The Dog That Did Not Bark: A Defense of Return Predictability

John Cochrane
October 23, 2006

Presented By: Michelle Zemel
September 25, 2007
Motivation

- Are stock returns predictable?
Motivation

- Are stock returns predictable?
- What is the (statistically) correct way to answer this question?
## Motivation

<table>
<thead>
<tr>
<th>Regression</th>
<th>b</th>
<th>t</th>
<th>$R^2$ (%)</th>
<th>$\sigma$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{t+1} = a + b(D_t/P_t) + \epsilon_{t+1}$</td>
<td>3.39</td>
<td>2.28</td>
<td>5.8</td>
<td>4.9</td>
</tr>
<tr>
<td>$R_{t+1} - R^f_t = a + b(D_t/P_t) + \epsilon_{t+1}$</td>
<td>3.83</td>
<td>2.61</td>
<td>7.4</td>
<td>5.6</td>
</tr>
<tr>
<td>$D_{t+1}/D_t = a + b(D_t/P_t) + \epsilon_{t+1}$</td>
<td>0.07</td>
<td>0.06</td>
<td>0.0001</td>
<td>0.001</td>
</tr>
<tr>
<td>$r_{t+1} = a_r + b_r(d_t - p_t) + \epsilon^r_{t+1}$</td>
<td>0.097</td>
<td>1.92</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>$\Delta d_{t+1} = a_d + b_d(d_t - p_t) + \epsilon^{dp}_{t+1}$</td>
<td>0.008</td>
<td>0.18</td>
<td>0.00</td>
<td>0.003</td>
</tr>
</tbody>
</table>
Main Findings

- The relationship between the dividend yield, dividend growth and return variables can be used to form more powerful statistical tests to evaluate return predictability.
The relationship between the dividend yield, dividend growth and return variables can be used to form more powerful statistical tests to evaluate return predictability.

The dividend price ratio does predict returns, while it does not predict dividend growth.
Main Findings

- The relationship between the dividend yield, dividend growth and return variables can be used to form more powerful statistical tests to evaluate return predictability.
- The dividend price ratio does predict returns, while it does not predict dividend growth.
- Long-horizon return regressions provide stronger evidence for return predictability than regressions on the one-period return.
Predictability - What are we testing?

1st order VAR system

\[ r_{t+1} = a_r + b_r(d_t - p_t) + \epsilon_{t+1}^r \quad (1) \]

\[ \Delta d_{t+1} = a_d + b_d(d_t - p_t) + \epsilon_{t+1}^d \quad (2) \]

\[ d_{t+1} - p_{t+1} = a_{dp} + \phi(d_t - p_t) + \epsilon_{t+1}^{dp} \quad (3) \]
## Estimation

<table>
<thead>
<tr>
<th></th>
<th>Estimates</th>
<th>Error Correlation and S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{b}, \hat{\phi}$</td>
<td>$\hat{\sigma}(\hat{b})$</td>
</tr>
<tr>
<td>$r$</td>
<td>0.097 0.050</td>
<td>1.92</td>
</tr>
<tr>
<td>$\triangle d$</td>
<td>0.008 0.044</td>
<td>.182</td>
</tr>
<tr>
<td>$dp$</td>
<td>.941 0.047</td>
<td>20.02</td>
</tr>
</tbody>
</table>

**Table:** Forecasting Regressions. The data was taken from the CRSP database, the sample period covers 1927-2004. The length of the period is 1 year. Coefficients were estimated using OLS. The standard errors include a GMM correction for heteroskedasticity.
Dividend yields, Dividend Growth and Returns

By the definition of a return ⇒
Dividend yields, Dividend Growth and Returns

By the definition of a return $\Rightarrow$

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{(1 + \frac{P_{t+1}}{D_{t+1}}) D_{t+1}}{P_t/D_t}$$
Dividend yields, Dividend Growth and Returns

By the definition of a return ⇒

\[ R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{(1 + \frac{P_{t+1}}{D_{t+1}}) D_{t+1}}{P_t D_t} \]

Log linearizing as in Cambell and Shiller (1988) ⇒
Dividend yields, Dividend Growth and Returns

By the definition of a return \( \Rightarrow \)

\[
R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{(1 + \frac{P_{t+1}}{D_{t+1}})}{P_t} \frac{D_{t+1}}{D_t}
\]

Log linearizating as in Cambell and Shiller (1988) \( \Rightarrow \)

\[
r_{t+1} = \log[1 + e^{(p_{t+1} - d_{t+1})}] + \triangle d_{t+1} - (p_t - d_t)
\]
Dividend yields, Dividend Growth and Returns

By the definition of a return ⇒

\[ R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{(1 + \frac{P_{t+1}}{D_{t+1}})}{P_t} \times \frac{D_{t+1}}{D_t} \]

Log linearizing as in Cambell and Shiller (1988) ⇒

\[ r_{t+1} = log[1 + e^{(p_{t+1}-d_{t+1})}] + \triangle d_{t+1} - (p_t - d_t) \]

\[ r_{t+1} \approx log[1 + \bar{P}/\bar{D}] + \frac{\bar{P}/\bar{D}}{1 + \bar{P}/\bar{D}}(p_{t+1}-d_{t+1}-(\bar{p}-\bar{d}))+\triangle d_{t+1}-(p_t-d_t) \]
By the definition of a return \( \Rightarrow \)

\[
R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{(1 + \frac{P_{t+1}}{D_{t+1}})}{\frac{P_t}{D_t}} D_{t+1}
\]

