Stock Return Predictability: comparing Macro- and Micro-Approaches

Arthur Stalla-Bourdillon

Université Paris Dauphine & Banque de France

09 July 2021
Motivation

• Efficient Market Hypothesis often implies no predictability:
  \[ r_{t+1} = \alpha + \beta' X_t + u_{t+1} \]

• But aggregate returns may differ from individual ones due to diversification:

  “Modern markets show considerable micro efficiency. [But] I had hypothesized considerable macro inefficiency” (Samuelson)

• What would micro-predictability give compared with macro-predictability?

• Interpretation is sensitive, predictability can both come from:
  1. **alpha-predictability**: market inefficiencies
  2. **beta-predictability**: time-varying expected returns
Contribution

- Literature on predictability heavily focuses on macro-returns.
- Studies on micro-predictability do not report time variation (Rapach et al., 2011) or do not draw micro-macro comparisons (Chinco et al., 2019).

**Table: Comparison with Literature**

<table>
<thead>
<tr>
<th></th>
<th>Aggregate</th>
<th>Individual</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lettau and Ludvigson (2001)</td>
<td>Rapach et al. (2011)</td>
</tr>
<tr>
<td></td>
<td>van Binsbergen and Koijen (2010)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dangl and Halling (2012)</td>
<td>Paper</td>
</tr>
<tr>
<td></td>
<td>Kelly and Pruitt (2013)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Farmer et al. (2019)</td>
<td></td>
</tr>
</tbody>
</table>
1. **Methodology**
   - Three working hypotheses on return predictability.
   - Postwar US monthly excess returns. 23 models estimated.
   - Building then a metric of predictability theoretically linked only with market inefficiencies: $R^2_{\alpha,t}$.

2. **Results**
   - Raw micro-predictability is **not** structurally lower than macro-predictability ($\neq$ Samuelson).
   - Micro/macro-predictability appear to follow a model where **both** alpha- and beta-predictability are at play.
   - Decomposing return predictability into $R^2_{\alpha,t}$ and $R^2_{\beta,t}$ match the theoretical explanations.
1. Theoretical Background

- We build **3 different hypotheses** regarding the behaviour of micro/macro-predictability.
- Remember, return predictability can emerge from **alpha-predictability** or from **beta-predictability**.
- More formally, following Rapach et al. (2011):

\[
\begin{align*}
    r_{t+1} &= \alpha(X_t) + \beta'_t f_{t+1} + \epsilon_{t+1} \\
    f_{t+1} &= g(X_t) + u_{t+1}
\end{align*}
\]

- This system constitutes the basis for the 3 hypotheses.
2. $H_1$, Samuelson’s view

- Micro-inefficiencies are **arbitrated away**, and micro-efficient components are **averaged out** in the aggregate.
- Macro-inefficiencies subsist, particularly for aggregate returns.
- Macro-returns should especially be predictable in times of **elevated market inefficiencies** (speculative bubbles or recessions).
3. $H_2$, Cochrane’s view

- **Markets are efficient**, but micro/macro-predictability persist due to time-variation in expected returns.

- As micro- and macro-predictability emerge from the same phenomenon, they evolve similarly.

- The mechanism is especially at play during recessions (Henkel et al., 2011).
4. $H_3$, Third view

- Micro-returns are affected **both** by idiosyncratic efficient and inefficient components that are **averaged out** in the aggregate.
- Macro-returns are affected **both** by alpha- and beta-predictability.
- Consequently: micro-predictability bounces around macro-pred. Strong Macro-predictability during **market booms and recessions**.
Data and Models

Data:

- Postwar US monthly excess returns - K.French website.
- 25 PF returns vs. Aggregate returns.

Models:

- 23 models estimated. Econometric (DESH, AR...), forecast averaging, ML (ANN), factor models...
- Each period, chosen model is the one with the best previous Out-of-Sample performance.

