Predictive Regressions: A Present Value Approach

Jules van Binsbergen and Ralph Koijen

Sept 25, 2007
Most studies have relied on log approximations to derive the present value relationship between the Price-Dividend ratio, and future expected returns and dividend growth.

Empirically, the Price-Dividend ratio seems to predict returns but not dividend growth.

van Binsbergen and Koijen want to avoid the log linear approximation framework, instead develop their own to analyze these predictive relationships.
Results

- Develop a framework that does not rely on the log approximation, Price-Dividend ratio is an exact linear function of expected returns and expected dividend growth rates.

- Find that Price-Dividend ratio predicts both future returns and dividend growth, \( R^2 \) of 16% and 8% respectively.

- Find evidence of a transient and persistent component of the expected dividend growth rate, \( R^2 \) rise to 18% and 16%.
The Model

- Price-Dividend Ratio: \( PD_t = \frac{P_t}{D_t} \)

- Returns: \( R_{t+1} = \frac{P_{t+1}}{P_t - D_t} \)

- Expected returns: \( \mu_t^* = E_t[R_{t+1}] - 1 \)
Do not observe the expected return sequence $\{\mu_t^*\}$ or sequence of expected dividend growth rate $\{g_t\}$

How one models these sequences is extremely important!
Model Dividend Growth as:

\[
\frac{D_{t+1}}{D_t} = (1 + g_t)(1 + \epsilon_{t+1}^D)
\]

Define \( \mu_t \) such that:

\[
\frac{1 + g_t}{1 + \mu_t^*} = 1 + g_t - \mu_t
\]

Introduce two new variables, \( \hat{g}_t \) and \( \hat{\mu}_t \) such that:

\[
g_t = \gamma_0 + \hat{g}_t
\]

\[
\mu_t = \delta_0 + \hat{\mu}_t
\]
Model these as non-linear equations:

\[ \hat{g}_{t+1} = \gamma_1 \frac{1 + \gamma_0 - \delta_0}{1 + \gamma_0 - \delta_0 + \hat{g}_t - \hat{\mu}_t} \hat{g}_t + \epsilon_{g,t+1} \]

\[ \hat{\mu}_{t+1} = \delta_1 \frac{1 + \gamma_0 - \delta_0}{1 + \gamma_0 - \delta_0 + \hat{g}_t - \hat{\mu}_t} \hat{\mu}_t + \epsilon_{\mu,t+1} \]
Re-label as in the paper:

\[ \hat{g}_{t+1} = \gamma_1 t \hat{g}_t + \epsilon_{t+1}^g \]

\[ \hat{\mu}_{t+1} = \delta_1 t \hat{\mu}_t + \epsilon_{t+1}^\mu \]

Where:

\[ \gamma_1 t = \gamma_1 \lambda_t \]

\[ \delta_1 t = \delta_1 \lambda_t \]

\[ \lambda_t = \frac{\alpha}{\alpha + \hat{g}_t - \hat{\mu}_t} \]

\[ \alpha = 1 + \gamma_0 - \delta_0 \]
The covariance between innovations in expected dividend growth and expected return is:

$$\text{Cov}(\epsilon_{g,t+1}^g, \epsilon_{\mu,t+1}^\mu) = \sigma_{g\mu}$$

The covariance between innovations in unexpected dividend growth and expected return is:

$$\text{Cov}(\epsilon_{t+1}^D, \epsilon_{t+1}^\mu) = \sigma_{D\mu}$$

For identification, the covariance between innovations in unexpected and expected dividend growth is zero:

$$\text{Cov}(\epsilon_{t+1}^D, \epsilon_{t+1}^g) = \sigma_{Dg} = 0$$
The Price-Dividend Ratio can be written as:

$$PD_t = A + B_1 \hat{\mu}_t + B_2 \hat{g}_t$$

Where

$$A = \left(1 - \alpha + \frac{\sigma D \mu \alpha}{1 - \delta_1 \alpha + \sigma D \mu}\right)^{-1}$$

$$B_1 = \frac{-A}{1 - \delta_1 \alpha + \sigma D \mu}$$

$$B_2 = \frac{A + B_1 \sigma D \mu}{1 - \gamma_1 \alpha}$$
Why do variables move?