Log linearizing as in Cambell and Shiller (1988) \( \Rightarrow \)

\[
r_{t+1} = \log[1 + e^{(p_{t+1} - d_{t+1})}] + \Delta d_{t+1} - (p_t - d_t)
\]

\[
r_{t+1} \approx \log[1 + \bar{P}/\bar{D}] + \frac{\bar{P}/\bar{D}}{1 + \bar{P}/\bar{D}} (p_{t+1} - d_{t+1} - (\bar{p} - \bar{d})) + \Delta d_{t+1} - (p_t - d_t)
\]

\[
r_{t+1} \approx k + \rho (p_{t+1} - d_{t+1}) + \Delta d_{t+1} - (p_t - d_t)
\]
Important Identities

\[ r_{t+1} \approx k + \rho (p_{t+1} - d_{t+1}) + \Delta d_{t+1} - (p_t - d_t) \quad (4) \]
Important Identities

\[ r_{t+1} \approx k + \rho (p_{t+1} - d_{t+1}) + \Delta d_{t+1} - (p_t - d_t) \quad (4) \]

Iterate Forward ....
Important Identities

\[ r_{t+1} \approx k + \rho (p_{t+1} - d_{t+1}) + \Delta d_{t+1} - (p_t - d_t) \]  \hspace{1cm} (4)

Iterate Forward ....

\[ d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \]  \hspace{1cm} (5)
Relationship between the regression coefficients

Use (4) to link the regressions (1)-(3):

\[ r_{t+1} = k - \rho (d_{t+1} - p_{t+1}) + \Delta d_{t+1} + (d_t - p_t) \]
Relationship between the regression coefficients

Use (4) to link the regressions (1)-(3):

\[ r_{t+1} = k - \rho (d_{t+1} - p_{t+1}) + \Delta d_{t+1} + (d_t - p_t) \]

\[ = k - \rho^* (a_{dp} + \phi (d_t - p_t) + \epsilon_{t+1}^{dp}) + a_d + b_d (d_t - p_t) + \epsilon_{t+1}^d + (d_t - p_t) \]
Relationship between the regression coefficients

Use (4) to link the regressions (1)-(3):

\[ r_{t+1} = k - \rho (d_{t+1} - p_{t+1}) + \Delta d_{t+1} + (d_t - p_t) \]

\[ = k - \rho (a_d + \phi (d_t - p_t)) + \epsilon_{t+1}^d \]

\[ r_{t+1} = (k - \rho a_{dp} + a_d) + (-\rho \phi + b_d + 1)(d_t - p_t) + (-\rho \epsilon_{t+1}^{dp} + \epsilon_{t+1}^d) \]
Relationship between the regression coefficients

Use (4) to link the regressions (1)-(3):

\[ r_{t+1} = k - \rho (d_{t+1} - p_{t+1}) + \Delta d_{t+1} + (d_t - p_t) \]

\[ = k - \rho^* (a_{dp} + \phi (d_t - p_t) + \epsilon_{t+1}^{dp}) + a_d + b_d (d_t - p_t) + \epsilon_{t+1}^{d} + (d_t - p_t) \]

\[ r_{t+1} = (k - \rho a_{dp} + a_d) + (-\rho \phi + b_d + 1)(d_t - p_t) + (-\rho \epsilon_{t+1}^{dp} + \epsilon_{t+1}^{d}) \]

\[ = a_r + b_r (d_t - p_t) + \epsilon_r^{t+1} \]
Relationship between the regression coefficients

Use (4) to link the regressions (1)-(3):

\[ r_{t+1} = k - \rho (d_{t+1} - p_{t+1}) + \Delta d_{t+1} + (d_t - p_t) \]

\[ = k - \rho \ast (a_{dp} + \phi (d_t - p_t) + \epsilon_{t+1}^{dp}) + a_d + b_d (d_t - p_t) + \epsilon_{t+1}^d + (d_t - p_t) \]

\[ r_{t+1} = (k - \rho a_{dp} + a_d) + (-\rho \phi + b_d + 1)(d_t - p_t) + (-\rho \epsilon_{t+1}^{dp} + \epsilon_{t+1}^d) \]

\[ = a_r + b_r (d_t - p_t) + \epsilon_r^{t+1} \]

\[ \Rightarrow \]
Relationship between the regression coefficients

Use (4) to link the regressions (1)-(3):

\[ r_{t+1} = k - \rho(d_{t+1} - p_{t+1}) + \Delta d_{t+1} + (d_t - p_t) \]

\[ = k - \rho^* (a_{dp} + \phi(d_t - p_t) + \epsilon_{t+1}^{dp}) + a_d + b_d(d_t - p_t) + \epsilon_{t+1}^d + (d_t - p_t) \]

\[ r_{t+1} = (k - \rho a_{dp} + a_d) + (-\rho \phi + b_d + 1)(d_t - p_t) + (-\rho \epsilon_{t+1}^{dp} + \epsilon_{t+1}^d) \]

\[ = a_r + b_r(d_t - p_t) + \epsilon_{t+1}^r \]

\[ \Rightarrow \]
Relationship between the regression coefficients

\[ b_r = 1 - \rho \phi + b_d \]  \hspace{1cm} (6)

and

\[ \epsilon_{t+1}^r = \epsilon_{t+1}^d - \rho \epsilon_{t+1}^{dp} \]  \hspace{1cm} (7)
Forming the Null Hypothesis

Any two coefficients in the system $b_r$, $b_d$, $\phi$ uniquely determine the third coefficient. We would like to say something about the return predictability coefficient $b_r$. Let's ask about $b_r$ using what we know about $\phi$. $b_r = 0$ and $\phi = 0$. 

$\Rightarrow b_d = b_r + \rho \phi - 1 = -0.1$.

The evaluation of the statement $b_r = 0$ is equivalent in this framework to an evaluation of the null hypothesis $H_0$: $b_r = 0$ and $b_d = -0.1$. 

John Cochrane  October 23, 2006
The Dog That Did Not Bark: A Defense of Return Predictability
Forming the Null Hypothesis

- Any two coefficients in the system $b_r, b_d, \phi$ uniquely determine the third coefficient.
Forming the Null Hypothesis

- Any two coefficients in the system $b_r, b_d, \phi$ uniquely determine the third coefficient.
- We would like to say something about the return predictability coefficient $b_r$. 

\[ b_r = 0 \Rightarrow b_d = b_r + \rho_\phi - 1 = -1. \]
Any two coefficients in the system $b_r, b_d, \phi$ uniquely determine the third coefficient.

