even if we include them in pairs.
4.6.2 Out-of-sample predictability

Previous sections have shown that the returns are predictable within the present-value approach when combined with several predictive variables. To address the complaints of several papers that argue poor out-of-sample return predictability (e.g. Goyal and Welch, 2003), I now study how well the present-value approach performs out-of-sample. I run the analysis as follows: I use the data up to time-t to estimate the model with an appropriate instrument, and then forecast the next-year’s return ($r_{t+1}$) using the filtered expected returns ($\hat{\mu}_t$). Then, I iterate the process until the terminal year in my sample ($T$), 2010. I start with data up to 1974, predicting the realized return on 1975. This ensures that the initial sample contains enough observations to estimate the present-value model. I use the two metrics. First, I use the out-of-sample version of the R-squared value:

$$R^2_{OOS} = 1 - \frac{\text{var}(r_{t+1} - \hat{\mu}_t)}{\text{var}(r_t)}$$

where $\hat{\mu}_t$ is computed using data up to time $t$. I then compute the sample variances. Second, I use the cumulative sum of squared errors:

$$\text{cumSSE}_t = \sum_{s=0}^{s=t} (r_{s+1} - \hat{\mu}_s)$$

I compare the cumulative SSE’s for different models. The difference shows the relative performance of the present-value model over time.

Table 10 shows the out-of-sample predictability results for each predictive variable. I report both the predictive regression and the present-value model out-of-sample R-squared values for comparison. Note that bm, cay and infl generates positive out-of-sample R-squared values. For bm and cay, returns are significantly predictable out-of-sample. They outperform the corresponding predictive regressions as well as the price-dividend ratio regression. Note that the standard predictive regressions with all predictive variables except cay have poor out-of-sample predictability, consistent with Goyal and Welch (2003). Hence, out-of-sample predictability of bm and infl justifies the present-value approach. More importantly, the plots of differences in cumulative SSE show that the present-value models with various predictive variables, including those with negative out-of-sample R-squared values, perform better in the later part of the sample, i.e. with longer samples. See Figure 13 and 14. In fact, many predictive variables would have produced positive out-of-sample R-squared values had I used a larger initial sample. The relative performance of the present-value models shows how the present-value approach provides a more efficient estimator of realized returns than the corresponding predictive regressions.
5 Conclusion

I propose a present-value approach to study which variables help forecast returns and dividend growth rates along with the price-dividend ratio. Using information contained in additional predictive variables within a present-value framework, the approach produces better estimates of the expected returns and expected dividend growth rates, which strongly predicts realized returns and realized dividend growth rates, respectively. I document significantly higher R-squared values for both returns and dividend growth rates. From the likelihood ratio tests, I conclude that returns and dividend growth rates are predictable at a statistically significant level, and the following predictive variables help forecast them along with the price-dividend ratio: Consumption-Wealth-Income Ratio, equity issuing Activity Ratio, Long Term Rate of Returns. I also note that despite high return R-square values, most variables imply similar long-run return forecasts within the present-value model, and mostly forecast “business cycle” frequency variations in returns. The Wold decomposition shows that some predictors require many lags to accurately capture their effects on expected returns and expected dividend growth rates, which is how the present-value approach outperforms standard predictive regressions. Out-of-sample predictability test further supports the present-value approach.

Even though I apply the approach to the aggregate stock market, it has broader applications. We can use the cross-sectional data within this framework. In such a case, my approach can be even more useful as we can account for other variables that explain cross-sectional variations in expected returns and expected dividend growth rates. See Kelly and Pruitt (2011a) for a similar approach. The state-space model used in this paper can be extended to allow more general VAR specifications among state variables. For example, we can allow additional predictors $z_{t+1}$ to depend on past values of expected returns $(\mu_t)$ and/or expected dividend growth rates $(g_t)$. One can also allow more lags in the VAR system, capturing richer dynamics. For example, some predictors can predict returns and dividend growth rates in different horizons. This would shed more light on how different predictors work relative to each other and the price-dividend ratio. Another natural extension of the approach would be to incorporate all predictive variables at once. This would be a multivariate extension of my present-value approach. See Kelly and Pruitt (2011b) for a related approach, where they use many predictive variables to forecast a single time series.
References


Cochrane, John H., 2008(B), State-Space vs. VAR models for Stock Returns, Working Paper, University of Chicago

Cochrane, John H. 2011, Presidential Address: Discount Rates, University of Chicago


Kelly, Bryan, and Seth Pruitt, 2011a, Market Expectations in the Cross Section of Present Values, Working Paper, University of Chicago


Menzly, Lior, Tano Santos, and Pietro Veronesi, 2004, Understanding Pre-
dictability, *Journal of Political Economy* 112(1), 1-47


Rytchkov, Oleg, 2008, Filtering Out Expected Dividends and Expected Returns, Working Paper, University of Texas


Appendix A: Present-Value Model Derivation

I model the expected returns, expected dividend growth rates, and a predictive variable as:

\[
\begin{align*}
\mu_{t+1} &= \delta_0 + \delta_1(\mu_t - \delta_0) + \delta_2(z_t - \xi_0) + \epsilon^\mu_{t+1} \\
g_{t+1} &= \gamma_0 + \gamma_1(g_t - \gamma_0) + \gamma_2(z_t - \xi_0) + \epsilon^g_{t+1} \\
z_{t+1} &= \xi_0 + \xi_1(z_t - \xi_0) + \epsilon^z_{t+1}
\end{align*}
\]

where

\[
\begin{align*}
\mu_t &= E_t[r_{t+1}] \\
g_t &= E_t[\Delta d_{t+1}]
\end{align*}
\]

and

\[
\begin{align*}
r_{t+1} &= \log \left( \frac{P_{t+1} + D_{t+1}}{P_t} \right) \\
\Delta d_{t+1} &= \log \left( \frac{D_{t+1}}{D_t} \right) \\
pd_t &= \log(PD_t) \equiv \log \left( \frac{P_t}{D_t} \right)
\end{align*}
\]

Now, following Campbell and Shiller (1988), I obtain:

\[
\begin{align*}
r_{t+1} &= \log(1 + \exp(pd_{t+1})) + \Delta d_{t+1} - pd_t \\
&= \log(1 + \exp(\tilde{p}d)) + \frac{\exp(\tilde{p}d)}{1 + \exp(\tilde{p}d)}pd_{t+1} + \Delta d_{t+1} - pd_t \\
&= \kappa + \rho pd_{t+1} + \Delta d_{t+1} - pd_t
\end{align*}
\]

where \(\tilde{p}d = E[pd_t]\), \(\kappa = \log(1 + \exp(pd)) - \tilde{p}d\) and \(\rho = \frac{\exp(\tilde{p}d)}{1 + \exp(\tilde{p}d)}\). Now iterating this equation:

\[
\begin{align*}
pd_t &= \kappa + \rho pd_{t+1} + \Delta d_{t+1} - r_{t+1} \\
&= \kappa + \rho(\kappa + \rho pd_{t+2} + \Delta d_{t+2} - r_{t+2}) + \Delta d_{t+1} - r_{t+1} \\
&= \kappa + \rho \kappa + \rho^2 pd_{t+2} + \Delta d_{t+1} - r_{t+1} + \rho(\Delta d_{t+2} - r_{t+2}) \\
&= \ldots \\
&= \frac{\kappa}{1 - \rho} + \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}
\end{align*}
\]