- Changes in the dividend growth, returns, or the price dividend ratio can be attributed to changes in expected returns, expected growth rates, or unexpected growth rates.
- Expected change in $PD_t$ is:
  \[
  E_t[PD_t] - PD_t = B_1(\delta_{1t} - 1)\hat{\mu}_t + B_2(\gamma_{1t} - 1)\hat{g}_t
  \]
- Unexpected change in price dividend ratio is:
  \[
  PD_{t+1} - E_t[PD_{t+1}] = B_1\epsilon^{\mu}_{t+1} + B_2\epsilon^{g}_{t+1}
  \]
Unexpected Returns

- Unexpected return is:

\[ R_{t+1} - (1 + \mu_t^*) = h_t \left[ B_1 \epsilon_{\mu t+1} + B_2 \epsilon_{g t+1} + E_t [PD_{t+1}] \epsilon_{\mu t+1} 
+ B_1 (\epsilon_{t+1} \epsilon_{t+1} - \sigma_{D \mu}) + B_2 \epsilon_{t+1} \epsilon_{t+1} \right] \]

- Where:

\[ h_t = \frac{1 + g_t}{PD_t - 1} \]
Implications for Variance of Returns

If we assume that $(\epsilon_{t+1}^g, \epsilon_{t+1}^\mu, \epsilon_{t+1}^D)$ are jointly normal, then the conditional variance of returns is:

$$\sigma^2_{R,t} = h_t^2 \left[ B_1^2 \sigma_{\mu}^2 + B_2^2 \sigma_g^2 + (E_t[PD_{t+1}])^2 \sigma_D^2 + B_1^2 (\sigma_D^2 \sigma_{\mu}^2 + 2 \sigma_D^2 \sigma_{\mu g}) 
+ B_2^2 \sigma_g^2 \sigma_D^2 + B_1 B_2 \sigma_{g \mu}^2 + B_1 E_t[PD_{t+1}] \sigma_{D,\mu} + B_1 B_2 \sigma_D^2 \sigma_{g \mu} \right]$$

The conditional covariance between expected returns and unexpected returns is:

$$\text{Cov}_t(R_{t+1} - E_t[R_{t+1}], \epsilon_{t+1}^\mu) = h_t (B_1 \sigma_{\mu}^2 + B_2 \sigma_{\mu g} + E_t[PD_{t+1}] \sigma_{\mu d})$$
Data

- Annual data, 1946-2005
- Use total payout data as in some previous studies
First, want to characterize small sample performance of estimator given constant dividend growth and compare their maximum likelihood estimator to that of conventional predictive regressions.

**PV model:** \( PD_t = A + B_1 \hat{\mu}_t \Rightarrow L(\Theta; \Delta D^T, PD^T) \)

**Standard Predictive Regression:** \( R_{t+1} = \alpha + \beta PD_t + \eta_{t+1} \)

Can link the two as follows:

\[
\begin{align*}
\delta_0 &= \alpha + \beta A \\
1 &= \beta B_1
\end{align*}
\]
Empirical Exercise

- Generate 10000 series of 60 observations, given some set of parameters
- Estimate model
- Look at distribution of estimators
Results