We would like to say something about the return predictability coefficient $b_r$.

Let’s ask about $b_r$ using what we know about $\phi$. 

The evaluation of the statement $b_r = 0$ is equivalent in this framework to an evaluation of the null hypothesis $H_0: b_r = 0$ and $b_d = -0.1$. 

John Cochrane  October 23, 2006
The Dog That Did Not Bark: A Defense of Return Predictability
Forming the Null Hypothesis

- Any two coefficients in the system $b_r, b_d, \phi$ uniquely determine the third coefficient.
- We would like to say something about the return predictability coefficient $b_r$.
- Let’s ask about $b_r$ using what we know about $\phi$.
- $b_r = 0$ and $\phi = 0.941 \Rightarrow b_d = b_r + \rho \phi - 1 = -0.1$. 

The evaluation of the statement $b_r = 0$ is equivalent in this framework to an evaluation of the null hypothesis $H_0$: $b_r = 0$ and $b_d = -0.1$. 

John Cochrane  October 23, 2006
The Dog That Did Not Bark: A Defense of Return Predictability
Forming the Null Hypothesis

■ Any two coefficients in the system $b_r, b_d, \phi$ uniquely determine the third coefficient

■ We would like to say something about the return predictability coefficient $b_r$

■ Let’s ask about $b_r$ using what we know about $\phi$

■ $b_r = 0$ and $\phi = 0.941 \Rightarrow b_d = b_r + \rho\phi - 1 = -0.1$

■ The evaluation of the statement $b_r = 0$ is equivalent in this framework to an evaluation of the null hypothesis $H_0 : b_r = 0$ and $b_d = -0.1$
Testing the Null Hypothesis

- We would like to test the hypothesis
  \[ H_0 : b_r = 0 \text{ and } b_d = -0.1 \]
We would like to test the hypothesis
\[ H_0 : b_r = 0 \text{ and } b_d = -0.1 \]

What is the chance of seeing the sample coefficients if our data came from the null distribution?
Testing the Null Hypothesis

- We would like to test the hypothesis
  $H_0: b_r = 0$ and $b_d = -0.1$
- What is the chance of seeing the sample coefficients if our data came from the null distribution?
- The hypothesis will be tested by simulation
Testing the Null Hypothesis

Simulation Procedure:

Simulation Procedure:

Simulate dividend growth and dividend yield and let the return follow from the identity

To create one simulated data set:

Step 1: Draw initial observation of \( d_0 - p_0 \). For \( \phi < 1 \), draw \( d_0 - p_0 \) from the unconditional distribution \( d_0 - p_0 \sim N[0, \sigma^2(\epsilon dp/(1 - \phi^2))] \). If \( \phi \geq 1 \) start at \( d_0 - p_0 = 0 \).

Step 2: Sample \( \epsilon dp \) and \( \epsilon d \) using the estimated covariance matrix of the VAR system.

Step 3: Using \( d_t - p_t \) and the random errors calculate dividend yield and dividend growth.

Step 4: Using the dividend yield and dividend growth calculate the return.
Testing the Null Hypothesis

Simulation Procedure:

- Simulate dividend growth and dividend yield and let the return follow from the identity
Testing the Null Hypothesis

Simulation Procedure:

- Simulate dividend growth and dividend yield and let the return follow from the identity
- To create one simulated data set:
Testing the Null Hypothesis

Simulation Procedure:

- Simulate dividend growth and dividend yield and let the return follow from the identity
- To create one simulated data set:
- Step 1: Draw initial observation of $d_0 - p_0$. For $\phi < 1$, draw $d_0 - p_0$ from the unconditional distribution $d_0 - p_0 \sim N[0, \sigma^2(\epsilon^{dp}/(1 - \phi^2))]$. If $\phi \geq 1$ start at $d_0 - p_0 = 0$. 
Testing the Null Hypothesis

Simulation Procedure:

- Simulate dividend growth and dividend yield and let the return follow from the identity.
- To create one simulated data set:
  - Step 1: Draw initial observation of $d_0 - p_0$. For $\phi < 1$, draw $d_0 - p_0$ from the unconditional distribution $d_0 - p_0 \sim N[0, \sigma^2(\epsilon^{dp}/(1 - \phi^2))]$. If $\phi \geq 1$ start at $d_0 - p_0 = 0$.
  - Step 2: Sample $\epsilon^{dp}$ and $\epsilon^d$ using the estimated covariance matrix of the VAR system.
Testing the Null Hypothesis

Simulation Procedure:

- Simulate dividend growth and dividend yield and let the return follow from the identity
- To create one simulated data set:
  - Step 1: Draw initial observation of $d_0 - p_0$. For $\phi < 1$, draw $d_0 - p_0$ from the unconditional distribution $d_0 - p_0 \sim N[0, \sigma^2(\epsilon^{dp}/(1 - \phi^2))]$. If $\phi \geq 1$ start at $d_0 - p_0 = 0$.
  - Step 2: Sample $\epsilon^{dp}$ and $\epsilon^d$ using the estimated covariance matrix of the VAR system
  - Step 3: Using $d_t - p_t$ and the random errors calculate dividend yield and dividend growth
Testing the Null Hypothesis

Simulation Procedure:

- Simulate dividend growth and dividend yield and let the return follow from the identity

- To create one simulated data set:

  - Step 1: Draw initial observation of \( d_0 - p_0 \). For \( \phi < 1 \), draw \( d_0 - p_0 \) from the unconditional distribution
    \( d_0 - p_0 \sim N[0, \sigma^2(\epsilon^{dp}/(1 - \phi^2))] \). If \( \phi \geq 1 \) start at \( d_0 - p_0 = 0 \).