Next, I take expectations on both sides conditional upon time-\(t\):

\[
\begin{align*}
pd_t &= \frac{\kappa}{1 - \rho} + E_t \left[ \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \right] - E_t \left[ \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \right]
\end{align*}
\]
Using the VAR(1) assumptions:

\[ pd_t = \frac{\kappa}{1 - \rho} + E_t \left[ \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \right] - E_t \left[ \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \right] \]

\[ = \frac{\kappa}{1 - \rho} + \sum_{j=0}^{\infty} \rho^j E_t [g_{t+j} - \mu_{t+j}] \]

\[ = \frac{\kappa}{1 - \rho} + \sum_{j=0}^{\infty} \rho^j E_t [\gamma_0 + \gamma_1 (g_{t+j-1} - \gamma_0) + \gamma_2 (z_{t+j-1} - \xi_0) - \mu_{t+j}] \]

\[ = \cdots \kappa \]

\[ = \frac{\kappa}{1 - \rho} \]

\[ + \sum_{j=0}^{\infty} \rho^j E_t [\gamma_0 + \gamma_1 (g_t - \gamma_0) + \gamma_2 (\xi_1^{j-1} + \xi_1^{j-2} + \cdots + \gamma_1^{j-1}) (z_t - \xi_0) - \mu_{t+j}] \]

\[ = \frac{\kappa}{1 - \rho} \]

\[ + \sum_{j=0}^{\infty} \rho^j E_t [\gamma_0 + \gamma_1 (g_t - \gamma_0) + \gamma_2 \left( \frac{\xi_1^j - \gamma_1^j}{\xi_1 - \gamma_1} \right) (z_t - \xi_0) - \mu_{t+j}] \]

\[\]

\[ = A_0 + [A_1 (g_t - \gamma_0) + A_2 (z_t - \xi_0)] + [A_3 (\mu_t - \delta_0) + A_4 (z_t - \xi_0)] \]

where

\[ A_0 = \frac{\kappa + \gamma_0 - \delta_0}{1 - \rho} \]

\[ A_1 = \frac{1}{1 - \rho \gamma_1} \]

\[ A_2 = \frac{\gamma_2}{\gamma_1 - \xi_1} \left( \frac{1}{1 - \rho \gamma_1} - \frac{1}{1 - \rho \xi_1} \right) \]

\[ A_3 = -\frac{1}{1 - \rho \delta_1} \]

\[ A_4 = -\frac{\delta_2}{\delta_1 - \xi_1} \left( \frac{1}{1 - \rho \delta_1} - \frac{1}{1 - \rho \xi_1} \right) \]

Therefore, I get the equation (3).
Appendix B: Kalman filter

I discuss the Kalman Filtering procedure of my model. First, express the model in a standard state-space form. Define an expanded state vector:

\[
X_t = \begin{bmatrix}
\hat{g}_{t-1} \\
\epsilon_t^p \\
\epsilon_t^q \\
\epsilon_t^r \\
\epsilon_t^z
\end{bmatrix}
\]

which satisfies:

\[
X_{t+1} = FX_t + Bu_{t+1} + \Gamma \epsilon_{t+1}^X
\]

with

\[
F = \begin{bmatrix}
\gamma_1 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

\[
B = \begin{bmatrix}
\gamma_2 \\
0 \\
0 \\
0
\end{bmatrix}
\]

\[
\Gamma = \begin{bmatrix}
0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

and where

\[
u_t = z_{t-1} - \xi_0
\]

\[
\epsilon_{t+1}^X = \begin{bmatrix}
\epsilon_t^p \\
\epsilon_t^q \\
\epsilon_t^r \\
\epsilon_t^z
\end{bmatrix}
\]

which I assume to be multivariate normal.

The measurement equation, which has the observables \(Y_t = (\Delta d_t, pd_t, z_t)\), is:

\[
Y_t = M_0 + M_1 Y_{t-1} + M_2 X_t
\]
with

\[
M_0 = \begin{bmatrix}
\gamma_0 \\
(1 - \delta_1)A_0 \\
0
\end{bmatrix}
\]

\[
M_1 = \begin{bmatrix}
0 & 0 & \delta_2A_2 + (\xi_1 - \delta_1)A_4 + \gamma_2A_1 + (\xi_1 - \delta_1)A_3 \\
0 & 0 & \xi_1
\end{bmatrix}
\]

\[
M_2 = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 \\
(\gamma_1 - \delta_1)A_1 & 0 & A_2 & A_3 + A_4 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

The Kalman procedure is given by:

\[
X_{0|0} = E[X_0] = 0_{6 \times 1}
\]

\[
P_{0|0} = E[X_tX_t']
\]

\[
X_{t|t-1} = FX_{t-1|t-1} + Bu_t
\]

\[
P_{t|t-1} = FP_{t-1|t-1}F' + \Gamma \Sigma \Gamma'
\]

\[
\eta_t = Y_t - M_0 - M_1Y_{t-1} - M_2X_{t|t-1}
\]

\[
S_t = M_2P_{t|t-1}M_2'
\]

\[
K_t = P_{t|t-1}M_2'\Sigma_t^{-1}
\]

\[
X_{t|t} = X_{t|t-1} + K_t\eta_t
\]

\[
P_{t|t} = (I - K_tM_2)P_{t|t-1}
\]

The likelihood function is computed as follows:

\[
L = -\sum_{t=1}^{T} \log(\det(S_t)) - \sum_{t=1}^{T} \eta_t'\Sigma_t^{-1}\eta_t
\]

The covariance matrix is the shocks is:

\[
\Sigma \equiv \text{var}\left(\begin{bmatrix}
\epsilon^g_{t+1} \\
\epsilon^{\mu}_{t+1} \\
\epsilon^{D}_{t+1} \\
\epsilon^{z}_{t+1}
\end{bmatrix}\right) = \begin{bmatrix}
\sigma^2_g & \sigma_g\mu & \sigma_gD & \sigma_gz \\
\sigma_g\mu & \sigma^2_\mu & \sigma_\muD & \sigma_\muz \\
\sigma_gD & \sigma_\muD & \sigma^2_D & \sigma_Dz \\
\sigma_gz & \sigma_\muz & \sigma_Dz & \sigma^2_z
\end{bmatrix}
\]

I maximize the likelihood over the parameters using simulated annealing:

\[
\Phi = \{\gamma_0, \delta_0, \xi_0, \gamma_1, \delta_1, \xi_1, \gamma_2, \delta_2, \sigma_g, \sigma_\mu, \sigma_D, \sigma_z, \sigma_{g\mu}, \sigma_{gD}, \sigma_{gz}, \sigma_{\muD}, \sigma_{\muz}, \sigma_{Dz}, \sigma_{Dz}\}
\]

**Appendix C: Wold decomposition**

I start by using the Kalman filter in Appendix A in stationary state \((K_t = K)\).