Methodology:

- Models are estimated on 120-month windows (Timmermann, 2008).
- Raw predictability: Out-of-Sample $R^2$ (wrt. prevailing mean)
Disentangling Return Predictability (i)

We first estimate raw micro/macro-predictability with:

$$R^2_{os,i,t} = 1 - \sum_{i=t-n}^{t-1} \frac{(r_{i+1} - r_{i+1}^f)^2}{(r_{i+1} - \bar{r}_i)^2}$$

We then build a constrained return-forecast:

- First by forecasting risk factors $f_{t+1}$
- Then by computing:

$$r_{i,t+1}^\beta = \hat{\beta}_{i,t} f_{t+1}^f$$

All predictability stemming from the risk factors is embedded in $r_{i,t+1}^\beta$ (Rapach et al., 2011).
Disentangling Return Predictability (ii)

We can thus build estimates of alpha- and beta-predictability:

\[
R^2_{i,\beta,t} = 1 - \sum_{i=t-n}^{t-1} \frac{(r_{i+1} - r_{i+1}^\beta)^2}{(r_{i+1} - \bar{r}_i)^2}
\]

\[
R^2_{i,\alpha,t} = 1 - \sum_{i=t-n}^{t-1} \frac{(r_{i+1} - r_{i+1}^f)^2}{(r_{i+1} - r_{i+1}^\beta)^2}
\]

and show that:

\[
R^2_{i,os,t} \sim R^2_{i,\alpha,t} + R^2_{i,\beta,t}
\]

\(R^2_{i,\alpha,t}\) assesses the **extra-predictability** that can be gained beyond the exposition to predictable risk factors.
1. Raw Pred.: Individual variances > Agg. variance
2. Raw Pred.: Micro-pred. isn’t lower than macro-pred.
3. Raw Pred.: Aggregating PFs

Pooling individual raw predictability series:

- Sharply reduces the variance.
- Increases the **correlation with macro-predictability**.
4. Disentangling alpha- and beta-predictability: $R_{i,\alpha,t}^2$

- $R_{i,\alpha,t}^2$ high during “Kennedy-Johnson peak” and during the dotcom bubble.
- Relatively strong dispersion along the mean.
5. Disentangling alpha- and beta-predictability: $R_{i,\beta,t}^2$

- $\overline{R}^2_{i,\alpha,t}$ rises during the 1960-61 recession or during the GFC.
- $R^2_{i,\beta,t}$ series are less dispersed, they reflect the same mechanism.
### 6. Drivers of alpha-Predictability

\[ R^2_{i,\alpha,t} = c + \gamma'_{IE} X_{IE,t} + \gamma'_{RA} X_{RA,t} + \gamma'_{FC} X_{FC,t} + \epsilon_t \]

<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>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( ec_y_t )</td>
<td>24.0***</td>
<td>44.3***</td>
<td>24.3***</td>
<td>24.0***</td>
<td>24.0***</td>
<td>24.0***</td>
<td>24.0***</td>
<td>24.0***</td>
<td>24.0***</td>
</tr>
<tr>
<td></td>
<td>(6.8)</td>
<td>(13.0)</td>
<td>(6.9)</td>
<td>(6.8)</td>
<td>(13.0)</td>
<td>(6.9)</td>
<td>(6.8)</td>
<td>(13.0)</td>
<td>(6.9)</td>
</tr>
<tr>
<td>( p_e_t )</td>
<td></td>
<td></td>
<td></td>
<td>0.1***</td>
<td>0.1***</td>
<td>0.1***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.02)</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Yale_t )</td>
<td></td>
<td></td>
<td></td>
<td>0.2***</td>
<td>0.2***</td>
<td>0.2***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.04)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( unemp_t )</td>
<td>0.7***</td>
<td>0.6**</td>
<td>0.7***</td>
<td>0.8***</td>
<td>1.0***</td>
<td>0.7***</td>
<td>1.3***</td>
<td>1.5***</td>
<td>1.4***</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.3)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.4)</td>
<td>(0.4)</td>
<td>(0.4)</td>
</tr>
<tr>
<td>( vol_{1,t} )</td>
<td>-0.001</td>
<td></td>
<td>-0.1</td>
<td></td>
<td>0.3***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td></td>
<td>(0.1)</td>
<td></td>
<td>(0.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( LF_t )</td>
<td>0.01</td>
<td>0.002</td>
<td></td>
<td>0.1*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td></td>
<td>(0.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( Baa_t )</td>
<td>0.3</td>
<td></td>
<td></td>
<td>-0.4</td>
<td>2.1***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.8)</td>
<td></td>
<td></td>
<td>(0.8)</td>
<td>(0.7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Const.</td>
<td>3.2***</td>
<td>3.2***</td>
<td>3.2***</td>
<td>1.8*</td>
<td>3.3***</td>
<td>1.7*</td>
<td>-10.1***</td>
<td>-5.0*</td>
<td>-11.3***</td>
</tr>
<tr>
<td></td>
<td>(0.8)</td>
<td>(0.8)</td>
<td>(0.7)</td>
<td>(0.8)</td>
<td>(0.7)</td>
<td>(0.8)</td>
<td>(0.8)</td>
<td>(0.8)</td>
<td>(0.8)</td>
</tr>
</tbody>
</table>