Jules van Binsbergen and Ralph Koijen

Predictive Regressions: A Present Value Approach
<table>
<thead>
<tr>
<th></th>
<th>True value</th>
<th>Mean</th>
<th>St.dev</th>
<th>5%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>95%</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_{OLS}$</td>
<td>-0.0064</td>
<td>-0.0106</td>
<td>0.0060</td>
<td>-0.0211</td>
<td>-0.0141</td>
<td>-0.0101</td>
<td>-0.0066</td>
<td>-0.0017</td>
<td>0.00730</td>
</tr>
<tr>
<td>$\beta_{OLS, \text{ Adapted}}$</td>
<td>-0.0064</td>
<td>-0.0096</td>
<td>0.0053</td>
<td>-0.0191</td>
<td>-0.0127</td>
<td>-0.0091</td>
<td>-0.0060</td>
<td>-0.0019</td>
<td>0.00616</td>
</tr>
<tr>
<td>$\beta_{ML}$</td>
<td>-0.0064</td>
<td>-0.0088</td>
<td>0.0032</td>
<td>-0.0148</td>
<td>-0.0107</td>
<td>-0.0084</td>
<td>-0.0065</td>
<td>-0.0044</td>
<td>0.00405</td>
</tr>
<tr>
<td>$\gamma_{0,OLS}$</td>
<td>0.02500</td>
<td>0.0249</td>
<td>0.0186</td>
<td>-0.0059</td>
<td>0.0123</td>
<td>0.0248</td>
<td>0.0376</td>
<td>0.0551</td>
<td>0.01862</td>
</tr>
<tr>
<td>$\gamma_{0,ML}$</td>
<td>0.02500</td>
<td>0.0248</td>
<td>0.0188</td>
<td>-0.0065</td>
<td>0.0124</td>
<td>0.0249</td>
<td>0.0375</td>
<td>0.0555</td>
<td>0.01882</td>
</tr>
<tr>
<td>$\delta_{0,ML}$</td>
<td>0.0600</td>
<td>0.0619</td>
<td>0.0206</td>
<td>0.0292</td>
<td>0.0485</td>
<td>0.0616</td>
<td>0.0749</td>
<td>0.0945</td>
<td>0.02071</td>
</tr>
<tr>
<td>$\delta_{1,ML}$</td>
<td>0.8500</td>
<td>0.7848</td>
<td>0.1001</td>
<td>0.6072</td>
<td>0.7274</td>
<td>0.7960</td>
<td>0.8514</td>
<td>0.9198</td>
<td>0.11947</td>
</tr>
<tr>
<td>$\sigma_{D,ML}$</td>
<td>0.1400</td>
<td>0.1385</td>
<td>0.132</td>
<td>0.1172</td>
<td>0.1293</td>
<td>0.1383</td>
<td>0.1474</td>
<td>0.1607</td>
<td>0.01332</td>
</tr>
<tr>
<td>$\sigma_{\mu,ML}$</td>
<td>0.0180</td>
<td>0.0241</td>
<td>0.0090</td>
<td>0.0118</td>
<td>0.0178</td>
<td>0.0229</td>
<td>0.0293</td>
<td>0.0401</td>
<td>0.01086</td>
</tr>
<tr>
<td>$\rho_{\mu,D,ML}$</td>
<td>0.0516</td>
<td>-0.0006</td>
<td>0.1329</td>
<td>-0.2187</td>
<td>-0.0919</td>
<td>-0.0008</td>
<td>0.0902</td>
<td>0.2169</td>
<td>0.14279</td>
</tr>
</tbody>
</table>
Estimation of Full Model

- **Transition Equations:**
  \[
  \hat{g}_{t+1} = \gamma_1 t \hat{g}_t + \epsilon_{t+1}^g \\
  \hat{\mu}_{t+1} = \delta_1 t \hat{\mu}_t + \epsilon_{t+1}^g
  \]

- **Measurement equations:**
  \[
  PD_t = \frac{D_{t+1}}{D_t} = A + B_1 \hat{\mu}_t + B_2 \hat{g}_t \\
  (1 + \delta_0 + \hat{g}_t) + (1 + \delta_0 + \hat{g}_t) \epsilon_{t+1}^D
  \]
Want to evaluate the likelihood:

\[ L(y^t; \Theta) = \prod_{t=1}^{T} l(y_t | y^{t-1}; \Theta) \]

\[ L(y^t; \Theta) = \prod_{t=1}^{T} \int \int l(y_t | y^{t-1}, \epsilon_g^t, g_0, \Theta) l(\epsilon_g^t, g_0 | y^{t-1}, \Theta) d\epsilon_g^t dg_0 \]

This is hard in a non linear world!
Estimating the Model

To evaluate the likelihood function use 1 of:
- Unscented Kalman Filter
- Particle Filter