I can express the filtered state in terms of historical growth rates, price-dividend
ratios and predictive variable:

\[
X_{t|t} = X_{t|t-1} + K\eta_t
\]

\[
= X_{t|t-1} + K(Y_t - M_0 - M_1 Y_{t-1} - M_2 X_{t|t-1})
\]

\[
= (I - KM_2) X_{t|t-1} + K(Y_t - M_0 - M_1 Y_{t-1})
\]

\[
= (I - KM_2) F X_{t-1|t-1} + K(Y_t - M_0 - M_1 Y_{t-1})
\]

\[
\cdots
\]

\[
= \sum_{i=0}^{\infty} [(I - KM_2) F]^i K(Y_{t-i} - M_0 - M_1 Y_{t-1-i})
\]

Recall that the first element of \(X_t\) is \(\hat{g}_{t-1} = g_{t-1} - \gamma_0\). Hence, we focus on \(X_{t|t-1}\):

\[
X_{t|t-1} = FX_{t-1|t-1}
\]

\[
= F \sum_{i=0}^{\infty} [(I - KM_2) F]^i K(Y_{t-1-i} - M_0 - M_1 Y_{t-2-i})
\]

This shows that the first element of \(X_{t|t-1}\) equals \(\hat{g}_{t-1|t-1}\), which is the filtered value of expected growth rates up to time \(t - 1\). Now we define \(\epsilon_t^d = \Delta d_t - \gamma_0 - \hat{g}_{t-1|t-1}\), and obtain:

\[
\Delta d_t = \gamma_0 + e'_1 X_{t|t-1} + \epsilon_t^d
\]

\[
= \gamma_0 + e'_1 F \sum_{i=0}^{\infty} [(I - KM_2) F]^i K(Y_{t-1-i} - M_0 - M_1 Y_{t-2-i}) + \epsilon_t^d
\]

\[
= a_0^d + \sum_{i=0}^{\infty} a_{i1}^d p d_{t-1-i} + \sum_{i=0}^{\infty} a_{2i}^d \Delta d_{t-1-i} + \sum_{i=0}^{\infty} a_{3i}^d (z_{t-1-i} - \xi_0) + \epsilon_t^d
\]

with:

\[
a_0^d = \gamma_0 - e'_1 F \sum_{i=0}^{\infty} [(I - KM_2) F]^i KM_0
\]

\[
a_{i1}^d = e'_1 F Ke_2, \text{ if } i = 0
\]

\[
= e'_1 F [(I - KM_2) F]^{i-1} ((I - KM_2) FK - KM_1) e_2, \text{ if } i \neq 0
\]

\[
a_{2i}^d = e'_1 F Ke_1, \text{ if } i = 0
\]

\[
= e'_1 F [(I - KM_2) F]^{i-1} ((I - KM_2) FK - KM_1) e_1, \text{ if } i \neq 0
\]

\[
a_{3i}^d = e'_1 F Ke_3, \text{ if } i = 0
\]

\[
= e'_1 F [(I - KM_2) F]^{i-1} ((I - KM_2) FK - KM_1) e_3, \text{ if } i \neq 0
\]

For returns, note that from the expression for the log price-dividend ratio, we have:

\[
\hat{\mu}_{t-1|t-1} = -A_2^{-1} [pd_{t-1} - A_0 - A_1 \hat{g}_{t-1|t-1} - A_3 (z_{t-1} - \xi_0)]
\]
Hence,

\[ r_t = \delta_0 + \mu_{t-1|t-1} + \epsilon_t^* \]

\[ = \delta_0 - A_2^{-1} [pd_{t-1} - A_0 - A_1 \hat{g}_{t-1|t-1} - A_3(z_{t-1} - \xi_0)] + \epsilon_t^* \]

\[ = a_r^* + \sum_{i=0}^{\infty} a_{1i}^* p_{d_{t-1-i}} + \sum_{i=0}^{\infty} a_{2i}^* \Delta d_{t-1-i} + \sum_{i=0}^{\infty} a_{3i}^* (z_{t-1-i} - \xi_0) + \epsilon_t^* \]

with:

\[ a_r^* = \delta_0 + A_2^{-1} A_0 - A_2^{-1} A_1 \epsilon_1^* F \sum_{i=0}^{\infty} [(I - \mathbf{K} \mathbf{M}_2) \mathbf{F}]^i \mathbf{K} \mathbf{M}_0 \]

\[ a_{1i}^* = A_1 \frac{a_{1i}^d}{A_2} - \frac{1}{A_2}, \quad \text{if } i = 0 \]

\[ = A_1 \frac{a_{1i}^d}{A_2}, \quad \text{if } i \neq 0 \]

\[ a_{2i}^* = A_1 \frac{a_{2i}^d}{A_2}, \]

\[ a_{3i}^* = A_1 \frac{a_{3i}^d}{A_2} + \frac{A_3}{A_2}, \quad \text{if } i = 0 \]

\[ = A_1 \frac{a_{3i}^d}{A_2}, \quad \text{if } i \neq 0 \]
Table 1. Summary Statistics

The table shows the summary statistics for returns, dividend growth rates, price-dividend ratio and predictive variables. I use data from 1945-2010. The predictive variables are: risk-free rate (rfree), stock variance (svar), book-to-market ratio (bm), net equity expansion (ntis), percent equity issuing (eqis), long term yield (lty), long term rate of returns (ltr), corporate bond yields on AAA and BAA-rated bonds (AAA and BAA), inflation (infl), and consumption, wealth, income ratio (cay). I explain each variable in Section 3.1.

<table>
<thead>
<tr>
<th></th>
<th>Definition</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>75%</th>
<th>25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta d_t$</td>
<td>log dividend growth rate</td>
<td>0.0545</td>
<td>0.0678</td>
<td>0.0477</td>
<td>0.0820</td>
<td>0.0205</td>
</tr>
<tr>
<td>pd$_t$</td>
<td>log price-dividend ratio</td>
<td>3.4542</td>
<td>0.4282</td>
<td>3.3927</td>
<td>3.6503</td>
<td>3.1490</td>
</tr>
<tr>
<td>$r_t$</td>
<td>log return</td>
<td>0.0998</td>
<td>0.1670</td>
<td>0.1331</td>
<td>0.2227</td>
<td>0.0011</td>
</tr>
<tr>
<td>bm</td>
<td>book-to-market ratio</td>
<td>0.5495</td>
<td>0.2473</td>
<td>0.5257</td>
<td>0.7213</td>
<td>0.3550</td>
</tr>
<tr>
<td>tbl</td>
<td>treasury bill rate</td>
<td>0.0444</td>
<td>0.0303</td>
<td>0.0437</td>
<td>0.0577</td>
<td>0.0219</td>
</tr>
<tr>
<td>AAA</td>
<td>AAA-rated bond yield</td>
<td>0.0655</td>
<td>0.0283</td>
<td>0.0634</td>
<td>0.0831</td>
<td>0.0435</td>
</tr>
<tr>
<td>BAA</td>
<td>BAA-rated bond yield</td>
<td>0.0754</td>
<td>0.0315</td>
<td>0.0739</td>
<td>0.0912</td>
<td>0.0502</td>
</tr>
<tr>
<td>lty</td>
<td>long-term yield</td>
<td>0.0391</td>
<td>0.0273</td>
<td>0.0557</td>
<td>0.0760</td>
<td>0.0395</td>
</tr>
<tr>
<td>cay</td>
<td>consumption, wealth, income ratio</td>
<td>0.0000</td>
<td>0.0234</td>
<td>-0.0023</td>
<td>0.0153</td>
<td>-0.0137</td>
</tr>
<tr>
<td>ntis</td>
<td>net equity expansion</td>
<td>0.0164</td>
<td>0.0180</td>
<td>0.0205</td>
<td>0.0270</td>
<td>0.0105</td>
</tr>
<tr>
<td>rfree</td>
<td>risk-free rate</td>
<td>0.0453</td>
<td>0.0311</td>
<td>0.0436</td>
<td>0.0577</td>
<td>0.0179</td>
</tr>
<tr>
<td>infl</td>
<td>inflation</td>
<td>0.0392</td>
<td>0.0343</td>
<td>0.0305</td>
<td>0.0471</td>
<td>0.0170</td>
</tr>
<tr>
<td>eqis</td>
<td>equity issuing activity ratio</td>
<td>0.1915</td>
<td>0.0811</td>
<td>0.1953</td>
<td>0.2270</td>
<td>0.1390</td>
</tr>
<tr>
<td>ltr</td>
<td>long-term rate of returns</td>
<td>0.0626</td>
<td>0.1052</td>
<td>0.0368</td>
<td>0.1210</td>
<td>-0.0092</td>
</tr>
<tr>
<td>corpr</td>
<td>corporate bond return</td>
<td>0.0639</td>
<td>0.0237</td>
<td>0.0446</td>
<td>0.1076</td>
<td>0.0020</td>
</tr>
<tr>
<td>svar</td>
<td>stock variance</td>
<td>0.0213</td>
<td>0.0241</td>
<td>0.0133</td>
<td>0.0250</td>
<td>0.0081</td>
</tr>
</tbody>
</table>
Table 2. Present Value Model Estimation

The table reports the estimation results of the present-value model (4) with each predictive variable used as an additional instrument. The models are estimated by maximizing conditional maximum likelihood values using data from 1945 to 2010. The R-squared values for returns and dividend growth rates are reported in percentage. They are computed as in (5).

