Mandatory Assignment 3 AEF

ncx951 & dsc579

2023-06-11

Exercise 1

Parametric Portfolio Policies is a method introduced by Brandt et al. (2009)[1] as an alternative to the traditional mean-variance approach of Markowitz (1952)[2]. The method simplifies the portfolio optimization process by leveraging firm characteristics, such as market capitalization, book-to-market ratio, or lagged return, which are related to the stock's expected return, variance, and covariance with other stocks.

The intuition behind the parameter vector θ is to capture the relationship between a stock's firm characteristics and its portfolio weight. θ represents a vector of coefficients that, when combined with the standardized firm characteristics ($\hat{x}_{i,t}$), generates the optimal portfolio weights for each stock. This linear specification allows for a more stable and less complex optimization process compared to the mean-variance approach.

Brandt et al. (2009) propose to estimate θ by maximizing the utility that would have been obtained by implementing the policy over the sample period. This is done by solving the investor's problem of choosing portfolio weights to maximize the expected utility of the portfolio return:

$$\max_{\{w_{i,t}\}_{i=1}^{N_t}} E_t[u(r_{p,t+1})] = E_t\left[u(\sum_{i=1}^{N_t} w_{i,t} r_{i,t+1})\right]$$

The portfolio weights ($w_{i,t}$) are parametrized as a function of firm characteristics:

$$w_{i,t} = \bar{w}_{i,t} + \frac{1}{N_t}\theta'\hat{x}_{i,t}$$

Here, $\bar{w}_{i,t}$ represents the weight of a benchmark portfolio (in this case, the market capitalization-weighted portfolio), and $\hat{x}_{i,t}$ are the cross-sectionally standardized firm characteristics of stock i.

Describing the portfolio policy as "active portfolio management relative to a performance benchmark" makes sense because the Parametric Portfolio Policies method adjusts the benchmark portfolio weights based on the firm characteristics. The market capitalization-weighted portfolio serves as a starting point, and the parametrized weights $\frac{1}{N_t}\theta'\hat{x}_{i,t}$ represent the active management component of the portfolio. By estimating θ and incorporating it into the portfolio weights, investors can create a portfolio that actively adjusts to changes in firm characteristics, which can potentially lead to better risk-adjusted returns compared to a passive benchmark portfolio.

---

[1] *Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns* Brandt et al. 2009
[2] *Portfolio Selection* Markowitz 1952

Exercise 2

The Parametric Portfolio Policies (PPP) method has its pros and cons compared to the traditional two-step procedure:

Pros:

1. Simplicity: The two-step procedure involves estimating the moments of the return distribution (expected returns, variances, and covariances) and then computing optimal weights based on these estimates. This can be complex and computationally intensive, especially for large portfolios. PPP simplifies the process by directly linking firm characteristics to portfolio weights, making the optimization process more straightforward.

2. Stability: The estimates of moments in the two-step procedure can be notoriously noisy and unstable, especially when dealing with a large number of assets. Direct weight parameterization in PPP leads to more stable estimates, as it leverages firm characteristics that are more easily observable and less prone to estimation errors.