R\(^2\) 0.2 0.2 0.2 0.2 0.2 0.2 0.4 0.4 0.4
Adj.R\(^2\) 0.2 0.2 0.2 0.2 0.2 0.2 0.4 0.4 0.4

*Note:* 
* \( p < 0.1; \) ** \( p < 0.05; \) *** \( p < 0.01 \)
### 7. Drivers of beta-Predictability

\[
R^2_{i,\beta,t} = c + \gamma'_IE X_{IE,t} + \gamma'_RA X_{RA,t} + \gamma'_FC X_{FC,t} + \epsilon_t
\]

<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>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ecy_t</td>
<td>-58.2***</td>
<td>-21.4*</td>
<td>-59.8***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(10.1)</td>
<td>(12.5)</td>
<td>(10.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pet</td>
<td></td>
<td>-0.1***</td>
<td>-0.04***</td>
<td>-0.1***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.03)</td>
<td>(0.01)</td>
<td>(0.03)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Yale_t</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-unempt</td>
<td>-0.5***</td>
<td>-1.1***</td>
<td>-0.7***</td>
<td>-0.8***</td>
<td>-1.3***</td>
<td>-0.8***</td>
<td>-1.4***</td>
<td>-1.7***</td>
<td>-1.5***</td>
</tr>
<tr>
<td></td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.2)</td>
<td>(0.3)</td>
<td>(0.3)</td>
<td>(0.3)</td>
<td>(0.3)</td>
<td>(0.3)</td>
</tr>
<tr>
<td>vol_1,_t</td>
<td>-0.02</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-LF_t</td>
<td>-0.02</td>
<td></td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baa_t</td>
<td>-1.3</td>
<td></td>
<td></td>
<td>-0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.6</td>
</tr>
<tr>
<td></td>
<td>(0.8)</td>
<td></td>
<td></td>
<td>(0.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.9)</td>
</tr>
<tr>
<td>Const.</td>
<td>-6.7***</td>
<td>-9.4***</td>
<td>-6.7***</td>
<td>-4.2***</td>
<td>-9.3***</td>
<td>-4.2***</td>
<td>7.3*</td>
<td>2.6</td>
<td>7.6**</td>
</tr>
<tr>
<td>R^2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.3</td>
<td>0.1</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
</tr>
<tr>
<td>Adj. R^2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>0.3</td>
<td>0.1</td>
<td>0.4</td>
<td>0.5</td>
<td>0.4</td>
</tr>
</tbody>
</table>

**Note:**

* p<0.1; ** p<0.05; *** p<0.01
Conclusion

Several findings **corroborate the Third view**. On the raw predictability side:

1. Micro-predictability is **not structurally lower** than macro-predictability, but exhibits a **stronger variance**.

2. Pooling the micro-predictability series yields an index that mimics the macro-predictability estimate (evidence of **diversification**).