<table>
<thead>
<tr>
<th></th>
<th>none</th>
<th>bm</th>
<th>tbl</th>
<th>AAA</th>
<th>BAA</th>
<th>lty</th>
<th>cay</th>
<th>ntis</th>
<th>rfree</th>
<th>infl</th>
<th>eqis</th>
<th>ltr</th>
<th>corpr</th>
<th>svar</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_0$</td>
<td>0.0819</td>
<td>0.0784</td>
<td>0.0820</td>
<td>0.0802</td>
<td>0.0808</td>
<td>0.0796</td>
<td>0.0822</td>
<td>0.0810</td>
<td>0.0823</td>
<td>0.0829</td>
<td>0.0838</td>
<td>0.0835</td>
<td>0.0832</td>
<td>0.0794</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.9288</td>
<td>0.8861</td>
<td>0.9273</td>
<td>0.9210</td>
<td>0.9205</td>
<td>0.9281</td>
<td>0.9273</td>
<td>0.9328</td>
<td>0.9195</td>
<td>0.9009</td>
<td>0.8742</td>
<td>0.9155</td>
<td>0.9147</td>
<td>0.9641</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>0.0549</td>
<td>0.0514</td>
<td>0.0550</td>
<td>0.0541</td>
<td>0.0549</td>
<td>0.0338</td>
<td>0.0556</td>
<td>0.0547</td>
<td>0.0548</td>
<td>0.0547</td>
<td>0.0552</td>
<td>0.0556</td>
<td>0.0554</td>
<td>0.0564</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>0.2657</td>
<td>0.2062</td>
<td>0.2281</td>
<td>0.2549</td>
<td>0.2456</td>
<td>0.2676</td>
<td>0.2921</td>
<td>0.2613</td>
<td>-0.4381</td>
<td>0.2331</td>
<td>0.2773</td>
<td>-0.1991</td>
<td>-0.3958</td>
<td>0.3890</td>
</tr>
<tr>
<td>$\sigma_\mu$</td>
<td>0.0160</td>
<td>0.0174</td>
<td>0.0163</td>
<td>0.0181</td>
<td>0.0193</td>
<td>0.0167</td>
<td>0.0380</td>
<td>0.0164</td>
<td>0.0185</td>
<td>0.0241</td>
<td>0.0461</td>
<td>0.0394</td>
<td>0.0303</td>
<td>0.0207</td>
</tr>
<tr>
<td>$\sigma_g$</td>
<td>0.0653</td>
<td>0.0640</td>
<td>0.0618</td>
<td>0.0649</td>
<td>0.0645</td>
<td>0.0652</td>
<td>0.0639</td>
<td>0.0614</td>
<td>0.0174</td>
<td>0.0645</td>
<td>0.0651</td>
<td>0.0972</td>
<td>0.1071</td>
<td>0.0709</td>
</tr>
<tr>
<td>$\sigma_D$</td>
<td>0.0025</td>
<td>0.0123</td>
<td>0.0174</td>
<td>0.0053</td>
<td>0.0101</td>
<td>0.0069</td>
<td>0.0138</td>
<td>0.0163</td>
<td>0.0632</td>
<td>0.0069</td>
<td>0.0864</td>
<td>0.0592</td>
<td>0.0594</td>
<td>0.0158</td>
</tr>
<tr>
<td>$\rho_{gz}$</td>
<td>0.2653</td>
<td>0.2542</td>
<td>0.2325</td>
<td>0.2465</td>
<td>0.1741</td>
<td>0.2946</td>
<td>-0.0113</td>
<td>0.2489</td>
<td>-0.7342</td>
<td>0.1677</td>
<td>0.2358</td>
<td>0.8139</td>
<td>0.7415</td>
<td>-0.3485</td>
</tr>
<tr>
<td>$\rho_{zd}$</td>
<td>-0.9642</td>
<td>-0.1625</td>
<td>-0.2067</td>
<td>-0.2424</td>
<td>-0.2577</td>
<td>-0.3228</td>
<td>-0.7412</td>
<td>-0.1023</td>
<td>0.3863</td>
<td>-0.0973</td>
<td>0.8658</td>
<td>-0.0053</td>
<td>0.0420</td>
<td>-0.7218</td>
</tr>
<tr>
<td>$\xi_0$</td>
<td>0.5053</td>
<td>0.0447</td>
<td>0.0725</td>
<td>0.0831</td>
<td>0.0643</td>
<td>0.0007</td>
<td>0.0161</td>
<td>0.0448</td>
<td>0.0392</td>
<td>0.1897</td>
<td>0.0616</td>
<td>0.0636</td>
<td>0.0217</td>
<td></td>
</tr>
<tr>
<td>$\xi_1$</td>
<td>0.9166</td>
<td>0.8631</td>
<td>0.9443</td>
<td>0.9358</td>
<td>0.9350</td>
<td>0.6770</td>
<td>0.6002</td>
<td>0.8999</td>
<td>0.5454</td>
<td>0.6602</td>
<td>-0.0790</td>
<td>-0.0646</td>
<td>0.4410</td>
<td></td>
</tr>
<tr>
<td>$\rho_{gz}$</td>
<td>-0.2390</td>
<td>0.2429</td>
<td>-0.0092</td>
<td>-0.2303</td>
<td>0.2135</td>
<td>-0.1573</td>
<td>-0.0834</td>
<td>0.3334</td>
<td>0.0174</td>
<td>-0.1357</td>
<td>0.9616</td>
<td>0.9701</td>
<td>-0.6015</td>
<td></td>
</tr>
<tr>
<td>$\rho_{pz}$</td>
<td>0.6030</td>
<td>0.0013</td>
<td>0.3776</td>
<td>0.3642</td>
<td>0.3390</td>
<td>0.9228</td>
<td>-0.3981</td>
<td>0.3677</td>
<td>-0.6339</td>
<td>-0.8551</td>
<td>0.8934</td>
<td>0.8143</td>
<td>0.8741</td>
<td></td>
</tr>
<tr>
<td>$\rho_{dz}$</td>
<td>0.5618</td>
<td>0.9583</td>
<td>0.7927</td>
<td>0.7390</td>
<td>0.7808</td>
<td>-0.9218</td>
<td>-0.9532</td>
<td>0.2885</td>
<td>0.7624</td>
<td>-0.9874</td>
<td>-0.1981</td>
<td>-0.1879</td>
<td>-0.7808</td>
<td></td>
</tr>
<tr>
<td>$\sigma_z$</td>
<td>0.1012</td>
<td>0.0152</td>
<td>0.0081</td>
<td>0.0097</td>
<td>0.0087</td>
<td>0.0169</td>
<td>0.0143</td>
<td>0.0138</td>
<td>0.0287</td>
<td>0.0609</td>
<td>0.1051</td>
<td>0.0926</td>
<td>0.0217</td>
<td></td>
</tr>
<tr>
<td>$\delta_2$</td>
<td>0.0179</td>
<td>-0.0214</td>
<td>-0.0472</td>
<td>-0.0476</td>
<td>-0.0545</td>
<td>-0.5633</td>
<td>0.1976</td>
<td>-0.0372</td>
<td>0.3955</td>
<td>0.2277</td>
<td>-0.3869</td>
<td>-0.3353</td>
<td>-0.3661</td>
<td></td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>0.0629</td>
<td>-0.3345</td>
<td>-0.0910</td>
<td>-0.0150</td>
<td>-0.1479</td>
<td>0.3810</td>
<td>0.7856</td>
<td>-0.2885</td>
<td>0.2176</td>
<td>0.0171</td>
<td>-1.0118</td>
<td>-1.2002</td>
<td>0.9863</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. Predictive Regressions