3. Active management: PPP incorporates an active management component by adjusting portfolio weights based on firm characteristics. This allows investors to potentially achieve better risk-adjusted returns compared to passive benchmark portfolios.
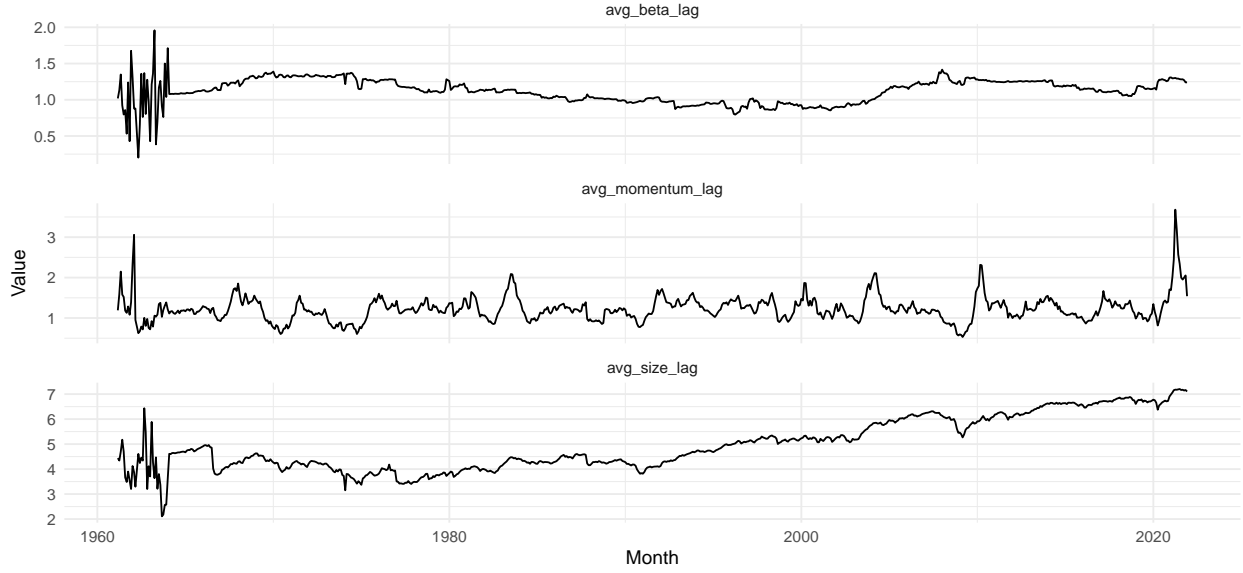
Cons:

1. Limited scope: Direct weight parameterization focuses on using firm characteristics to determine portfolio weights. This approach may exclude other relevant information or market factors that could be useful in optimizing the portfolio. In contrast, the two-step procedure attempts to capture a broader range of information through the estimation of moments.

2. Model misspecification: The linear specification of portfolio weights in the PPP method is a simplification, and it may not accurately capture the true relationship between firm characteristics and optimal portfolio weights. This can lead to suboptimal portfolios if the assumed functional form does not accurately represent the underlying relationships.

3. Estimation risk: Although the PPP method may lead to more stable estimates, it is still subject to estimation risk. The choice of firm characteristics and the estimation of the θ parameter can have a significant impact on the resulting portfolio weights, and errors in either of these components can affect portfolio performance.

Exercise 3

We retrieve monthly stock returns from the `tidy_finance.sqlite` database and compute three lagged monthly firm characteristics: `size_lag` (the log of the price per share times the number of outstanding), `beta_lag` (The CAPM-beta estimated using daily data on a rolling windows basis), and `momentum_lag` (The compounded return between months $t-2$ and $t-13$ for each firm). We will then standardize these characteristics to have zero mean and unit standard deviation across all stocks at each month for optimizing the vector $\hat{\theta}$.

The figure below displays the means of the firm characteristics `momentum_lag`, `size_lag` and `beta_lag` every month for the entire sample period from the `crsp_monthly` dataset. The means are computed across firms each month.

Figure 1: Average Values for Momentum_lag, Size_lag, and Beta_lag

Exercise 4

We now wish to estimate the parameter vector θ for an investor who aims to maximize the expected utility with a power utility function:

$$u_\gamma(r) = \frac{(1+r)^{(1-\gamma)}}{1-\gamma}$$

With $\gamma = 5$ and no constraints regarding short positions. We base the benchmark portfolio weights, $\bar{w}_{i,t}$ on the lagged market capitalization. In order to find the optimal vector $\hat{\theta}$ we use an optimizer that seeks to calibrate the weights of $\hat{\theta}$ that maximized the evaluated expected utility:

$$\max_\theta \frac{1}{T} \sum_{t=0}^{T-1} u(\theta, r_{p,t+1}) = \frac{1}{T} \sum_{t=0}^{T-1} u\left(\sum_{i=1}^{N_t} \left(\bar{w}_{i,t} + \frac{1}{N_t} \theta' \hat{x}_{i,t} \right) r_{i,t+1} \right)$$

Where U is the power utility function given above. We provide the estimated coefficients from $\hat{\theta}$ below.

$$\hat{\theta} = (0.8491 \quad -2.5774 \quad -1.9586)'$$

The resulting values of our optimal parameter choice are easy to interpret:

- The expected utility increases by tilting the weights from a portfolio with large-cap stocks toward smaller stocks (negative coefficient for size). This observation is consistent with the "size effect", where smaller stock tends to outperform larger stocks on average over long horizons. We need to remember that smaller-cap stock also comes with higher risks, so the higher return of the "size effect" observation isn't free of cost.

