Anomalies and Links to Market Return Predictability: Supranational Evidence Based on New Market Efficiency Measures

Xi Dong, Yan Li, Liana Yanran Li, David E Rapach and Guofu Zhou

Stock price are jointly determined in the cross-sectional and time-series dimension. But return predictability emerges in two separated strands in most existing literature.

Cross-sectional: Anomalies discover
Time-series: Market return Prediction

Can we link them together?

Little evidence of individual characteristics aggregated across firms predicting market return in US (Engelberg et al. (2023))

Yes, aggregated cross-sectional anomaly portfolio return can predict market return in US both in-sample and out-of-sample (Dong et al. (2022))

* Thus, we wonder: Any linkage in the international markets?

With 44 non-US countries and 100 anomalies:

* Little evidence in country level prediction

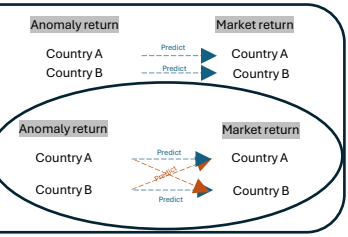
Out-of-sample R^2:

	Avg	PC	PLS	C-Enet	Enet	Combine
G6	1.4*	1.03	-1.16	-0.77	-1.67	0.46
G18	-0.75	-1.54	-5.12	-0.71	-2.65	0.11
World	-0.39	-0.65	-5.31	-0.53	-2.58	0.07

* Strong prediction power shows up once we group and aggregate country level anomaly return into **"supra-national" level** to capture mispricing that operates both within and beyond country borders

Specifically, to construct a supra-national group, we compute each anomaly and market excess return by taking a market-cap-weighted average across the countries within this group.

	Avg	PC	PLS	C-Enet	Enet	Combine
G6	5.58***	2.72**	6.92***	4.23**	2.15**	0.91**
G18	5.09***	2.45**	7.46***	3.46**	3.8**	0.79**
World	5.13***	2.49*	7.22***	1.89**	4.16**	0.79**



* How to understand this linkage ?

Consider a representative anomaly whose long and short leg return can be decomposed into a martingale increment $\boldsymbol{f_t}$ (serially uncorrelated) and a mispricing correction component $\boldsymbol{\Delta u_{L,t}}$ (serially correlated):

$$r_{L,t} = k_L f_t + \Delta u_{L,t}\,, \quad for\ l = Long$$
$$r_{S,t} = k_S f_t + \Delta u_{S,t}\,, \quad for\ l = Short$$

Thus, the long-short anomaly return is (for simplicity, assume $k_S = 1$ and $k_L = k$. Empirically, we find $k \approx 1$):

$$r_{LS,t} = (k-1)f_t + \Delta u_{L,t} - \Delta u_{S,t}$$

Suppose the market is the combination of long and short legs:

$$r_{M,t} = 0.5(k+1)f_t + 0.5(\Delta u_{L,t} + \Delta u_{S,t})$$

- Based on the framework above, the standardized coefficient for using long, short and long-short leg:
- Predictability comes from systematic mispricing correction persistence. (In other words, the autocovariance of $\Delta u_{L,t}$ component.) Due to reasons like short-selling constraints, arbitragers correct mispricing gradually.

$$\tilde{\beta}_L = \frac{0.5cov(\Delta u_{L,t+1}, \Delta u_{L,t})}{\sqrt{var(r_{M,t+1}) * \sqrt{k^2 var(f_t) + var(\Delta u_{S,t})}}}$$

$$\tilde{\beta}_S = \frac{0.5cov(\Delta u_{S,t+1}, \Delta u_{S,t})}{\sqrt{var(r_{M,t+1}) * \sqrt{var(f_t) + var(\Delta u_{S,t})}}}$$

$$\tilde{\boldsymbol{\beta}}_{\boldsymbol{LS}} = \frac{\mathbf{0.5}[cov(\Delta u_{L,t+1}, \Delta u_{L,t}) - cov(\Delta u_{S,t+1}, \Delta u_{S,t})]}{\sqrt{\boldsymbol{var(r_{M,t+1})}} * \sqrt{(k-1)^2 var(f_t) + \boldsymbol{var(\Delta u_{L,t})} + \boldsymbol{var(\Delta u_{S,t})}}}$$