  - Step 2: Sample \( \epsilon^{dp} \) and \( \epsilon^d \) using the estimated covariance matrix of the VAR system

  - Step 3: Using \( d_t - p_t \) and the random errors calculate dividend yield and dividend growth

  - Step 4: Using the dividend yield and dividend growth calculate the return
Testing the Null Hypothesis

Simulate 50,000 data sets
Run regressions (1)-(3) on each of the simulated samples, calculate coefficient estimates and t statistics

Using the sample distribution of coefficient estimates and t statistics, what is the probability of observing $b_d$, $b_r$ estimate pairs and t statistics more extreme than our original sample estimates?

| Real Returns | 22.3% | 10.3% | 1.77% | 1.67% |
| Excess Returns | 17.4% | 6.32% | 1.11% | 0.87% |

John Cochrane October 23, 2006
The Dog That Did Not Bark: A Defense of Return Predictability
Testing the Null Hypothesis

- Simulate 50,000 data sets
Testing the Null Hypothesis

- Simulate 50,000 data sets
- Run regressions (1)-(3) on each of the simulated samples, calculate coefficient estimates and t statistics
Testing the Null Hypothesis

- Simulate 50,000 data sets
- Run regressions (1)-(3) on each of the simulated samples, calculate coefficient estimates and t statistics
- Using the sample distribution of coefficient estimates and t statistics, what is the probability of observing \( b_d, b_r \) estimate pairs and t statistics more extreme than our original sample estimates?

<table>
<thead>
<tr>
<th></th>
<th>( b_r )</th>
<th>( t_r )</th>
<th>( b_d )</th>
<th>( t_d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real Returns</td>
<td>22.3%</td>
<td>10.3%</td>
<td>1.77%</td>
<td>1.67%</td>
</tr>
<tr>
<td>Excess Returns</td>
<td>17.4%</td>
<td>6.32%</td>
<td>1.11%</td>
<td>0.87%</td>
</tr>
</tbody>
</table>
Testing the Null Hypothesis
Recall identity (7): \( b_r = 1 - \rho \phi + b_d \)

- \( b_r \) has two interpretations:
  1) Regression coefficient on long-run returns on dividend yields
  2) Fraction of variance of dividend yields that can be attributed to time-varying expected returns and time-varying expected dividend growth
Long-run Coefficients

- Recall identity (7): \( b_r = 1 - \rho \phi + b_d \)
- Dividing through by \( 1 - \rho \phi \) yields
Recall identity (7): $b_r = 1 - \rho \phi + b_d$

Dividing through by $1 - \rho \phi$ yields

\[
\frac{b_r}{1 - \rho \phi} - \frac{b_d}{1 - \rho \phi} = 1
\]
Long-run Coefficients

- Recall identity (7): \( b_r = 1 - \rho \phi + b_d \)
- Dividing through by \( 1 - \rho \phi \) yields

\[
\frac{b_r}{1 - \rho \phi} - \frac{b_d}{1 - \rho \phi} = 1
\]

\( b_r^{lr} - b_d^{lr} = 1 \)
Long-run Coefficients

- Recall identity (7): $b_r = 1 - \rho \phi + b_d$
- Dividing through by $1 - \rho \phi$ yields

$$\frac{b_r}{1 - \rho \phi} - \frac{b_d}{1 - \rho \phi} = 1$$

$$b_r^{lr} - b_d^{lr} = 1$$

- $b_r^{lr}$ has two interpretations:
Long-run Coefficients

- Recall identity (7): \( b_r = 1 - \rho \phi + b_d \)
- Dividing through by \( 1 - \rho \phi \) yields

\[
\frac{b_r}{1 - \rho \phi} - \frac{b_d}{1 - \rho \phi} = 1
\]

\[
b_r^{lr} - b_d^{lr} = 1
\]

- \( b_r^{lr} \) has two interpretations:
- 1) Regression coefficient on long run returns on dividend yields
Recall identity (7): \[ b_r = 1 - \rho \phi + b_d \]

Dividing through by \( 1 - \rho \phi \) yields

\[
\frac{b_r}{1 - \rho \phi} - \frac{b_d}{1 - \rho \phi} = 1
\]

\[ b_{lr}^r - b_{lr}^d = 1 \]

\( b_{lr}^r \) has two interpretations:

1) Regression coefficient on long run returns on dividend yields

2) Fraction of variance of dividend yields that can be attributed to time-varying expected returns and time-varying expected dividend growth
Long-run Coefficients

\[ d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \]  \hspace{1cm} (5)
Long-run Coefficients

\[ d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \] (5)

After some manipulation ...
Long-run Coefficients

\[ d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \] (5)

After some manipulation ...

\[ \text{var}(d_t - p_t) = \text{cov}(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t) - \text{cov}(\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}, d_t - p_t) \]
Long-run Coefficients

\[ d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \quad (5) \]

After some manipulation ...

\[ \text{var}(d_t - p_t) = \text{cov}(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t) - \text{cov}(\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}, d_t - p_t) \]
Long-run Coefficients
Long-run Coefficients

\[ \text{var}(d_t - p_t) = \text{cov}(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t) - \text{cov}(\sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}, d_t - p_t) \]
Long-run Coefficients

\[
\text{var}(d_t - p_t) = \text{cov}\left(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t\right) - \text{cov}\left(\sum_{j=1}^{\infty} \rho^{j-1} \triangle d_{t+j}, d_t - p_t\right)
\]

\[
\beta\left(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t\right) - \beta\left(\sum_{j=1}^{\infty} \rho^{j-1} \triangle d_{t+j}, d_t - p_t\right) = 1
\]
Long-run Coefficients

\[ \text{var}(d_t - p_t) = \text{cov}(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t) - \text{cov}(\sum_{j=1}^{\infty} \rho^{j-1} \triangle d_{t+j}, d_t - p_t) \]