And by further disentangling the estimates:

1. Alpha- and beta-predictability **match with their theoretical drivers** (rise during market booms and recessions).

2. Beta-predictability series are less dispersed than alpha-predictability ones as they reflect **the same mechanism**.

The alpha-predictability index appears as a theoretically based and easily updatable metric to spot irrational exuberance.
Estimated Model (i)

- \( p_{t+1} = \alpha p_t + (1 - \alpha) r_t \)
- With \( p_1 = r_1 \)

- \( p_{t+1} = \alpha (p_t + \lambda_{t-1}) + (1 - \alpha) r_t \)
- \( \alpha_t = \beta (p_{t+1} - p_t) + (1 - \beta) \lambda_{t-1} \)
- With \( p_1 = 0, f_2 = r_2 \) and \( \lambda_2 = r_2 - r_1 \)

- \( r_{t+1} = \alpha + \beta (L) r_t + u_t \)
- Number of lags chosen with the Bayesian Information Criterion

- \( r_{t+1} = \alpha + \beta (L) r_t + u_t \)
- Number of lags chosen with the Aikake Information Criterion
Estimated Model (ii)


- \( r_{t+1} = \theta'_0 \eta_t d_t + \theta'_1 \eta_t + u_{t+1} \)
- \( d_t = 1/(1 + \exp(\gamma_0 + \gamma_1(r_t - r_{t-6})) \)
- With \( \eta_t = (1, r_t)' \)


- \( r_{t+1} = \theta'_0 \eta_t d_t + \theta'_1 \eta_t + u_{t+1} \)
- \( d_t = 1/(1 + \exp(\gamma_0 + \gamma_1 r_{t-3}) \)
- With \( \eta_t = (1, r_t)' \)


- \( r_{t+1} = \theta_0 + \sum_{i=1}^n \theta_i g(\beta'_i \eta_t) + u_{t+1} \)
- With \( g \) the logistic function, \( \eta_t = (1, r_t, r_{t-1}, r_{t-2})' \) and \( n = 2 \)


- \( r_{t+1} = \theta_0 + \sum_{i=1}^{n_1} \theta_i g(\sum_{j=1}^{n_2} \beta_j g(\alpha'_j \eta_t)) + u_{t+1} \)
- With \( g \) the logistic function, \( \eta_t = (1, r_t, r_{t-1}, r_{t-2})' \), \( n_1 = 2 \) and \( n_2 = 1 \)
Estimated Model (iii)

Model 9 to Model 18, *Univariate regressions*, Goyal and Welch (2008)

- \[ r_{t+1} = \theta_0 + \theta_1 x_t + u_{t+1} \]
- With \( x_t \) (univariate) exogenous regressors

Model 19, *“Kitchen sink” regression*, Goyal and Welch (2008)

- \[ r_{t+1} = \theta_0 + \theta_1' X_t + u_{t+1} \]
- With \( X_t \) the exogenous regressors

Model 20, *“Model selection” from* Goyal and Welch (2008)

- With all the potential combinations \( X_{i,t} \), we evaluate:
- \[ r_{t+1} = \theta_{i,0} + \theta_{i,1} X_{i,t} + u_{i,t+1} \]
- At each point in time, we choose the model with the smallest out-of-sample \( R^2 \)
Estimated Model (iv)

Model 21, *Factor model from*, Kelly and Pruitt (2013)
- Only for aggregate return predictions
- With $bm_{it}$ the book-to-market ratio of portfolio $i$ and $F_t$ the estimated factor, we run the following three regressions:
  - $bm_{i,t} = \theta_{i,0} + \theta_{i,1} r_{t+1} + e_{i,t}$ (time series)
  - $bm_{i,t} = c_t + F_t \hat{\theta}_{i,1} + u_{i,t}$ (cross section)
  - $r_{t+1} = \gamma_1 + \gamma_2 \hat{F}_t + \epsilon_{i,t+1}$ (time series)

- Let $p_{j,t+1}$ the forecasts from the $J$ precedent models, we use a simple equally-weighted forecast averaging of the form:
  - $p_{t+1} = \sum_{j=1}^{J} p_{j,t+1}$