- We also want to tilt the weights from a portfolio with high beta stocks to low beta stocks (negative coefficient for beta). This means that stock with lower systematic risk are expected to yield higher returns compared to larger-cap stocks. This suggestion can strike as counterintuitive since we generally

3

would associate higher beta stocks with higher expected returns according to the CAPM. However it is worth noticing that there is empirical evidence that shows that low-beta stocks can outperform high-beta stocks - a phenomenon that is known as the "low-beta anamoly".

- The results also suggests that we should shift the weights toward past winners (positive coefficient for momentum), so past winners will be expected to continue performing well, which is also aligned with the well-known "momentum effect", where stocks that have outperformed in the past tends to outperform in the future.

We provide the portfolio performance below:

Table 1: Portfolio performance (full sample)

measure	benchmark	tilt
Expected utility	-0.2496	-0.2427
Average return	7.0058	15.8617
SD return	15.8293	16.8323
Sharpe ratio	0.4426	0.9423
CAPM alpha	0.0000	0.0108
Market beta	0.9920	0.4052
Absolute weight	0.0347	0.0859
Max. weight	3.7785	3.5193
Min. weight	0.0008	-0.3418
Avg. sum of negative weights	0.0000	77.8012
Avg. fraction of negative weights	0.0000	42.8195

From the summary table of the performance of the two strategies we note that the "tilt" investment strategy achieves higher expected utility and a higher Sharpe ratio of 0.942. Which is more than double the benchmark portfolio of 0.44. This suggests that deviating from the bench mark can improve risk adjusted returns for the full sample period.

Exercise 5

In the initial implementation, there might be a look-ahead bias because we used the full dataset to estimate the parameter vector $\hat{\theta}$, which includes using future information to make decisions in the past. This can lead to overfitting and overly optimistic performance measures that are not representative of the true investment risks.

In this analysis, we aim to evaluate the robustness of our parametric portfolio policy by conducting an out-of-sample experiment. This will allow us to assess the performance of the investment strategy in a more realistic scenario, where the policy is tested on data that wasn't used during the estimation process.

Inspired by Brand et al. (2009), we split our dataset into two equal sub-samples: Sample 1, which spans from March 1961 to June 1993, and Sample 2, which covers July 1993 to December 2021. We then estimate the optimal portfolio policy for each sub-sample and use the resulting policy coefficients to form portfolios in the other sub-sample. This helps us ensure that our investment strategy is tested on unseen data, reducing the risk of overfitting and providing a more reliable evaluation of its performance.

More specifically, we first estimate the optimal portfolio policy for Sample 1 by finding the θ values that maximize the expected utility of the portfolio. Next, we apply this policy to Sample 2 by forming portfolios using the firm characteristics observed at the beginning of each month and the policy coefficients estimated from Sample 1. The resulting returns from these portfolios in Sample 2 represent the out-of-sample performance for the policy estimated from Sample 1.

We then repeat this process by estimating the optimal portfolio policy for Sample 2 and applying it to Sample 1, allowing us to evaluate the out-of-sample performance for the policy estimated from Sample 2.

By conducting this out-of-sample experiment, we can assess the generalization and robustness of our parametric portfolio policy. This approach provides a more convincing evaluation of the policy's performance than simply relying on in-sample metrics, as it demonstrates how well the policy performs on data it has not encountered during the estimation process. However it should be noted that splitting the data still contains two relatively large samples, and the model could thus still be considered unstable and is still prone to overfitting.

We report the results for the θ vectors and portfolio performances based on the two subsamples below:

Table 2: Optimal theta (subsamples)

	momentum_lag	size_lag	beta_lag
Full sample	0.8491	-2.577	-1.959
Sample 1	2.7024	-3.026	-2.888
Sample 2	-1.0127	-2.778	-1.870

From the estimated θ vectors we notice that the estimated coefficients have changed compared to the full sample period. The major difference is the negative coefficient in sample 2 for *momentum_lag*, which suggests that in sample 2 (July 1993 - December 2021) stocks with lower momentum should receive more weight. This seems to be inconsistent with the "momentum effect" from the full sample and sample 1. This difference could be due to changes in market conditions or other factors during different time periods. We also note that the *momentum_lag* is estimated much higher in Sample 1 (The optimal coefficients trained on sample 1) which tells us that the momentum effect might've been much greater from 1963-1993 than the time period after.

