Cross-sectional Predictability

Kevin Crotty

BUSI 448: Investments

Where are we?

Last time:

- CAPM and the Market Model Regression
- CAPM: Theory
- CAPM: Practice

Today:

- The cross-section of expected stock returns
- Portfolio sorts
- Cross-sectional regression

Expected returns

Estimating a stock's expected return

- How should we think about estimating $E[r_i]$?
- Core finance: use CAPM

$$E[r_i] = r_f + \beta_i E[r_{\text{mkt}} - r_f]$$

- But we saw last time that this does poorly empirically
  - as does using historical returns
- Today, we'll discuss some firm characteristics and how they relate to expected returns

Portfolio sorts

Sorting stocks

- Consider a characteristic of a firm, like its beta.
- How can we test if high beta firms have high expected returns?
- One method:
  - sort stocks on betas
  - form portfolios (b/c stock returns are noisy)
  - test if high beta portfolios have higher ex post returns than low beta portfolios

A litle history on sorting stocks

- Market beta (Black, Jensen, Scholes 1972)
  - If anything, high beta firms *underperform* low beta
- Size (Banz 1981)
  - Small firms outperform large firms
- Book-to-market ratio (Fama-French 1992)
  - High B/M (value) outperform low B/M (growth)

Is the CAPM dead?

	All	Low-β	β-2	β-3	β-4	β-5	β-6	β-7	β-8	β-9	High-β
Panel A: Average Monthly Returns (in Percent)											
All	1.25	1.34	1.29	1.36	1.31	1.33	1.28	1.24	1.21	1.25	1.14
Small-ME	1.52	1.71	1.57	1.79	1.61	1.50	1.50	1.37	1.63	1.50	1.42
ME-2	1.29	1.25	1.42	1.36	1.39	1.65	1.61	1.37	1.31	1.34	1.11
ME-3	1.24	1.12	1.31	1.17	1.70	1.29	1.10	1.31	1.36	1.26	0.76
ME-4	1.25	1.27	1.13	1.54	1.06	1.34	1.06	1.41	1.17	1.35	0.98
ME-5	1.29	1.34	1.42	1.39	1.48	1.42	1.18	1.13	1.27	1.18	1.08
ME-6	1.17	1.08	1.53	1.27	1.15	1.20	1.21	1.18	1.04	1.07	1.02
ME-7	1.07	0.95	1.21	1.26	1.09	1.18	1.11	1.24	0.62	1.32	0.76
ME-8	1.10	1.09	1.05	1.37	1.20	1.27	0.98	1.18	1.02	1.01	0.94
ME-9	0.95	0.98	0.88	1.02	1.14	1.07	1.23	0.94	0.82	0.88	0.59
Large-ME	0.89	1.01	0.93	1.10	0.94	0.93	0.89	1.03	0.71	0.74	0.56

Source: Fama-French 1992

More history on sorting stocks

- Liquidity (Amihud and Mendelson 1986)
  - less liquid $\rightarrow$ high r_{t+1}
- Momentum (Jegadeesh-Titman 1993)
  - past winners beat past losers
- Idiosyncratic volatility (Ang, Hodrick, Xing, Zhang 2006)
  - High `idiovol` $\rightarrow$ low r_{t+1}

Visualizing anomalies

One-way sorts

Two-way sorts

Sorting in python

```
# Sorting function
def cut_quintiles(x):
    try:
        out = pd.qcut(x, 5, labels=["Lo 20", "Qnt 2", "Qnt 3", "Qnt 4", "Hi 20"])
    except:
        out = pd.Series(np.nan, index=x.index)
    return out
CHAR = 'beta'
df["quintile"] = df.groupby("date")[CHAR].apply(cut_quintiles)
```

Cross-sectional regressions

Cross-sectional regression

- An alternative approach to sorting is a cross-sectional regression

Regress each stock's average return (a time-series average) on its average characteristic:

$$\bar{r}_i = a + b \cdot \text{characteristic}_i + e_i$$

- If the characteristic is associated with higher returns, $\hat{b}$ should be different from zero!

Fama-MacBeth

- Characteristics are often time-varying, so it is preferable to use a series of cross-sectional regressions.

For each time period, run a cross-sectional regression:

$$r_{i,t} = a_t + b_t \cdot \text{characteristic}_{i,t-1} + e_{i,t}$$

- This produces a time-series of coefficients b_t. If the characteristic is associated with higher returns, the time-series average of $\hat{b}_t$ should be different from zero!

Multivariate Fama-MacBeth

- We can also characterize returns as a linear function of multiple characteristics:

For each time period, run a cross-sectional regression:

$$r_{i,t} = a_t + b_{1,t} \cdot \text{X1}_{i,t-1} + b_{2,t} \cdot \text{X2}_{i,t-1} + e_{i,t}$$

Regressions or sorts?

- Regression restricts the relation between returns and the characteristic to be linear
- Sorting allows for a nonlinear relation across stocks
  - It is more flexible, but it is reassuring if the relation is monotonically increasing or decreasing.
- Recent literature has applied various machine learning techniques to better characterize the cross-section of expected returns.

Persistence of anomalies

- What should happen as market participants become aware of these results?
- They should trade until the differences disappear.
- Anomalies that do not go away may be compensation for risk and should be accounted for as part of the asset's risk premium

$$E[r] = r_f + \text{risk premium}$$

- More on this next time…

For next time: Multi-factor Models