\[ \beta(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t) - \beta(\sum_{j=1}^{\infty} \rho^{j-1} \triangle d_{t+j}, d_t - p_t) = 1 \]

\[ \beta(\sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t) = \sum_{j=1}^{\infty} \rho^{j-1} \beta(r_{t+j}, d_t - p_t) = \sum_{j=1}^{\infty} \rho^{j-1} \phi^{j-1} b_r \]
Long-run Coefficients

\[ \text{var}(d_t - p_t) = \text{cov}\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right) - \text{cov}\left( \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}, d_t - p_t \right) \]

\[ \beta\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right) - \beta\left( \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}, d_t - p_t \right) = 1 \]

\[ \beta\left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}, d_t - p_t \right) = \sum_{j=1}^{\infty} \rho^{j-1} \beta(r_{t+j}, d_t - p_t) = \sum_{j=1}^{\infty} \rho^{j-1} \phi^{j-1} b_r \]

\[ = \frac{b_r}{1 - \rho \phi} = b_r^{lr} \]
## Long-run estimates and tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>$b^{lr}$</th>
<th>s.e.</th>
<th>t</th>
<th>% p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>1.09</td>
<td>0.44</td>
<td>2.48</td>
<td>1.39-1.83</td>
</tr>
<tr>
<td>$\Delta d$</td>
<td>0.09</td>
<td>.44</td>
<td>2.48</td>
<td>1.39-1.83</td>
</tr>
<tr>
<td>Excess r</td>
<td>1.23</td>
<td>0.47</td>
<td>2.62</td>
<td>0.47 - 0.69</td>
</tr>
</tbody>
</table>

**Table:** The long run coefficients are calculated from the one period coefficients. The standard errors are calculated by the delta method using the standard errors for $b^{lr}_r$, $b^{lr}_d$, $\phi$. The t statistic for $\Delta d$ is the statistic for the hypothesis $b^{lr}_d = -1$. Percent probability values were generated by Monte Carlo under the $\phi = 0.941$ null.
Long-run Coefficients

Using either of the long run coefficients, it is easy to reject the null of no return predictability. The tests on $b_{lr}$ and $b_{lr,d}$ give the same results, and there is no need to choose between return and dividend growth tests.
Using either of the long run coefficients, it is easy to reject the null of no return predictability.
Using either of the long run coefficients, it is easy to reject the null of no return predictability.

The tests on $b_{lr}^r$ and $b_{lr}^d$ give the same results, and there is no need to choose between return and dividend growth tests.
Power, $\Phi$ and negative correlation

Why are the dividend growth and long run coefficient tests more powerful?
Power, $\Phi$ and negative correlation

$\hat{b}_r$ and $\hat{\phi}$ are highly negatively correlated. High $b_{lr}$ requires high $\hat{b}_r$ and $\hat{\phi}$, which is hard under the null. Negative correlation of coefficient estimates comes from the correlation structure of the error terms.
Power, $\Phi$ and negative correlation

- $\hat{b}_r$ and $\hat{\phi}$ are highly negatively correlated
Power, $\Phi$ and negative correlation

- $\hat{b}_r$ and $\hat{\phi}$ are highly negatively correlated
- High $b^{lr}_r$ requires high $\hat{b}_r$ and $\hat{\phi}$, which is hard under the null
Power, $\Phi$ and negative correlation

- $\hat{b}_r$ and $\hat{\phi}$ are highly negatively correlated
- High $b^l_r$ requires high $\hat{b}_r$ and $\hat{\phi}$, which is hard under the null
- Negative correlation of coefficient estimates comes from the correlation structure of the error terms
Power, $\Phi$ and negative correlation

From identity (7), the error terms in this system are related as follows:

$$\epsilon_{t+1}^r = \epsilon_{t+1}^d - \rho \epsilon_{t+1}^{dp}$$
Power, $\Phi$ and negative correlation

From identity (7), the error terms in this system are related as follows:

$$\epsilon_{t+1}^r = \epsilon_{t+1}^d - \rho \epsilon_{t+1}^{dp}$$

$$\text{cov}(\epsilon_{t+1}^r, \epsilon_{t+1}^{dp}) = \text{cov}(\epsilon_{t+1}^{dp}, \epsilon_{t+1}^d) - \rho \sigma^2(\epsilon_{t+1}^{dp})$$
Power, $\Phi$ and negative correlation

From identity (7), the error terms in this system are related as follows:

$$\epsilon^r_{t+1} = \epsilon^d_{t+1} - \rho \epsilon^{dp}_{t+1}$$

$$\text{cov}(\epsilon^r_{t+1}, \epsilon^{dp}_{t+1}) = \text{cov}(\epsilon^d_{t+1}, \epsilon^d_{t+1}) - \rho \sigma^2(\epsilon^{dp}_{t+1})$$

So,

$$\text{cov}(\epsilon^{dp}_{t+1}, \epsilon^d_{t+1}) = 0 \Rightarrow$$

$$\text{cov}(\epsilon^r_{t+1}, \epsilon^{dp}_{t+1}) = -\rho \sigma^2(\epsilon^{dp}_{t+1})$$
It all comes down to $\phi$

The value of $\phi$ can be bounded by theoretical arguments
It all comes down to $\phi$

The value of $\phi$ can be bounded by theoretical arguments

$$\phi > \frac{1}{\rho} \Rightarrow \text{infinite dividend yield}$$
It all comes down to $\phi$

The value of $\phi$ can be bounded by theoretical arguments

$$\phi > \frac{1}{\rho} \Rightarrow \text{infinite dividend yield}$$

$$d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \triangle d_{t+j} + \lim_{k \to \infty} \rho^k E_t(p_{t+k}-d_{t+k})$$
It all comes down to $\phi$