The table reports the OLS regression results of log returns and log dividend growth rates on the lagged log price-dividend ratio and each predictive variable using data from 1945 to 2010. I run the following regressions, where $\beta$'s are the coefficients on the lagged log price-dividend ratio and $\gamma$'s are the coefficients on the lagged predictive variables.

$$
\begin{align*}
  r_{t+1} &= \alpha_r + \beta_r pd_t + \gamma_r z_t + \epsilon_{t+1} \\
  \Delta d_{t+1} &= \alpha_d + \beta_d pd_t + \gamma_d z_t + \epsilon_{t+1}
\end{align*}
$$

T-statistics are reported in the parentheses. R-squared values are reported in percentage.

<table>
<thead>
<tr>
<th></th>
<th>$R^2_{ret}$</th>
<th>$R^2_{div}$</th>
<th>$\beta_r$</th>
<th>$\gamma_r$</th>
<th>$\beta_d$</th>
<th>$\gamma_d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>9.31</td>
<td>0.90</td>
<td>-0.119</td>
<td>-0.015</td>
<td>(2.54)</td>
<td>(-0.75)</td>
</tr>
<tr>
<td>bm</td>
<td>9.69</td>
<td>4.24</td>
<td>-0.161</td>
<td>-0.082</td>
<td>(1.71)</td>
<td>(0.89)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.51)</td>
<td>(0.51)</td>
<td>(1.47)</td>
<td>(1.30)</td>
</tr>
<tr>
<td>tbl</td>
<td>9.52</td>
<td>1.01</td>
<td>-0.122</td>
<td>-0.258</td>
<td>(-2.55)</td>
<td>(-0.78)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.38)</td>
<td>(0.38)</td>
<td>(1.30)</td>
<td>(0.78)</td>
</tr>
<tr>
<td>AAA</td>
<td>9.59</td>
<td>1.07</td>
<td>-0.120</td>
<td>0.308</td>
<td>(-2.53)</td>
<td>(-0.75)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.44)</td>
<td>(0.44)</td>
<td>(1.30)</td>
<td>(0.75)</td>
</tr>
<tr>
<td>BAA</td>
<td>9.85</td>
<td>1.67</td>
<td>-0.119</td>
<td>0.386</td>
<td>(-2.52)</td>
<td>(-0.76)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.61)</td>
<td>(0.61)</td>
<td>(1.30)</td>
<td>(0.76)</td>
</tr>
<tr>
<td>lty</td>
<td>9.51</td>
<td>0.90</td>
<td>-0.119</td>
<td>0.270</td>
<td>(-2.51)</td>
<td>(-0.75)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.37)</td>
<td>(0.37)</td>
<td>(1.30)</td>
<td>(0.75)</td>
</tr>
<tr>
<td>cay</td>
<td>15.80</td>
<td>0.93</td>
<td>-0.097</td>
<td>1.850</td>
<td>(-2.08)</td>
<td>(-0.70)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.89)</td>
<td>(1.89)</td>
<td>(1.14)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>ntis</td>
<td>9.34</td>
<td>2.94</td>
<td>-0.121</td>
<td>0.155</td>
<td>(-2.08)</td>
<td>(-0.70)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.14)</td>
<td>(0.14)</td>
<td>(1.14)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>rfree</td>
<td>9.32</td>
<td>2.40</td>
<td>-0.120</td>
<td>-0.051</td>
<td>(-2.51)</td>
<td>(-0.87)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(1.14)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>infl</td>
<td>11.70</td>
<td>2.59</td>
<td>-0.141</td>
<td>-0.791</td>
<td>(-2.84)</td>
<td>(-0.36)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.29)</td>
<td>(1.29)</td>
<td>(1.14)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>eqis</td>
<td>18.10</td>
<td>3.21</td>
<td>-0.186</td>
<td>-0.699</td>
<td>(-3.58)</td>
<td>(-1.26)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.57)</td>
<td>(1.57)</td>
<td>(1.14)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>ltr</td>
<td>11.70</td>
<td>4.76</td>
<td>-0.126</td>
<td>0.243</td>
<td>(-2.68)</td>
<td>(-0.59)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.28)</td>
<td>(1.28)</td>
<td>(1.14)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>corpr</td>
<td>12.00</td>
<td>1.17</td>
<td>-0.128</td>
<td>0.296</td>
<td>(-2.71)</td>
<td>(-0.69)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.37)</td>
<td>(1.37)</td>
<td>(1.14)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>svar</td>
<td>11.80</td>
<td>7.66</td>
<td>-0.032</td>
<td>1.100</td>
<td>(-2.77)</td>
<td>(-0.34)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1.31)</td>
<td>(1.31)</td>
<td>(1.14)</td>
<td>(0.64)</td>
</tr>
</tbody>
</table>
Table 4. Persistence of Expected Returns and Expected Dividend Growth Rates

The table reports the persistence coefficients for the model-implied expected returns and expected dividend growth rates. That is, I compute the following:

\[ \rho_\mu = \frac{\text{cov}(\mu_{t+1}, \mu_t)}{\text{var}(\mu_t)}; \rho_g = \frac{\text{cov}(g_{t+1}, g_t)}{\text{var}(g_t)} \]

I report \( \rho_\mu \) and \( \rho_g \) for each predictive variable. I also report the persistence coefficient implied by the predictive regressions with the price-dividend ratio as a sole predictor. Note that as expected returns and expected dividend growth rates are not AR(1) processes, the following values can differ from \( \delta_1 \) and \( \gamma_1 \).