(1) Comparing to $\tilde{\beta}_L$ and $\tilde{\beta}_S$, $\tilde{\beta}_{LS}$ has part of underpricing correction persistence and overpricing correction canceling out each other. For most countries, $\tilde{\beta}_{LS}$ is negative as overpricing correction persistence > underpricing correction persistence.

(2) Comparing to $\tilde{\beta}_L$ and $\tilde{\beta}_S$, $\tilde{\beta}_{LS}$ remove the large variation from the martingale increment component f_t which makes the market so hard to predict, as k approximates to 1.

* What is the driving force underlying this linkage? Let's further decompose $\tilde{\beta}_{LS}$ into the following three components:

$$\log(-\tilde{\beta}_{LS}) = \log(\tilde{\beta}_S * \left(-\frac{\tilde{\beta}_{LS}}{\tilde{\beta}_S}\right)) = \mathbf{\log(\tilde{\beta}_S)} + \mathbf{\log(-\theta_{auto})} + \mathbf{\log(\theta_{var})}$$

(1) Component 1 $\tilde{\beta}_S$: **Systematic mispricing** -- Overpricing correction within the short-leg segment of the market foreshadows future overpricing correction in the aggregate market (**inter-temporal** pricing beta)

(2) Component 2 θ_{auto} : **Asymmetric mispricing persistence** -- the degree of asymmetric mispricing persistence in the short leg vis-a-vis the long leg (negative for most countries: indicating persistent bubbles are a more impactful than persistent underpricing)

$$\theta_{auto} = \frac{\boldsymbol{cov}(\Delta u_{L,t+1}, \Delta u_{L,t})}{\boldsymbol{cov}(\Delta u_{S,t+1}, \Delta u_{S,t})} - 1$$

(3) Component 3 θ_{var} : **Price randomness** -- the degree of pricing variability coming from the unpredictable martingale increment ($var(f_t)$) relative to mispricing correction ($\mathrm{var}(\Delta u_{L,t})$) (greater price randomness indicates a more efficient market closer to random walk)

$$\theta_{var} = \frac{\sqrt{\mathrm{var}(f_t) + \mathrm{var}(\Delta u_{S,t})}}{\sqrt{(k^2-1)\mathrm{var}(f_t) + \mathrm{var}(\Delta u_{L,t}) + \mathrm{var}(\Delta u_{S,t})}} \approx \sqrt{\frac{\boldsymbol{var}(f_t)}{\boldsymbol{var}(\Delta u_{S,t})} + 1}{2}$$

* Empirically, we verified that the supra-nations constructed by countries with the 3 components falling the top half outperforms the supra-nations constructed by countries falling in the bottom half in terms of market return predictability:

All three in top half	Two in top half	One in top half	All three in bottom half
6.47***	3.86**	-0.26	-0.44

*When we group countries by their market maturity, we find long-short anomaly returns predict Developed market, while long and short legs predict Emerging market:

Supra-national group	Long-Short	Long	Short
Developed	3.24**	0.93	1.57
Emerging	1.26*	3.60**	3.69**

* The reason why Developed and Emerging markets perform differently in market return predictability is because comparing to Emerging market, Developed market has less systematic mispricing ($\tilde{\beta}_S$), but high asymmetry in mispricing persistence (θ_{auto}) and high price randomness (θ_{var}) :

	$\log(\tilde{\beta}_S)$	$\log(-\theta_{auto})$	$\log(\theta_{var})$
Developed vs Emerging	-0.41***	0.70***	0.18***

To conclude, we contribute to the linkage between the cross-sectional and time-series pricing by documenting anomalies' predictive power for the market return in the international markets at "supra-national" level, whose underlying mechanism lies in three new measures of market efficiency.