The value of $\phi$ can be bounded by theoretical arguments

$$\phi > \frac{1}{\rho} \Rightarrow \text{infinite dividend yield}$$

$$d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta \Delta d_{t+j} + \lim_{k \to \infty} \rho^k E_t(p_{t+k} - d_{t+k})$$

$$\rho^k \phi^k (p_{t+k} - d_{t+k}) \text{ will explode if } \phi > \frac{1}{\rho}$$
It all comes down to $\phi$

$$\phi = \frac{1}{\rho} \Rightarrow \text{Rational Bubble}$$
It all comes down to $\phi$

$$\phi = \frac{1}{\rho} \Rightarrow \text{Rational Bubble}$$

$$E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} = \sum_{j=1}^{\infty} \rho^{j-1} \phi^{j-1} b_d(p_t - d_t)$$
It all comes down to $\phi$

$$\phi = \frac{1}{\rho} \Rightarrow \text{Rational Bubble}$$

$$E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} = \sum_{j=1}^{\infty} \rho^{j-1} \phi^{j-1} b_d (p_t - d_t)$$

For these terms to converge,

$$b_r = 0 \text{ and } b_d = 0$$
It all comes down to $\phi$

$$\phi = \frac{1}{\rho} \Rightarrow \text{Rational Bubble}$$

$$E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} = \sum_{j=1}^{\infty} \rho^{j-1} \phi^{j-1} b_d(p_t - d_t)$$

For these terms to converge,

$$b_r = 0 \text{ and } b_d = 0$$

Recalling that,

$$d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} + \lim_{k \to \infty} \rho^k E_t(p_t - d_t)$$
It all comes down to $\phi$

$$\phi = \frac{1}{\rho} \Rightarrow \text{Rational Bubble}$$

$$E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} = \sum_{j=1}^{\infty} \rho^{j-1} \phi^{j-1} b_d (p_t - d_t)$$

For these terms to converge,

$$b_r = 0 \text{ and } b_d = 0$$

Recalling that,

$$d_t - p_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} + \lim_{k \to \infty} \rho^k E_t (p_t - d_t)$$

The dividend yield in this case will depend only on the last term.
It all comes down to $\phi$

$$\phi = 1 \Rightarrow \text{dividend yield follows a random walk}$$
It all comes down to $\phi$

$\phi = 1 \implies \text{dividend yield follows a random walk}$

- The dividend yield does pass standard unit root tests, but statistical evidence is not compelling
It all comes down to $\phi$

$\phi = 1 \Rightarrow$ dividend yield follows a random walk

- The dividend yield does pass standard unit root tests, but statistical evidence is not compelling
- A random walk in dividend yields generates far more variation than we have seen over the history of stock trading
It all comes down to $\phi$

$$\phi = 1 \Rightarrow \text{dividend yield follows a random walk}$$

- The dividend yield does pass standard unit root tests, but statistical evidence is not compelling.
- A random walk in dividend yields generates far more variation than we have seen over the history of stock trading.
- If the dividend yield is not stationary, than either returns or dividend growth are not stationary.
It all comes down to $\phi$

<table>
<thead>
<tr>
<th>Null $\phi$</th>
<th>$b_r$</th>
<th>$b_d$</th>
<th>$b^l_{min}$</th>
<th>$b^l_{max}$</th>
<th>$b_r$</th>
<th>$b_d$</th>
<th>$b^l_{min}$</th>
<th>$b^l_{max}$</th>
<th>$\sigma(dp)$</th>
<th>$1/2$ life</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>24</td>
<td>0.6</td>
<td>0.3</td>
<td>0.6</td>
<td>19</td>
<td>0.4</td>
<td>0.1</td>
<td>0.2</td>
<td>0.35</td>
<td>6.6</td>
</tr>
<tr>
<td>0.941</td>
<td>22</td>
<td>1.6</td>
<td>1.2</td>
<td>1.7</td>
<td>17</td>
<td>1.1</td>
<td>0.5</td>
<td>0.7</td>
<td>0.45</td>
<td>11.4</td>
</tr>
<tr>
<td>0.96</td>
<td>22</td>
<td>2.6</td>
<td>2.0</td>
<td>2.8</td>
<td>17</td>
<td>1.6</td>
<td>0.8</td>
<td>1.2</td>
<td>0.55</td>
<td>17.0</td>
</tr>
<tr>
<td>0.98</td>
<td>21</td>
<td>4.9</td>
<td>4.3</td>
<td>5.5</td>
<td>17</td>
<td>2.7</td>
<td>1.8</td>
<td>2.5</td>
<td>0.77</td>
<td>34.3</td>
</tr>
<tr>
<td>0.99</td>
<td>21</td>
<td>6.3</td>
<td>5.9</td>
<td>7.4</td>
<td>17</td>
<td>3.6</td>
<td>2.7</td>
<td>3.6</td>
<td>1.09</td>
<td>69.0</td>
</tr>
<tr>
<td>1.00</td>
<td>22</td>
<td>8.7</td>
<td>8.1</td>
<td>10</td>
<td>16</td>
<td>4.4</td>
<td>3.7</td>
<td>4.8</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>1.01</td>
<td>19</td>
<td>11</td>
<td>11</td>
<td>13</td>
<td>14</td>
<td>5.1</td>
<td>5.1</td>
<td>6.3</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>Draw $\phi$</td>
<td>23</td>
<td>1.6</td>
<td>1.4</td>
<td>1.7</td>
<td>18</td>
<td>1.1</td>
<td>0.6</td>
<td>0.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Power in long-run regression coefficients

- How do we reconcile these findings with the literature on long horizon returns?
Power in long-run regression coefficients

- How do we reconcile these findings with the literature on long horizon returns?
- Three main differences between these and typical long horizon returns:
Power in long-run regression coefficients

- How do we reconcile these findings with the literature on long horizon returns?