<table>
<thead>
<tr>
<th>Predictive regression</th>
<th>Expected Returns</th>
<th>Expected Dividends</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.9230</td>
<td>0.9230</td>
</tr>
<tr>
<td>pd</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Present-value model</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>0.9288</td>
<td>0.2657</td>
</tr>
<tr>
<td>km</td>
<td>0.9291</td>
<td>0.2676</td>
</tr>
<tr>
<td>tbl</td>
<td>0.9282</td>
<td>0.2541</td>
</tr>
<tr>
<td>AAA</td>
<td>0.9253</td>
<td>0.2563</td>
</tr>
<tr>
<td>BAA</td>
<td>0.9250</td>
<td>0.2456</td>
</tr>
<tr>
<td>lty</td>
<td>0.9339</td>
<td>0.2712</td>
</tr>
<tr>
<td>cay</td>
<td>0.9470</td>
<td>0.3062</td>
</tr>
<tr>
<td>ntis</td>
<td>0.9399</td>
<td>0.2937</td>
</tr>
<tr>
<td>rfree</td>
<td>0.9226</td>
<td>-0.3337</td>
</tr>
<tr>
<td>infl</td>
<td>0.9340</td>
<td>0.2409</td>
</tr>
<tr>
<td>eqis</td>
<td>0.9111</td>
<td>0.2776</td>
</tr>
<tr>
<td>ltr</td>
<td>0.9089</td>
<td>-0.2405</td>
</tr>
<tr>
<td>corpr</td>
<td>0.9092</td>
<td>-0.4241</td>
</tr>
<tr>
<td>svar</td>
<td>0.9709</td>
<td>0.4331</td>
</tr>
</tbody>
</table>

32
Table 5. Autoregressive Regressions of Predictive Variables

The table reports the autoregressive OLS regression results of each predictive variables using data from 1945 to 2010. I estimate the following AR(1) and AR(3) regression models.

\[
AR(1): \quad z_{t+1} = a + b z_t + \epsilon_{t+1}
\]

\[
AR(3): \quad z_{t+1} = a + b z_t + c z_{t-1} + d z_{t-2} + \epsilon_{t+1}
\]

T-statistics are reported in the parentheses. R-squared values are reported in percentage.

<table>
<thead>
<tr>
<th></th>
<th>AR(1) R²</th>
<th>AR(1) a</th>
<th>AR(1) b</th>
<th>AR(3) R²</th>
<th>AR(3) a</th>
<th>AR(3) b</th>
<th>AR(3) c</th>
<th>AR(3) d</th>
</tr>
</thead>
<tbody>
<tr>
<td>pcl</td>
<td>83.6</td>
<td>0.0276</td>
<td>0.9230</td>
<td>84.3</td>
<td>0.2510</td>
<td>0.7970</td>
<td>-0.0153</td>
<td>0.1510</td>
</tr>
<tr>
<td></td>
<td>(1.54)</td>
<td>(17.9)</td>
<td>(6.35)</td>
<td>(13.5)</td>
<td>(6.28)</td>
<td>(-0.09)</td>
<td>(1.15)</td>
<td></td>
</tr>
<tr>
<td>bm</td>
<td>83.3</td>
<td>0.0421</td>
<td>0.9170</td>
<td>84.6</td>
<td>0.0227</td>
<td>0.7870</td>
<td>-0.0890</td>
<td>0.2460</td>
</tr>
<tr>
<td></td>
<td>(1.35)</td>
<td>(17.8)</td>
<td>(6.70)</td>
<td>(10.7)</td>
<td>(6.25)</td>
<td>(-0.55)</td>
<td>(1.96)</td>
<td></td>
</tr>
<tr>
<td>tbl</td>
<td>73.8</td>
<td>0.0063</td>
<td>0.8610</td>
<td>75.5</td>
<td>0.0064</td>
<td>1.1300</td>
<td>-0.4940</td>
<td>0.2660</td>
</tr>
<tr>
<td></td>
<td>(1.79)</td>
<td>(13.3)</td>
<td>(8.88)</td>
<td>(13.7)</td>
<td>(8.88)</td>
<td>(-2.7)</td>
<td>(1.77)</td>
<td></td>
</tr>
<tr>
<td>AAA</td>
<td>91.5</td>
<td>0.0040</td>
<td>0.9440</td>
<td>91.0</td>
<td>0.0043</td>
<td>0.9920</td>
<td>-0.0773</td>
<td>0.0264</td>
</tr>
<tr>
<td></td>
<td>(1.56)</td>
<td>(26.1)</td>
<td>(1.65)</td>
<td>(8.88)</td>
<td>(7.61)</td>
<td>(-0.42)</td>
<td>(0.21)</td>
<td></td>
</tr>
<tr>
<td>BAA</td>
<td>90.2</td>
<td>0.0053</td>
<td>0.9370</td>
<td>90.0</td>
<td>0.0056</td>
<td>1.1000</td>
<td>-0.2540</td>
<td>0.0880</td>
</tr>
<tr>
<td></td>
<td>(1.65)</td>
<td>(24.1)</td>
<td>(1.64)</td>
<td>(8.47)</td>
<td>(7.61)</td>
<td>(-1.31)</td>
<td>(0.66)</td>
<td></td>
</tr>
<tr>
<td>lty</td>
<td>89.5</td>
<td>0.0043</td>
<td>0.9340</td>
<td>89.2</td>
<td>0.0041</td>
<td>0.7820</td>
<td>0.1830</td>
<td>-0.0267</td>
</tr>
<tr>
<td></td>
<td>(1.62)</td>
<td>(23.2)</td>
<td>(1.43)</td>
<td>(6.00)</td>
<td>(1.1)</td>
<td>(-0.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cay</td>
<td>45.9</td>
<td>0.0002</td>
<td>0.6740</td>
<td>57.9</td>
<td>-0.0014</td>
<td>0.7490</td>
<td>-0.1230</td>
<td>0.0946</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(7.32)</td>
<td>(-0.84)</td>
<td>(5.88)</td>
<td>(-0.84)</td>
<td>(0.93)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ntis</td>
<td>35.9</td>
<td>0.0064</td>
<td>0.6000</td>
<td>38.2</td>
<td>0.0053</td>
<td>0.6370</td>
<td>-0.1061</td>
<td>0.1090</td>
</tr>
<tr>
<td></td>
<td>(2.52)</td>
<td>(5.95)</td>
<td>(1.94)</td>
<td>(4.94)</td>
<td>(4.94)</td>
<td>(0.72)</td>
<td>(0.84)</td>
<td></td>
</tr>
<tr>
<td>rfree</td>
<td>79.5</td>
<td>0.0049</td>
<td>0.8030</td>
<td>81.7</td>
<td>0.0052</td>
<td>1.2400</td>
<td>0.6100</td>
<td>0.2630</td>
</tr>
<tr>
<td></td>
<td>(1.54)</td>
<td>(15.6)</td>
<td>(1.59)</td>
<td>(9.85)</td>
<td>(9.85)</td>
<td>(-3.23)</td>
<td>(2.09)</td>
<td></td>
</tr>
<tr>
<td>infl</td>
<td>29.6</td>
<td>0.0179</td>
<td>0.5450</td>
<td>50.6</td>
<td>0.0112</td>
<td>0.8680</td>
<td>-0.3360</td>
<td>0.1490</td>
</tr>
<tr>
<td></td>
<td>(3.23)</td>
<td>(5.15)</td>
<td>(2.29)</td>
<td>(7.09)</td>
<td>(7.09)</td>
<td>(-2.77)</td>
<td>(1.59)</td>
<td></td>
</tr>
<tr>
<td>eqis</td>
<td>43.3</td>
<td>0.0647</td>
<td>0.6590</td>
<td>44.4</td>
<td>0.0572</td>
<td>0.6280</td>
<td>0.0249</td>
<td>0.0377</td>
</tr>
<tr>
<td></td>
<td>(3.27)</td>
<td>(6.94)</td>
<td>(2.51)</td>
<td>(4.81)</td>
<td>(4.81)</td>
<td>(0.17)</td>
<td>(0.30)</td>
<td></td>
</tr>
<tr>
<td>ltr</td>
<td>1.3</td>
<td>0.0690</td>
<td>-0.1130</td>
<td>6.8</td>
<td>0.0501</td>
<td>-0.1290</td>
<td>0.1600</td>
<td>0.1910</td>
</tr>
<tr>
<td></td>
<td>(4.52)</td>
<td>(-0.90)</td>
<td>(2.58)</td>
<td>(-1.01)</td>
<td>(1.23)</td>
<td>(1.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>corpr</td>
<td>0.4</td>
<td>0.0604</td>
<td>0.0619</td>
<td>3.2</td>
<td>0.0498</td>
<td>0.0300</td>
<td>0.1410</td>
<td>0.0904</td>
</tr>
<tr>
<td></td>
<td>(4.3)</td>
<td>(0.49)</td>
<td>(2.84)</td>
<td>(0.23)</td>
<td>(1.1)</td>
<td>(0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>svar</td>
<td>17.1</td>
<td>0.0127</td>
<td>0.4150</td>
<td>18.8</td>
<td>0.0143</td>
<td>0.4570</td>
<td>-0.0645</td>
<td>-0.0748</td>
</tr>
<tr>
<td></td>
<td>(3.45)</td>
<td>(3.61)</td>
<td>(2.83)</td>
<td>(3.53)</td>
<td>(4.6)</td>
<td>(-0.46)</td>
<td>(-0.36)</td>
<td></td>
</tr>
</tbody>
</table>
Table 6. Model Implied Parameters