Then maximize the likelihood and bootstrap for standard errors.
Can reduce the dimension by subbing out one of the transition equations to get:

$$\hat{g}_{t+1} = \gamma_{1t} \hat{g}_t + \epsilon_{t+1}^g$$

Measurement equations:

$$PD_{t+1} = A + \delta_{1t}[PD_t - A - B_2\hat{g}_t] + B_2\gamma_{1t}\hat{g}_t + B_1\epsilon_{t+1}^\mu + B_2\epsilon_{t+1}^g$$

$$\frac{D_{t+1}}{D_t} = (1 + \delta_0 + \hat{g}_t) + (1 + \delta_0 + \hat{g}_t)\epsilon_{t+1}^D$$
## Results

**Panel A: Maximum-likelihood estimates**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_0$</td>
<td>0.1163</td>
<td>(0.0187)</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.8243</td>
<td>(0.0807)</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>0.0811</td>
<td>(0.0173)</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-0.3347</td>
<td>(0.4186)</td>
</tr>
<tr>
<td>$\sigma_\mu$</td>
<td>0.0282</td>
<td>(0.0192)</td>
</tr>
<tr>
<td>$\sigma_\gamma$</td>
<td>0.0692</td>
<td>(0.0360)</td>
</tr>
<tr>
<td>$\rho_{D\mu}$</td>
<td>0.0404</td>
<td>(0.1795)</td>
</tr>
<tr>
<td>$\rho_{g\mu}$</td>
<td>0.7913</td>
<td>(0.2810)</td>
</tr>
</tbody>
</table>

**Panel B: Implied present-value model parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>27.93</td>
</tr>
<tr>
<td>$B_1$</td>
<td>-136.31</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.96</td>
</tr>
<tr>
<td>$B_2$</td>
<td>29.10</td>
</tr>
</tbody>
</table>

**Panel C: R-squared values**

- $R^2_{\text{Returns}}$: 15.5%
- $R^2_{\text{Div}}$: 8.0%
Expected Returns

Jules van Binsbergen and Ralph Koijen

Predictive Regressions: A Present Value Approach
Expected Dividend Growth Rates

Jules van Binsbergen and Ralph Koijen

Predictive Regressions: A Present Value Approach
Decomposing Unexpected Returns: First Order Effects

Jules van Binsbergen and Ralph Koijen

Predictive Regressions: A Present Value Approach
Decomposing Unexpected Returns: Second Order Effects

Jules van Binsbergen and Ralph Koijen

Predictive Regressions: A Present Value Approach
## Variance Decomposition

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1): $h_t B_1 \varepsilon_{t+1}^\mu$</td>
<td>110.0%</td>
<td>-14.2%</td>
<td>-22.9%</td>
<td>0.2%</td>
<td>0.3%</td>
</tr>
<tr>
<td>(2): $h_t B_2 \varepsilon_{t+1}^g$</td>
<td>-14.2%</td>
<td>2.8%</td>
<td>-5.2%</td>
<td>0.4%</td>
<td>-0.1%</td>
</tr>
<tr>
<td>(3): $h_t E_t (PD_{t+1}) \varepsilon_{t+1}^D$</td>
<td>-22.9%</td>
<td>-5.2%</td>
<td>75.3%</td>
<td>-3.6%</td>
<td>0.3%</td>
</tr>
<tr>
<td>(4): $h_t B_1 (\varepsilon_{t+1}^\mu \varepsilon_{t+1}^D - \sigma_{D\mu})$</td>
<td>0.2%</td>
<td>0.36%</td>
<td>-3.6%</td>
<td>2.2%</td>
<td>-0.2%</td>
</tr>
<tr>
<td>(5): $h_t B_2 (\varepsilon_{t+1}^g \varepsilon_{t+1}^D)$</td>
<td>0.3%</td>
<td>-0.1%</td>
<td>0.3%</td>
<td>-0.2%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>
Decomposing $PD_t$

![Graph showing the decomposition of PD_t with contributions from expected returns and expected dividend growth rates.](image-url)
Want to allow for two growth rates

\[ g_t = \gamma_0 + \hat{g}_1 t + \hat{g}_2 t \]

\[ g_{it} = \gamma_{it} \hat{g}_{it} + \epsilon_{i,t+1} \]