- Three main differences between these and typical long horizon returns:
  - Infinite vs. Finite long run returns
    \[ \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \text{ on } d_t - p_t \]
  - Implied vs. Directly Calculated
Power in long-run regression coefficients

- How do we reconcile these findings with the literature on long horizon returns?
- Three main differences between these and typical long horizon returns:
  - Infinite vs. Finite long run returns
    \[ \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \text{ on } d_t - p_t \]
  - Implied vs. Directly Calculated
  - Weighted vs. Unweighted Returns
Power in long-run regression coefficients

- How do we reconcile these findings with the literature on long horizon returns?
- Three main differences between these and typical long horizon returns:
  - Infinite vs. Finite long run returns
    \[ \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \text{ on } d_t - p_t \]
  - Implied vs. Directly Calculated
  - Weighted vs. Unweighted Returns
## Power in long-run Regression Coefficients

The table below presents the weighted and unweighted regression coefficients for the model: \[ \sum_{j=1}^{k} \rho^{j-1} r_{t+j} = a + b^{(k)}_r (d_t - p_t) + \delta_{t+k} \]

<table>
<thead>
<tr>
<th>k</th>
<th>Direct Coefficient</th>
<th>Direct P-value</th>
<th>Implied Coefficient</th>
<th>Implied P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.10</td>
<td>0.94 0.99</td>
<td>0.10 22</td>
<td>0.10 22</td>
</tr>
<tr>
<td>5</td>
<td>0.35</td>
<td>0.94 0.99</td>
<td>0.40 17</td>
<td>0.37 29</td>
</tr>
<tr>
<td>10</td>
<td>0.80</td>
<td>0.94 0.99</td>
<td>0.65 10</td>
<td>0.92 16</td>
</tr>
<tr>
<td>15</td>
<td>1.38</td>
<td>0.94 0.99</td>
<td>0.80 6.2</td>
<td>1.68 4.8</td>
</tr>
<tr>
<td>20</td>
<td>1.49</td>
<td>0.94 0.99</td>
<td>0.89 4.1</td>
<td>1.78 7.8</td>
</tr>
<tr>
<td>\infty</td>
<td>1.04</td>
<td>1.8 7.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The model can also be expressed as an unweighted regression: \[ \sum_{j=1}^{k} r_{t+j} = a + b^{(k)}_r (d_t - p_t) + \delta_{t+k} \]
Power in long-run Regression Coefficients

- The key factor is horizon
Power in long-run Regression Coefficients

- The key factor is horizon
- A short finite horizon does not capture all the power of $\phi$
Power in long-run Regression Coefficients

- The key factor is horizon
- A short finite horizon does not capture all the power of $\phi$
- To see why, consider a short finite return

$$b_r^k = (1 + \phi \rho + \phi^2 \rho^2 + \ldots + \phi^k \rho^k)$$
Power in long-run Regression Coefficients

- The key factor is horizon
- A short finite horizon does not capture all the power of $\phi$
- To see why, consider a short finite return

\[
b_r^k = (1 + \phi \rho + \phi^2 \rho^2 + \ldots + \phi^k \rho^k)
\]

- Longer horizons give more weight to $\phi$, generating more power
Power in long-run Regression Coefficients
Out of Sample $R^2$

- Goyal and Welch (2005) show that the dividend yield does not do well forecasting out of sample.
Goyal and Welch (2005) show that the dividend yield does not do well forecasting out of sample.

The Goyal and Welch statistic is calculated as follows:

1. Run a regression from time 1 to time T, estimate the return predictability coefficient.
2. Use the coefficients estimated on time 1 to T to forecast the return in period T + 1.
3. Compute the sample mean return from time 1 to T.
4. Use the sample mean calculated on time 1 to T to forecast the return in period T + 1.
5. Compare the MSE of the two strategies.

Goyal and Welch find that the out of sample MSE is higher for the return forecast than for the sample mean.
Goyal and Welch (2005) show that the dividend yield does not do well forecasting out of sample. The Goyal and Welch statistic is calculated as follows:

- Run a regression from time 1 to time T, estimate the return predictability coefficient
- Use the coefficients estimated on time 1 to T to forecast the return in period T + 1
- Compute the sample mean return from time 1 to T
- Use the sample mean calculated on time 1 to T to forecast the return in period T + 1

Goyal and Welch find that the out of sample MSE is higher for the return forecast than for the sample mean.
Goyal and Welch (2005) show that the dividend yield does not do well forecasting out of sample. The Goyal and Welch statistic is calculated as follows:

- Run a regression from time 1 to time T, estimate the return predictability coefficient.
- Use the coefficients estimated on time 1 to T to forecast the return in period T + 1.
Goyal and Welch (2005) show that the dividend yield does not do well forecasting out of sample.

The Goyal and Welch statistic is calculated as follows:

- Run a regression from time 1 to time $T$, estimate the return predictability coefficient.
- Use the coefficients estimated on time 1 to $T$ to forecast the return in period $T + 1$.
- Compute the sample mean return from time 1 to $T$.

Comparing the two strategies, Goyal and Welch find that the out of sample MSE is higher for the return forecast than for the sample mean.
Goyal and Welch (2005) show that the dividend yield does not do well forecasting out of sample.

The Goyal and Welch statistic is calculated as follows:
- Run a regression from time 1 to time T, estimate the return predictability coefficient.
- Use the coefficients estimated on time 1 to T to forecast the return in period T + 1.
- Compute the sample mean return from time 1 to T.
- Use the sample mean calculated on time 1 to T to forecast the return in period T + 1.

Goyal and Welch find that the out of sample MSE is higher for the return forecast than for the sample mean.
Goyal and Welch (2005) show that the dividend yield does not do well forecasting out of sample.

