The Restrictions on Predictability Implied by Rational Asset Pricing Models

Chris Kirby
Introduction

• Return predictability has been documented in regression-based empirical studies (e.g. Keim and Stambaugh (1986), Fama and French (1989)), but it is not clear what the source of predictability is:
  – market inefficiency?
  – rational variation in expected returns?
  – data mining?
• Rational asset pricing models impose restrictions on the intercept, slope coefficients and $R^2$ from predictive regressions
• These restrictions provide a way to assess whether the predictability uncovered using regression analysis is consistent with rational pricing
Related Literature

• Hansen and Singleton (1983) - use a representative agent model with CRRA utility to derive restrictions on the joint process for consumption growth and asset returns
  – the predictable component of asset returns is proportional to the predictable component of consumption growth
• Ferson and Harvey (1991) - use a multi-beta CAPM to decompose the variance of the fitted values from a regression of returns on a set of instrumental variables into explained and unexplained components
  – conclude that return predictability is driven by rational variation in expected returns
  – errors-in-the-variables bias
• Hansen and Jagannathan (1991) – derive a lower bound for the variance of admissible sdf's
• Hansen and Jagannathan (1997) – use the least-squares distance between a candidate sdf and the minimum-variance sdf as a measure of model misspecification
Model

- There are no transaction costs, short-sale constraints and other barriers to trade
- Allow for heterogenous investors and nonstandard preferences
- Each investor solves a dynamic portfolio problem in order to arrive at an optimal holding of securities
- FOC: \( E[m_t R_t | \mathcal{F}_{t-1}] = 1 \) \( E[\tilde{m}_t r_i | \mathcal{F}_{t-1}] = 0 \) \( E[\tilde{m}_t r_i z_{t-1}] = 0 \)
  - This equation does not require that the returns themselves be unpredictable, except when \( m_t \) is constant (risk-neutrality)
- Restrictions on predictability given by: \( \text{cov}(r_t, z_{t-1}) = -\text{cov}(\tilde{m}_t, r_t(z_{t-1} - \mu_z)) \)
  - \( \text{cov}(r_t, z_{t-1}) = E[r_t(z_{t-1} - \mu_z)] \) is the ability to predict \( r_t \), as measured by its unconditional covariance with \( z_{t-1} \) is the same as the expected excess payoff on a dynamic trading strategy that exploits the information conveyed by the realization of \( (z_{t-1} - \mu_z) \)
  - If investors are rational, then the ability to predict \( r_t \) must be consistent with the exposure to systematic risk that an investor takes on by following this trading rule
Restrictions for Predictive Regressions

- The predictable variation in excess returns is typically measured in terms of the slope coefficients and $R^2$ from a multiple regression of the form $r_t = x_{t-1}^r b + e_t$

  - **unrestricted coefficients**
    \[ b = E[x_{t-1}x_{t-1}']^{-1} E[r_t x_{t-1}] \]

  - **unrestricted $R^2$**
    \[ R^2 = \left( \frac{b'_z \Sigma_{zz} b_z}{\sigma_r^2} \right) \]

  - **restricted coefficients**
    \[ b = -E[x_{t-1}x_{t-1}']^{-1} \text{cov}(\hat{m}_t, r_t x_{t-1}) \]

  - **restricted $R^2$**
    \[ R^2 = \left( \frac{\sigma'_{m,rz} \Sigma_{zz}^{-1} \sigma_{m,rz}}{\sigma_r^2} \right) \]
    \[ \sigma_{m,rz} \equiv \text{cov}(\hat{m}_t, r_t (z - \mu_z)) \]

- These restrictions summarize the implications of market efficiency for predictive regressions
- They make it easy to assess whether the evidence of predictability uncovered using regression analysis and a given set of instruments is consistent with any given specification for the sdf
Data and Econometric Approach

• Use returns on the 10 capitalization-based decile portfolios of NYSE stocks for the period 1963-1991 (342 observations)

• Five instrumental variables:
  – $x_{rew}$ – the excess return on the equally weighted NYSE index
  – $jan$ – a dummy variable for January
  – $term$ – the 1-month return from holding a 90-day Treasury bill less the return on a 30-day bill
  – $prem$ – the yield on Moody’s Baa rated bonds less the yield on Moody’s Aaa rated bonds
  – $x_{div}$ - the dividend yield on the S&P 500 stock index less the return on a 30-day Treasury bill