- From the $J$ precedent models (apart from Model 22), we evaluate the in-sample RMSE for each single model and take as a prediction the forecast of the model with the lowest RMSE.
Moments of the raw return predictability series

- Mean $R^2_{os, i, t}$
- Standard deviation $R^2_{os, i, t}$
- Correlation $R^2_{os, i, t}$ with $R^2_{os, t}$
Raw Predictability levels vs. Returns standard deviations
Raw Pred. standard deviations vs. Returns standard deviations
Alternative risk factors: alpha-predictability
Alternative risk factors: beta-predictability
Robustness checks: regressions

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Alpha-pred.: $R^2_{i,\alpha,t}$</th>
<th>Beta-pred.: $R^2_{i,\beta,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$pe_t$</td>
<td>$0.001^{**}$</td>
<td>$-0.001^{***}$</td>
</tr>
<tr>
<td></td>
<td>$(0.0003)$</td>
<td>$(0.0002)$</td>
</tr>
<tr>
<td>$Michigan_t$</td>
<td>$0.0004$</td>
<td>$-0.001^{***}$</td>
</tr>
<tr>
<td></td>
<td>$(0.0003)$</td>
<td>$(0.0002)$</td>
</tr>
<tr>
<td>$-unemp_t$</td>
<td>$0.008^{***}$</td>
<td>$-0.008^{***}$</td>
</tr>
<tr>
<td></td>
<td>$(0.002)$</td>
<td>$(0.002)$</td>
</tr>
<tr>
<td>$vol_t$</td>
<td>$-0.0004$</td>
<td>$0.0004$</td>
</tr>
<tr>
<td></td>
<td>$(0.001)$</td>
<td>$(0.001)$</td>
</tr>
<tr>
<td>$vol_{2,t}$</td>
<td>$-0.002$</td>
<td>$0.002$</td>
</tr>
<tr>
<td></td>
<td>$(0.002)$</td>
<td>$(0.001)$</td>
</tr>
<tr>
<td>Const.</td>
<td>$-0.074^{***}$</td>
<td>$0.017^{*}$</td>
</tr>
<tr>
<td></td>
<td>$(0.026)$</td>
<td>$(0.010)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs.</td>
<td>496</td>
<td>856</td>
<td>496</td>
<td>856</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.077</td>
<td>0.168</td>
<td>0.086</td>
<td>0.107</td>
</tr>
<tr>
<td>Adj. $R^2$</td>
<td>0.071</td>
<td>0.165</td>
<td>0.086</td>
<td>0.107</td>
</tr>
</tbody>
</table>

Note: $^*$ p<0.1; $^{**}$ p<0.05; $^{***}$ p<0.01
Working Hypothesis Systems

$H_1$ Samuelson’s view:
\[
\begin{align*}
    r_{i,t+1} &= \omega_i \alpha(X_t) + \beta'_{i,t} f_{t+1} + \epsilon_{i,t+1} + \delta_i \epsilon_{t+1} \\
    r_{t+1} &= \alpha(X_t) + \beta' f_{t+1} + \epsilon_{t+1} \\
    f_{t+1} &= c + u_{t+1}
\end{align*}
\]

$H_2$ Cochrane’s view:
\[
\begin{align*}
    r_{i,t+1} &= \beta'_{i,t} f_{t+1} + \epsilon_{i,t+1} + \delta_i \epsilon_{t+1} \\
    r_{t+1} &= \beta' f_{t+1} + \epsilon_{t+1} \\
    f_{t+1} &= g(X_t) + u_{t+1}
\end{align*}
\]

$H_3$ Third view:
\[
\begin{align*}
    r_{i,t+1} &= \alpha_i(X_t) + \omega_i \alpha(X_t) + \beta'_{i,t} f_{t+1} + \epsilon_{i,t+1} + \delta_i \epsilon_{t+1} \\
    r_{t+1} &= \alpha(X_t) + \beta' f_{t+1} + \epsilon_{t+1} \\
    f_{t+1} &= g(X_t) + u_{t+1}
\end{align*}
\]