Table 3: Portfolio perfomance (subsamples)

Measure	Sample 1 Benchmark	Sample 1 Tilt	Sample 2 Benchmark	Sample 2 Tilt
Expected utility	-0.2521	-0.2503	-0.2469	-0.2526
Average return	4.7739	13.2641	9.5315	12.5215
SD return	16.6981	23.0334	14.7753	23.0440
Sharpe ratio	0.2859	0.5759	0.6451	0.5434
CAPM alpha	-0.0004	0.0087	0.0005	0.0115
Market beta	1.0130	0.5427	0.9667	-0.1392
Absolute weight	0.0415	0.1059	0.0293	0.1036
Max. weight	4.4636	4.1577	3.2310	3.1992
Min. weight	0.0018	-0.7359	0.0000	-0.5223
Avg. sum of negative weights	0.0000	82.5781	0.0000	127.3702
Avg. fraction of negative weights	0.0000	42.5932	0.0000	45.2193

From table 4 we notice that the "tilt" strategy now seem to perform worse than using the full sample as in exercise 4. This also could also indicate the look-ahead bias previously stated. We notice that in Sample 1 the "tilt" strategy still manages to deliver higher expected utility, average returns and Sharpe ratio than the benchmark value weighted portfolio. However the "tilt" strategy also imposes fairly greater standard deviation. In sample 2 however the "tilt" strategy now delivers lower expected utility than the benchmark. It also has lower Sharpe ratio and higher volatility. It still manages to deliver higher average returns. But since the volatility is much larger than the benchmark the risk adjusted returns now perform worse.

These results tells us that using an active trading strategy is ambigous. If we estimate θ on sample 2 and apply them to sample 1 we perform better than the benchmark but vice-versa we don't. It isn't clear if basing a trading strategy based on the PPP method is desirable or not. As stated before this could be do to the data sampling periods, regime changes and other factors that the model doesn't capture. Which always seem to be issues when trying to model/forecast economic behavior.

Exercise 6

Turnover has a significant effect on how a portfolio performs, mainly through its impact on transaction costs. The more often assets within a portfolio are bought and sold (the higher the turnover), the more transaction costs you'll have. This is because each transaction comes with a cost. So, one of the strategies we want to use is to discourage these costs. In this section we want to propose a procedure that should reduce the deviation from the benchmark portfolio in an economically meaningful way and have a control over the impact of turnover on portfolio performance. Essentially, we're looking for a way to "penalize" high transaction costs in our system. If we do that, we automatically keep turnover lower because fewer transactions mean fewer costs. We would like to look at the optimization problem and draw some inspiration from the slide deck *Volatility estimation and Advanced Portfolio Optimization*:

$$\max_{\theta} \frac{1}{T} \sum_{t=0}^{T-1} u(\theta, r_{p,t+1}) = \frac{1}{T} \sum_{t=0}^{T-1} u\left(\sum_{i=1}^{N_t} \left(\bar{w}_{i,t} + \frac{1}{N_t}\theta'\hat{x}_{i,t}\right) r_{i,t+1}\right)$$

To reduce the deviation from the benchmark portfolio in an economically meaningful way, we can modify the optimization problem by adding a penalty term that incorporates the impact of turnover. The approach is to use the positive-definite matrix B for asset-specific transaction costs:

$$\nu_t(w_{t+1}, w_t, B) = (w_{t+1} - w_{t+})'B(w_{t+1} - w_{t+})$$

Where we have that w_{t+} are the weights prior to the adjustment at time $t + 1$, that is:

$$w_{t+}(w_t, r_{t+1}) = \frac{w_t \odot (1 + r_t)}{\iota'(w_t \odot (1 + r_t))}$$

We can now modify our optimization problem to include a penalty term on transaction costs by setting $\nu_{-1} = w_0'Bw_0$:

$$\max_{\theta} \frac{1}{T} \sum_{t=0}^{T-1} u(\theta, r_{p,t+1}) = \frac{1}{T} \sum_{t=0}^{T-1} u\left(\left[\sum_{i=1}^{N_t} \left(\bar{w}_{i,t} + \frac{1}{N_t}\theta'\hat{x}_{i,t}\right) r_{i,t+1}\right] - \nu_{-1}\right)$$

By incorporating a penalty on the impact of the turnover in the optimization function, we obtain a procedure that solves the problem in a economically meaningful.