• Approach:
  – Select a candidate sdf and estimate a set of predictive regressions under the restrictions it implies
  – Test whether the coefficients from these restricted regressions are significantly different from those obtained via OLS
Candidate sdf

- Consumption-based specifications

  - Standard power utility:
    \[
    m_t^C = \frac{1}{1 + \delta} \left( \frac{C_t}{C_{t-1}} \right)^{-\gamma}
    \]

  - Habit persistence: eg Abel (1990) – each agent’s utility depends on his level of consumption relative to some time-varying benchmark
    • the model can generate large equity risk premiums and substantial time-series variation in conditional expected rates of return

    \[
    m_t^C = \frac{1}{1 + \delta} \left( \frac{C_t}{C_{t-1}} \right)^{-\gamma} \frac{\left( \frac{C_{t-1}}{C_{t-2}} \right)^{1-\gamma}}
    \]

  - Epstein and Zin (1991) preferences: current utility depends both on current consumption and the certainty equivalent of future lifetime utility

    \[
    m_t^C = \left[ \frac{1}{1 + \delta} \left( \frac{C_t}{C_{t-1}} \right)^{p-1} \right]^{\alpha/p} \left( \frac{1}{R_{ml}} \right)^{1-\alpha/p}
    \]

  - \( \alpha = 0 \) → log utility
  - \( \alpha = p \) → power utility
Candidate sdfs (cont’d)

• Linear factor specifications
  – Conditional CAPM
    \[ m^C_t = 1 - \lambda_{m,t-1} r_{mt} \]
    where \( \lambda_{m,t-1} \equiv \frac{E[r_{mt} | F_{t-1}]}{E[r^2_{mt} | F_{t-1}]} \)
    
    • Constant price of risk specification: empirically implausible
    
    • Constant beta specification: cannot use the sdf representation nor the HJ (1997) distance measure
      \[ b_m = E[x_{t-1}'x'_{t-1}]^{-1}E[r_{mt}x_{t-1}] \]
      
      • The predictable component of the market return is taken as given, so the tests could have low power to reject
        the restrictions of the conditional CAPM
  
  – Conditional Fama and French (1993) model
    \[ m^C_t = 1 - \lambda_{m,t} 1 r_{mt} - \lambda_{s,t} 1 r_{st} - \lambda_{o,t} 1 r_{ot} \]
    
    • Constant price of risk specification
    \[ b = \beta_m b_m + \beta_s b_s + \beta_o b_o \]
    
    • Constant beta specification
Estimation and Tests

- GMM estimation (for log utility specification)

\[
\begin{align*}
  h(y_{it}, \theta_i) &= \begin{pmatrix}
    m_t^c - \mu m_c \\
    (r_{it} - x_{i-1}^' \theta_{iu}) x_{t-1} \\
    (-r_{it}(m_t^c - \mu m_c) - \mu m_c x_{i-1}^' \theta_{ir}) x_{t-1}
  \end{pmatrix} \\
  \sqrt{T}(\hat{\theta}_i - \theta_i) &\to_d N(0, (D_i' S_i^{-1} D_i)^{-1}),
\end{align*}
\]

where

\[
D_i \equiv E \left[ \frac{\partial h(y_{it}, \theta_i)}{\partial \theta_i'} \right] \quad \text{and} \quad S_i \equiv \sum_{j=-\infty}^{\infty} E[h(y_{it}, \theta_i)h(y_{i,t-j}, \theta_i)'].
\]

- Wald statistic for testing \( \theta_{iu} = \theta_{ir} \) (equivalent to the HJ bound under risk-neutrality)

\[
W_{iT} = T(\hat{\theta}_{iu} - \hat{\theta}_{ir})'(Q' \Omega_{ib} Q)^{-1}(\hat{\theta}_{iu} - \hat{\theta}_{ir})
\]

- Hansen and Jagannathan distance measure

\[
\hat{d}_{iT} = [(\hat{\theta}_{ir} - \hat{\theta}_{iu})' \hat{\Sigma}_{ix}^{-1} \hat{\Sigma}_{ix}(\hat{\theta}_{ir} - \hat{\theta}_{iu})]^{\frac{1}{2}} \quad \Sigma_{ix} \equiv E[r_{it}^2(x_{t-1} x_{t-1}')] \quad G \equiv E[x_{t-1} x_{t-1}']
\]
Empirical Evidence

- Power utility

• Similar results for the habit persistence and recursive utility specifications, CAPM with constant price of risk
Empirical Evidence