The table reports the model implied parameters. I report the implied values of $A_2$ and $A_4$, which are the coefficients of $z_t$ on the price-dividend ratio formula (3):

$$pd_t = A_0 + [A_1(g_t - \gamma_0) + A_2(z_t - \xi_0)] + [A_3(\mu_t - \delta_t) + A_4(z_t - \xi_0)]$$

I also report the implied correlation between unexpected returns and innovations to expected returns. The values are implied by the estimates of the present-value model parameters (4).

<table>
<thead>
<tr>
<th></th>
<th>none</th>
<th>bm</th>
<th>tbl</th>
<th>AAA</th>
<th>BAA</th>
<th>lty</th>
<th>cay</th>
<th>ntis</th>
<th>rfree</th>
<th>infl</th>
<th>eqis</th>
<th>ltr</th>
<th>corpr</th>
<th>svar</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_2 + A_4$</td>
<td>-0.4293</td>
<td>-1.2925</td>
<td>3.6613</td>
<td>-4.4109</td>
<td>3.5589</td>
<td>17.2728</td>
<td>-2.3454</td>
<td>1.0834</td>
<td>-5.8441</td>
<td>-3.9563</td>
<td>2.3327</td>
<td>1.9329</td>
<td>12.1597</td>
<td></td>
</tr>
<tr>
<td>$A_2$</td>
<td>0.6838</td>
<td>-2.5493</td>
<td>-1.3850</td>
<td>-0.2056</td>
<td>-2.0681</td>
<td>1.4900</td>
<td>2.4390</td>
<td>-1.5193</td>
<td>0.5782</td>
<td>0.0630</td>
<td>-0.7637</td>
<td>-0.7913</td>
<td>2.6811</td>
<td></td>
</tr>
<tr>
<td>$\rho_{rp}$</td>
<td>-0.0022</td>
<td>-0.0022</td>
<td>-0.0023</td>
<td>-0.0024</td>
<td>-0.0026</td>
<td>-0.0021</td>
<td>-0.0044</td>
<td>-0.0022</td>
<td>-0.0027</td>
<td>-0.0017</td>
<td>-0.0029</td>
<td>-0.0025</td>
<td>-0.0018</td>
<td>-0.0028</td>
</tr>
</tbody>
</table>
Table 7. Hypothesis Tests

The table reports the likelihood ratio test results. The tests are described in Section 4.2. R-squared values (%) are also reported. One asterisk (*), two asterisks (**), and 3 asterisks (***) indicate that the null hypothesis can be rejected at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th></th>
<th>bm</th>
<th>tbl</th>
<th>AAA</th>
<th>BAA</th>
<th>lty</th>
<th>cay</th>
<th>ntis</th>
<th>rfree</th>
<th>infl</th>
<th>eqis</th>
<th>ltr</th>
<th>corpr</th>
<th>svar</th>
</tr>
</thead>
</table>

No Predictive Variable

<table>
<thead>
<tr>
<th></th>
<th>LR</th>
<th>***</th>
<th>8.22</th>
<th>3.34</th>
<th>17.30</th>
<th>3.42</th>
<th>27.98</th>
<th>7.64</th>
<th>7.35</th>
<th>9.26</th>
<th>7.39</th>
<th>10.14</th>
<th>10.93</th>
<th>33.18</th>
</tr>
</thead>
</table>

No Return Predictability

|-------|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|

No Dividend Growth Rate Predictability

<table>
<thead>
<tr>
<th></th>
<th>LR</th>
<th>8.59</th>
<th>4.87</th>
<th>4.34</th>
<th>8.10</th>
<th>6.75</th>
<th>5.58</th>
<th>5.86</th>
<th>2.83</th>
<th>5.03</th>
<th>4.99</th>
<th>12.79</th>
<th>11.00</th>
<th>12.13</th>
</tr>
</thead>
</table>

* * * * * * * * * * * * *
Table 8. Variance Decomposition

The table reports the variance decomposition of the log price-dividend ratio implied by the estimated present-value model with each predictive variable. I decompose the variance into three components: expected dividend variance, expected return variance, covariance between the two. All values are reported in percentage.

<table>
<thead>
<tr>
<th></th>
<th>bm</th>
<th>tbl</th>
<th>AAA</th>
<th>BAA</th>
<th>lty</th>
<th>cay</th>
<th>ntis</th>
<th>rfree</th>
<th>infl</th>
<th>eqis</th>
<th>ltr</th>
<th>corpr</th>
<th>svar</th>
</tr>
</thead>
<tbody>
<tr>
<td>$var_\mu$</td>
<td>104.78</td>
<td>102.27</td>
<td>120.24</td>
<td>102.49</td>
<td>125.07</td>
<td>94.25</td>
<td>106.91</td>
<td>110.01</td>
<td>103.60</td>
<td>100.05</td>
<td>101.19</td>
<td>101.00</td>
<td>96.45</td>
</tr>
<tr>
<td>$var_g$</td>
<td>2.20</td>
<td>5.13</td>
<td>2.19</td>
<td>0.19</td>
<td>3.54</td>
<td>0.24</td>
<td>1.76</td>
<td>2.36</td>
<td>0.40</td>
<td>0.65</td>
<td>0.34</td>
<td>0.38</td>
<td>0.36</td>
</tr>
<tr>
<td>$2cov_{\mu,g}$</td>
<td>-6.98</td>
<td>-7.41</td>
<td>-22.44</td>
<td>-2.68</td>
<td>-28.61</td>
<td>5.52</td>
<td>-8.67</td>
<td>-12.36</td>
<td>-4.00</td>
<td>-0.70</td>
<td>-1.53</td>
<td>-1.38</td>
<td>3.19</td>
</tr>
</tbody>
</table>
Table 9. Hypothesis Tests for Extended Present-Value Model

The table reports the likelihood ratio test results for the extended present-value model in Section 4.6.1. The tests are described in the main text. R-squared values (%) are also reported. One asterisk (*), two asterisks (**), and 3 asterisks (***) indicate that the null hypothesis can be rejected at the 10%, 5%, and 1% levels, respectively.