The Goyal and Welch statistic is calculated as follows:

- Run a regression from time 1 to time T, estimate the return predictability coefficient.
- Use the coefficients estimated on time 1 to T to forecast the return in period T + 1.
- Compute the sample mean return from time 1 to T.
- Use the sample mean calculated on time 1 to T to forecast the return in period T + 1.
- Compare the MSE of the two strategies.
Out of Sample $R^2$

- Goyal and Welch (2005) show that the dividend yield does not do well forecasting out of sample.
- The Goyal and Welch statistic is calculated as follows:
  - Run a regression from time 1 to time T, estimate the return predictability coefficient.
  - Use the coefficients estimated on time 1 to T to forecast the return in period $T + 1$.
  - Compute the sample mean return from time 1 to T.
  - Use the sample mean calculated on time 1 to T to forecast the return in period $T + 1$.
  - Compare the MSE of the two strategies.
  - Goyal and Welch find that the out of sample MSE is higher for the return forecast than for the sample mean.
What is the chance of seeing such forecasting results if returns really are predictable?
Out of Sample $R^2$

- What is the chance of seeing such forecasting results if returns really are predictable?

- Simulate the data with $b_r = 1 - \rho \phi$ and $b_d = 0$
Out of Sample $R^2$

- What is the chance of seeing such forecasting results if returns really are predictable?
- Simulate the data with $b_r = 1 - \rho\Phi$ and $b_d = 0$
- Run the out of sample forecasts and calculate the Goyal and Welch statistic
Out of Sample $R^2$

- What is the chance of seeing such forecasting results if returns really are predictable?
- Simulate the data with $b_r = 1 - \rho \phi$ and $b_d = 0$
- Run the out of sample forecasts and calculate the Goyal and Welch statistic
- 30 - 40% of the draws show even worse results than the sample results
Out of Sample $R^2$

- What is the chance of seeing such forecasting results if returns really are predictable?
- Simulate the data with $b_r = 1 - \rho \phi$ and $b_d = 0$
- Run the out of sample forecasts and calculate the Goyal and Welch statistic
- 30 - 40% of the draws show even worse results than the sample results
- Poor $R^2$ cannot be used to reject the null hypothesis that returns are predictable
Out of Sample $R^2$

- What is the chance of seeing such forecasting results if returns really are predictable?
- Simulate the data with $b_r = 1 - \rho \phi$ and $b_d = 0$
- Run the out of sample forecasts and calculate the Goyal and Welch statistic
- 30 - 40% of the draws show even worse results than the sample results
- Poor $R^2$ cannot be used to reject the null hypothesis that returns are predictable
- However, these regressions are not likely to be useful in forming real-time forecasts because of the difficulty in estimating accurate coefficients in our short sample
The absence of dividend growth predictability provides strong evidence that returns are predictable.
The absence of dividend growth predictability provides strong evidence that returns are predictable.

There is even stronger evidence for predictability when long horizon returns are considered.
The absence of dividend growth predictability provides strong evidence that returns are predictable.

There is even stronger evidence for predictability when long horizon returns are considered.

The long run coefficient tests capture the three variables in one number and provide a powerful test.
Conclusion

- The absence of dividend growth predictability provides strong evidence that returns are predictable.
- There is even stronger evidence for predictability when long horizon returns are considered.
- The long run coefficient tests capture the three variables in one number and provide a powerful test.
- This testing framework addresses the issue of how to evaluate the question of return predictability; it does not suggest that this is the best model for forecasting.
Approximation Error:

Formulation of Null Hypothesis is based on the approximate linear relationship between dividend yields, dividend growth and returns. If the approximation does not hold, then the implication that $b_r = 0 \Rightarrow b_d = c$ breaks down.
Discussion

- Approximation Error:
- Formulation of Null Hypothesis is based on the approximate linear relationship between dividend yields, dividend growth and returns
Discussion

- Approximation Error:
- Formulation of Null Hypothesis is based on the approximate linear relationship between dividend yields, dividend growth and returns
- If the approximation does not hold then the implication that $b_r = 0 \Rightarrow b_d = c$ breaks down
Discussion

- Approximation Error:
- Formulation of Null Hypothesis is based on the approximate linear relationship between dividend yields, dividend growth and returns
- If the approximation does not hold then the implication that $b_r = 0 \Rightarrow b_d = c$ breaks down
Discussion

- Evaluation of Approximation Error
Discussion

- Evaluation of Approximation Error
- Using the data of the study, I created a series of return approximations using identity (4).
Discussion

- Evaluation of Approximation Error

  Using the data of the study, I created a series of return approximations using identity (4).

  I regressed the actual returns on the approximate returns to evaluate the accuracy of the approximation.
Discussion

The Dog That Did Not Bark: A Defense of Return Predictability
Discussion

The role of $\phi$ and stationnarity
The role of $\phi$ and stationnarity

Results are sensitive to the value of $\phi$ and the estimate of $\phi$ may be biased
Discussion

- The role of $\phi$ and stationarity
- Results are sensitive to the value of $\phi$ and the estimate of $\phi$ may be biased
- Cochrane claims that the stationarity of dividend yields implies predictability in either dividend growth or returns
Discussion

- The role of $\phi$ and stationnarity
- Results are sensitive to the value of $\phi$ and the estimate of $\phi$ may be biased
- Cochrane claims that the stationnarity of dividend yields implies predictability in either dividend growth or returns
- However, if the dividend yield is not stationnary, this relationship breaks down
The role of $\phi$ and stationnarity
Results are sensitive to the value of $\phi$ and the estimate of $\phi$ may be biased
Cochrane claims that the stationnarity of dividend yields implies predictability in either dividend growth or returns
However, if the dividend yield is not stationnary, this relationship breaks down
If for example, $\phi = 1$ then the stationnarity argument breaks down
Discussion

- The role of $\phi$ and stationnarity
- Results are sensitive to the value of $\phi$ and the estimate of $\phi$ may be biased
- Cochrane claims that the stationnarity of dividend yields implies predictability in either dividend growth or returns
- However, if the dividend yield is not stationnary, this relationship breaks down
- If for example, $\phi = 1$ then the stationnarity argument breaks down
- The implications are the you cannot estimate the covariance matrix from which you generate the parameters for the simulation