- Log utility

 Panel A: Individual portfolio tests

<table>
<thead>
<tr>
<th>Decile</th>
<th>$R^2$ (%)</th>
<th>$\hat{R}^2$</th>
<th>$\Delta \hat{b}_0$</th>
<th>$\Delta \hat{b}_1$</th>
<th>$\Delta \hat{b}_2$</th>
<th>$\Delta \hat{b}_3$</th>
<th>$\Delta \hat{b}_4$</th>
<th>$\Delta \hat{b}_5$</th>
<th>$t_{\Delta \hat{b}_0}$</th>
<th>$t_{\Delta \hat{b}_1}$</th>
<th>$t_{\Delta \hat{b}_2}$</th>
<th>$t_{\Delta \hat{b}_3}$</th>
<th>$t_{\Delta \hat{b}_4}$</th>
<th>$t_{\Delta \hat{b}_5}$</th>
<th>Wald test and HI distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.3</td>
<td>0.0</td>
<td>-1.19</td>
<td>13.55</td>
<td>8.46</td>
<td>8.86</td>
<td>25.59</td>
<td>8.89</td>
<td>-1.39</td>
<td>1.45</td>
<td>3.98</td>
<td>3.02</td>
<td>2.28</td>
<td>3.73</td>
<td>65.1</td>
</tr>
<tr>
<td>2</td>
<td>18.1</td>
<td>0.0</td>
<td>-1.25</td>
<td>8.74</td>
<td>6.22</td>
<td>9.09</td>
<td>27.09</td>
<td>8.87</td>
<td>-1.57</td>
<td>1.30</td>
<td>3.65</td>
<td>3.27</td>
<td>2.78</td>
<td>3.97</td>
<td>68.5</td>
</tr>
<tr>
<td>3</td>
<td>16.1</td>
<td>0.0</td>
<td>-1.19</td>
<td>9.78</td>
<td>4.92</td>
<td>8.16</td>
<td>26.29</td>
<td>8.33</td>
<td>-1.60</td>
<td>1.67</td>
<td>3.19</td>
<td>3.04</td>
<td>2.84</td>
<td>4.15</td>
<td>69.9</td>
</tr>
<tr>
<td>4</td>
<td>14.3</td>
<td>0.0</td>
<td>-0.82</td>
<td>7.84</td>
<td>3.68</td>
<td>8.69</td>
<td>23.98</td>
<td>8.44</td>
<td>-1.13</td>
<td>1.37</td>
<td>2.55</td>
<td>3.44</td>
<td>2.70</td>
<td>4.32</td>
<td>75.0</td>
</tr>
<tr>
<td>5</td>
<td>14.0</td>
<td>0.0</td>
<td>-1.03</td>
<td>7.05</td>
<td>3.12</td>
<td>9.01</td>
<td>24.08</td>
<td>7.90</td>
<td>-1.48</td>
<td>1.31</td>
<td>2.37</td>
<td>3.52</td>
<td>2.01</td>
<td>4.27</td>
<td>67.1</td>
</tr>
<tr>
<td>6</td>
<td>13.0</td>
<td>0.0</td>
<td>-0.91</td>
<td>4.47</td>
<td>2.65</td>
<td>9.26</td>
<td>23.01</td>
<td>7.63</td>
<td>-1.38</td>
<td>0.86</td>
<td>2.06</td>
<td>3.70</td>
<td>2.94</td>
<td>4.38</td>
<td>63.0</td>
</tr>
<tr>
<td>7</td>
<td>12.5</td>
<td>0.0</td>
<td>-1.02</td>
<td>3.52</td>
<td>1.65</td>
<td>8.78</td>
<td>25.08</td>
<td>7.95</td>
<td>-1.55</td>
<td>1.55</td>
<td>3.24</td>
<td>3.10</td>
<td>4.65</td>
<td>57.7</td>
<td>0.000</td>
</tr>
<tr>
<td>8</td>
<td>10.7</td>
<td>0.1</td>
<td>-0.67</td>
<td>-0.21</td>
<td>1.54</td>
<td>9.12</td>
<td>20.11</td>
<td>7.24</td>
<td>-1.00</td>
<td>-0.04</td>
<td>1.28</td>
<td>3.60</td>
<td>2.60</td>
<td>4.31</td>
<td>51.1</td>
</tr>
<tr>
<td>9</td>
<td>10.1</td>
<td>0.1</td>
<td>-0.75</td>
<td>-0.12</td>
<td>0.95</td>
<td>9.01</td>
<td>18.61</td>
<td>6.55</td>
<td>-1.26</td>
<td>-0.02</td>
<td>0.85</td>
<td>3.22</td>
<td>2.52</td>
<td>4.16</td>
<td>43.0</td>
</tr>
<tr>
<td>10</td>
<td>9.3</td>
<td>0.1</td>
<td>-0.53</td>
<td>-1.74</td>
<td>0.31</td>
<td>7.83</td>
<td>15.65</td>
<td>6.08</td>
<td>-1.09</td>
<td>-0.32</td>
<td>0.32</td>
<td>2.56</td>
<td>2.38</td>
<td>4.52</td>
<td>35.2</td>
</tr>
</tbody>
</table>