<table>
<thead>
<tr>
<th>z</th>
<th>bm</th>
<th>bm</th>
<th>bm</th>
<th>BAA</th>
<th>BAA</th>
<th>cay</th>
<th>cay</th>
<th>svar</th>
<th>svar</th>
<th>svar</th>
</tr>
</thead>
<tbody>
<tr>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2_{ret}$</td>
<td>8.95</td>
<td>10.57</td>
<td>8.57</td>
<td>13.74</td>
<td>8.78</td>
<td>13.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2_{div}$</td>
<td>23.24</td>
<td>21.80</td>
<td>21.15</td>
<td>15.29</td>
<td>21.80</td>
<td>19.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Joint Significance of Instruments z and w

<table>
<thead>
<tr>
<th>LR</th>
<th>83.52</th>
<th>106.66</th>
<th>110.47</th>
<th>40.44</th>
<th>47.94</th>
<th>47.79</th>
</tr>
</thead>
</table>

Significance of z

<table>
<thead>
<tr>
<th>LR</th>
<th>16.91</th>
<th>22.73</th>
<th>5.20</th>
<th>12.83</th>
<th>12.62</th>
<th>8.52</th>
</tr>
</thead>
</table>

Significance of w

<table>
<thead>
<tr>
<th>LR</th>
<th>13.11</th>
<th>27.32</th>
<th>24.53</th>
<th>15.30</th>
<th>19.58</th>
<th>17.84</th>
</tr>
</thead>
</table>
Table 10. Out-of-sample Return Predictability

The table reports the out-of-sample return R-squared values (%) for each predictive variable. The first row reports the predictive regression results and the second row reports the present-value approach results.

<table>
<thead>
<tr>
<th></th>
<th>none</th>
<th>bm</th>
<th>tbl</th>
<th>AAA</th>
<th>BAA</th>
<th>lty</th>
<th>cay</th>
<th>ntis</th>
<th>rfree</th>
<th>infl</th>
<th>eqis</th>
<th>ltr</th>
<th>corpr</th>
<th>svar</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2_{reg}$</td>
<td>2.36</td>
<td>-5.34</td>
<td>-3.63</td>
<td>-3.47</td>
<td>-3.36</td>
<td>-4.22</td>
<td>11.29</td>
<td>-13.53</td>
<td>-3.29</td>
<td>-0.22</td>
<td>-6.16</td>
<td>-14.44</td>
<td>-9.72</td>
<td>-0.75</td>
</tr>
<tr>
<td>$R^2_{pv}$</td>
<td>2.87</td>
<td>3.99</td>
<td>-5.97</td>
<td>-10.36</td>
<td>-2.48</td>
<td>-2.0</td>
<td>13.39</td>
<td>-13.43</td>
<td>-5.15</td>
<td>0.56</td>
<td>-5.72</td>
<td>-16.45</td>
<td>-5.38</td>
<td>-9.01</td>
</tr>
</tbody>
</table>
Figure 1. Filtered Expected Returns with (bm, BAA, cay)

I plot the filtered expected return series ($\hat{\mu}_t$) from the present-value model (4) with each of the following predictive variables: bm, BAA, cay. For comparison, I also plot the expected returns from the OLS regression of $r_{t+1}$ on $pd_t$ and from the present-value model without an instrument.
Figure 2. Filtered Expected Returns with (eqis, corpr, svar)

I plot the filtered expected return series ($\hat{\mu}_t$) from the present-value model (4) with each of the following predictive variables: eqis, corpr, svar. For comparison, I also plot the expected returns from the OLS regression of $r_{t+1}$ on $pd_t$ and from the present-value model without an instrument.

[Graph showing the expected returns over time for different models, including realized returns, eqis, corpr, svar, and OLS with pd]
Figure 3. Filtered Expected Dividend Growth Rates with (bm, BAA, cay)

I plot the filtered expected dividend growth rate series ($\hat{g}_t$) from the present-value model (4) with each of the following predictive variables: bm, BAA, cay. For comparison, I also plot the expected dividend growth rates from the OLS regression of $\Delta d_{t+1}$ on $pd_t$ and from the present-value model without an instrument.
Figure 4. Filtered Expected Dividend Growth Rates with (eqis, corpr, svar)

I plot the filtered expected dividend growth rate series ($\hat{\gamma}_t$) from the present-value model (4) with each of the following predictive variables: eqis, corpr, svar. For comparison, I also plot the expected dividend growth rates from the OLS regression of $\Delta d_{t+1}$ on $pd_t$ and from the present-value model without an instrument.
Figure 5. Wold Decomposition: Returns with (cay, eqis)

I plot the Wold decomposition of returns implied by the present-value model (4) with each of the following variables: cay, eqis. Each line denotes the coefficients \( \{a_{1,i}\} \) for \( i = 1, \ldots, 5 \) from the implied VAR-MA representation (8) from the present-value model with each predictive variable.
Figure 6. Wold Decomposition: Returns with (rfree, ltr, corpr)

I plot the Wold decomposition of returns implied by the present-value model (4) with each of the following variables: rfree, ltr, corpr. Each line denotes the coefficients ($\{a^{x}_{3,i}\}$ for $i = 1, \ldots, 5$) from the implied VAR-MA representation (8) from the present-value model with each predictive variable.
Figure 7. Wold Decomposition: Dividends with (cay, rfree, eqis, ltr, corpr)

I plot the Wold decomposition of dividend growth rates implied by the present-value model (4) with each of the following variables: cay, rfree, ltr, corpr. Each line denotes the coefficients \( \{a_{q,i}\} \) for \( i = 1, \ldots, 5 \) from the implied VAR-MA representation (8) from the present-value model with each predictive variable.
Figure 8. Long-run Return Forecasts with (cay, eqis)

I plot the long-run forecasts of returns (9) from the present-value model (4) with each of the following variables: cay, eqis. For comparison, I also plot the long-run forecasts from the first-order VAR with the P/D ratio and the present-value model without an instrument.
Figure 9. Long-run Return Forecasts with (bm, corpr)

I plot the long-run forecasts of returns (9) from the present-value model (4) with each of the following variables: bm, corpr. For comparison, I also plot the long-run forecasts from the first-order VAR with the P/D ratio and the present-value model without an instrument.
I plot the long-run forecasts of dividend growth rates (9) from the present-value model (4) with each of the following variables: cay, eqis. For comparison, I also plot the long-run forecasts from the first-order VAR with the P/D ratio and the present-value model without an instrument.
Figure 11. Long-run Dividend Growth Rate Forecasts with (bm, ltr, corpr)

I plot the long-run forecasts of dividend growth rates (9) from the present-value model (4) with each of the following variables: bm, ltr, corpr. For comparison, I also plot the long-run forecasts from the first-order VAR with the P/D ratio and the present-value model without an instrument.
Figure 12. Long-run Dividend Growth Rate Forecasts with (tbl, AAA, lty, rfree)

I plot the long-run forecasts of dividend growth rates (9) from the present-value model (4) with each of the following variables: tbl, AAA, lty, rfree. For comparison, I also plot the long-run forecasts from the first-order VAR with the P/D ratio and the present-value model without an instrument.
Figure 13. Out-of-sample Return Cumulative SSE Compared to Predictive Regressions

I plot the cumulative sum of squared errors from the present-value model (4) with each predictive variable minus the cumulative sum of squared errors from the corresponding predictive regression (6). It is positive at \( t \), if until time-\( t \), the present-value model outperforms the corresponding predictive regression in terms of the cumulative SSE, and vice versa.
Figure 14. Out-of-sample Return Cumulative SSE Compared to Predictive Regression with P/D Ratio

I plot the cumulative sum of squared errors from the present-value model (4) with each predictive variable minus the cumulative sum of squared errors from the predictive regression only with the log price-dividend ratio. It is positive at $t$, if until time-$t$, the present-value model outperforms the price-dividend-ratio regression in terms of the cumulative SSE, and vice versa.