Now the price dividend ratio

\[ PD_t = A + B_1 \hat{\mu}_t + B_2 \hat{g}_1 t + B_3 \hat{g}_2 t \]
### Results

Panel A: Maximum-likelihood estimates

<table>
<thead>
<tr>
<th>Estimate</th>
<th>S.e.</th>
<th>Estimate</th>
<th>S.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_0$</td>
<td>0.1172 (0.0227)</td>
<td>$\gamma_0$</td>
<td>0.0824 (0.0215)</td>
</tr>
<tr>
<td>$\delta_1$</td>
<td>0.8404 (0.4192)</td>
<td>$\gamma_1$</td>
<td>0.5984 (0.1875)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\gamma_2$</td>
<td>-0.7300 (0.2972)</td>
</tr>
<tr>
<td>$\sigma_\mu$</td>
<td>0.0399 (0.0259)</td>
<td>$\sigma_{g1}$</td>
<td>0.0437 (0.0321)</td>
</tr>
<tr>
<td>$\rho_{D\mu}$</td>
<td>0.0641 (0.2763)</td>
<td>$\sigma_{g2}$</td>
<td>0.0345 (0.0259)</td>
</tr>
<tr>
<td>$\rho_{\mu g1}$</td>
<td>0.9496 (0.3988)</td>
<td>$\sigma_D$</td>
<td>0.1026 (0.0181)</td>
</tr>
<tr>
<td>$\rho_{\mu g2}$</td>
<td>0.3069 (0.3829)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: Implied present-value model parameters

| $A$ | 26.82 | $\alpha$ | 0.96 |
| $B_1$ | -142.70 | $B_2$ | 15.85 |
| $B_3$ | 59.10 |        |        |

Panel C: R-squared values

| $R^2_{Returns}$ | 17.9% | $R^2_{Div}$ | 16.1% |
### Variance Decomposition

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$h_t B_1 \varepsilon_{t+1}^{\mu}$</td>
<td>269.9%</td>
<td>-117.9%</td>
<td>0.3%</td>
<td>-31.9%</td>
<td>-4.9%</td>
<td>2.3%</td>
</tr>
<tr>
<td>2</td>
<td>$h_t B_2 \varepsilon_{1,t+1}^g$</td>
<td>-117.9%</td>
<td>51.8%</td>
<td>0.2%</td>
<td>10.1%</td>
<td>2.2%</td>
<td>-1.0%</td>
</tr>
<tr>
<td>3</td>
<td>$h_t B_3 \varepsilon_{2,t+1}^g$</td>
<td>0.3%</td>
<td>0.2%</td>
<td>0.5%</td>
<td>-5.7%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>4</td>
<td>$h_t E_t (PD_{t+1}) \varepsilon_{t+1}^D$</td>
<td>-31.9%</td>
<td>10.1%</td>
<td>-5.7%</td>
<td>70.6%</td>
<td>-1.6%</td>
<td>0.6%</td>
</tr>
<tr>
<td>5</td>
<td>$h_t B_1 (\varepsilon_{t+1}^{\mu} - \sigma_{D\mu})$</td>
<td>-4.9%</td>
<td>2.2%</td>
<td>0.0%</td>
<td>-1.6%</td>
<td>4.8%</td>
<td>-2.1%</td>
</tr>
<tr>
<td>6</td>
<td>$h_t B_2 (\varepsilon_{1,t+1}^g - \varepsilon_{t+1}^D)$</td>
<td>2.3%</td>
<td>-1.0%</td>
<td>0.0%</td>
<td>0.6%</td>
<td>-2.1%</td>
<td>0.9%</td>
</tr>
<tr>
<td>7</td>
<td>$h_t B_3 (\varepsilon_{2,t+1}^g - \varepsilon_{t+1}^D)$</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.1%</td>
<td>0.1%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>
Perform a likelihood ratio test, \( LR = 2(L^1 - L^0) \)

- \( H_0 : \gamma_1 = \sigma_{g1} = \rho_{\mu g1} = 0 \) gives p value of 0.087
- \( H_0 : \delta_1 = \sigma_{\mu} = \rho_{\mu g1} = \rho_{D\mu} = 0 \) is rejected
Conclusion

- Developed a closed form present value model with time varying expected dividend growth rates and expected returns.
- Find that the Price-Dividend ratio can predict both returns and dividend growth rates.
- Decompose dividend growth rates and find a transient and persistent component.