Panel B: Joint test using all portfolios

<table>
<thead>
<tr>
<th>$W_T$</th>
<th>p-value</th>
<th>$\hat{d}_T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>149.7</td>
<td>0.000</td>
<td>0.484</td>
</tr>
</tbody>
</table>

- t-ratios and Wald tests are corrected for serial correlation and conditional heteroskedasticity (Andrews, 1991)
- Pricing errors are related to the market capitalization of the portfolio, suggesting that predictability is not due to market inefficiency
Empirical Evidence

- **CAPM with constant price of risk**

<table>
<thead>
<tr>
<th>Decile</th>
<th>$R^2$ (%)</th>
<th>$\Delta \beta_1$</th>
<th>$\Delta \beta_2$</th>
<th>$\Delta \beta_3$</th>
<th>$\Delta \beta_4$</th>
<th>$\Delta \beta_5$</th>
<th>$t_{\Delta \beta_j}$</th>
<th>Wald test and HJ distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.3 0.2</td>
<td>-1.28</td>
<td>16.10</td>
<td>7.94</td>
<td>8.88</td>
<td>23.24</td>
<td>8.60</td>
<td>64.1 0.000 0.318</td>
</tr>
<tr>
<td>2</td>
<td>18.1 0.2</td>
<td>-1.31</td>
<td>11.02</td>
<td>5.82</td>
<td>9.14</td>
<td>24.89</td>
<td>8.67</td>
<td>63.5 0.000 0.310</td>
</tr>
<tr>
<td>3</td>
<td>16.1 0.2</td>
<td>-1.28</td>
<td>11.96</td>
<td>4.57</td>
<td>8.19</td>
<td>24.33</td>
<td>8.17</td>
<td>62.2 0.000 0.316</td>
</tr>
<tr>
<td>4</td>
<td>14.3 0.2</td>
<td>-0.91</td>
<td>9.95</td>
<td>3.35</td>
<td>8.66</td>
<td>22.21</td>
<td>8.33</td>
<td>63.2 0.000 0.320</td>
</tr>
<tr>
<td>5</td>
<td>14.0 0.2</td>
<td>-1.13</td>
<td>9.18</td>
<td>2.83</td>
<td>8.93</td>
<td>22.43</td>
<td>7.79</td>
<td>55.7 0.000 0.304</td>
</tr>
<tr>
<td>6</td>
<td>13.0 0.2</td>
<td>-1.02</td>
<td>6.49</td>
<td>2.37</td>
<td>9.19</td>
<td>21.42</td>
<td>7.52</td>
<td>51.7 0.000 0.291</td>
</tr>
<tr>
<td>7</td>
<td>12.3 0.2</td>
<td>-1.12</td>
<td>5.86</td>
<td>1.38</td>
<td>8.62</td>
<td>23.52</td>
<td>7.85</td>
<td>46.9 0.000 0.284</td>
</tr>
<tr>
<td>8</td>
<td>10.7 0.2</td>
<td>-0.77</td>
<td>1.89</td>
<td>1.20</td>
<td>8.90</td>
<td>18.65</td>
<td>7.16</td>
<td>40.7 0.000 0.262</td>
</tr>
<tr>
<td>9</td>
<td>10.1 0.3</td>
<td>-0.85</td>
<td>1.97</td>
<td>0.72</td>
<td>8.68</td>
<td>17.37</td>
<td>6.50</td>
<td>34.0 0.000 0.248</td>
</tr>
<tr>
<td>10</td>
<td>9.5 0.3</td>
<td>-0.61</td>
<td>0.07</td>
<td>0.14</td>
<td>7.52</td>
<td>14.49</td>
<td>6.06</td>
<td>31.0 0.000 0.259</td>
</tr>
</tbody>
</table>

**Panel B: Joint test using all portfolios**

$W_T$  $p$-value  $d_T$
---
173.1  0.000  0.481

- **Fama and French with constant price of risk**

<table>
<thead>
<tr>
<th>Decile</th>
<th>$R^2$ (%)</th>
<th>$\Delta \beta_1$</th>
<th>$\Delta \beta_2$</th>
<th>$\Delta \beta_3$</th>
<th>$\Delta \beta_4$</th>
<th>$\Delta \beta_5$</th>
<th>$t_{\Delta \beta_j}$</th>
<th>Wald test and HJ distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20.3 4.5</td>
<td>-1.19</td>
<td>19.11</td>
<td>3.57</td>
<td>10.25</td>
<td>13.26</td>
<td>6.34</td>
<td>46.3 0.000 0.282</td>
</tr>
<tr>
<td>2</td>
<td>18.1 3.5</td>
<td>-1.37</td>
<td>13.63</td>
<td>2.39</td>
<td>10.18</td>
<td>18.71</td>
<td>6.90</td>
<td>46.6 0.000 0.277</td>
</tr>
<tr>
<td>3</td>
<td>16.1 2.9</td>
<td>-1.41</td>
<td>14.51</td>
<td>1.72</td>
<td>9.24</td>
<td>19.68</td>
<td>6.57</td>
<td>52.8 0.000 0.289</td>
</tr>
<tr>
<td>4</td>
<td>14.3 2.3</td>
<td>-1.05</td>
<td>11.66</td>
<td>0.98</td>
<td>9.47</td>
<td>18.98</td>
<td>6.98</td>
<td>51.5 0.000 0.294</td>
</tr>
<tr>
<td>5</td>
<td>14.0 1.9</td>
<td>-1.27</td>
<td>10.60</td>
<td>0.81</td>
<td>9.97</td>
<td>19.70</td>
<td>6.66</td>
<td>46.2 0.000 0.283</td>
</tr>
<tr>
<td>6</td>
<td>13.0 1.7</td>
<td>-1.17</td>
<td>8.06</td>
<td>0.56</td>
<td>10.17</td>
<td>19.15</td>
<td>6.47</td>
<td>46.4 0.000 0.279</td>
</tr>
<tr>
<td>7</td>
<td>12.3 1.2</td>
<td>-1.24</td>
<td>7.09</td>
<td>0.01</td>
<td>9.67</td>
<td>21.42</td>
<td>6.81</td>
<td>45.5 0.000 0.275</td>
</tr>
<tr>
<td>8</td>
<td>10.7 1.3</td>
<td>-0.89</td>
<td>3.02</td>
<td>-0.05</td>
<td>9.97</td>
<td>16.23</td>
<td>5.99</td>
<td>35.3 0.000 0.251</td>
</tr>
<tr>
<td>9</td>
<td>10.1 0.9</td>
<td>-0.97</td>
<td>2.46</td>
<td>-0.27</td>
<td>10.02</td>
<td>15.36</td>
<td>5.43</td>
<td>32.8 0.000 0.243</td>
</tr>
<tr>
<td>10</td>
<td>9.3 0.6</td>
<td>-0.70</td>
<td>0.07</td>
<td>-0.35</td>
<td>8.55</td>
<td>13.45</td>
<td>4.96</td>
<td>27.7 0.000 0.228</td>
</tr>
</tbody>
</table>

**Panel B: Joint test using all portfolios**

$W_T$  $p$-value  $d_T$
---
224.5  0.000  0.474
Conclusion

• Return predictability is often measured by regressing returns on a set of predetermined instrumental variables

• Rational asset pricing models impose restrictions on predictive regression parameters

• The empirical analysis indicates that returns are too predictable to be consistent with either consumption-based or linear factor models

• However, cross-sectional differences in predictability are reasonably consistent with market efficiency
Comments

• Potential data mining – Use out-of-sample predictions?
  – Eg evidence that the dividend yield fails to work in post-1990 data (Goyal and Welch, 2003, Schwert, 2003)
• Time-series properties of the instruments?
• Small sample bias?
• Survivorship?
• Regime changes?
• Using a correlation-based grouping into basis assets (Ahn et al, 2007) could give a more powerful test of the Fama and French model and of constant-beta models