Statistical Methodology in Alternative Risk Premia

Yang Ming

Supervisor:
Asst. Prof. Lim Chinghway

Contents Page

1. Aim of the FYP (9.5 minutes)
    I. Definitions
    II. Alternative Risk Premia
2. Literature Review (5.5 minutes)
    I. Current work
    II. Our contributions
3. Modern Portfolio Theory (8 minutes)
    I. Efficient Frontier
    II. Mean-Variance Optimization
        a. Limitations
        b. Solutions
4. Transaction Costs (9 minutes)
    I. Market Impact Model
    II. Optimized Market Impact Portfolio
    III. Optimized Mean-Variance Portfolio
5. Results from Optimized Strategies (9 minutes)
6. Conclusion and Further Work (3 minutes)

Link to slides: https://tinyurl.com/yangming-fyp

Definitions

Empirical Asset Pricing:

- Response variable is *excess return*
- *Return:* Percentage change of price over time
- *Excess return*: is the compensation given to the investor for the risk of holding the risky stock (equity)
- *Risk premium:* Long-only exposure on equities or bonds

$$R_{t,s} = \frac{P_{t+s}}{P_t} - 1$$

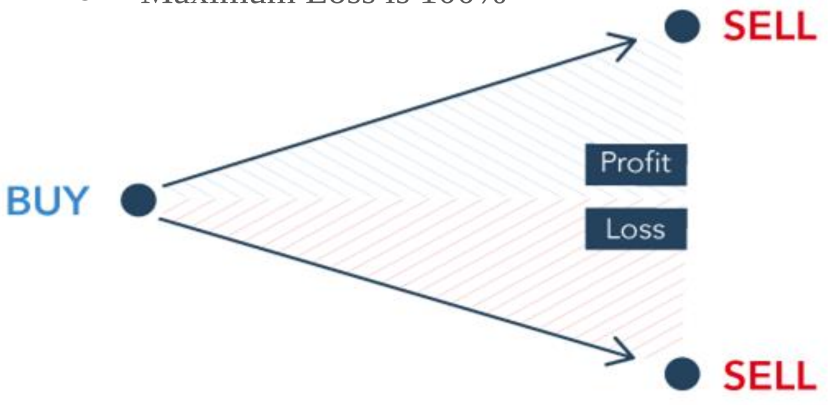
$R_{t,s}$:	Return from t over holding period $s \in \mathbb{R}$
P_{t+s}:	Price at time $t + s$
P_t:	Price at time t

$$y_{it} = R_{it} - R_f$$

y_{it}:	Excess Return of asset i
R_{it}:	Return of asset i
R_f:	Return of risk-free asset

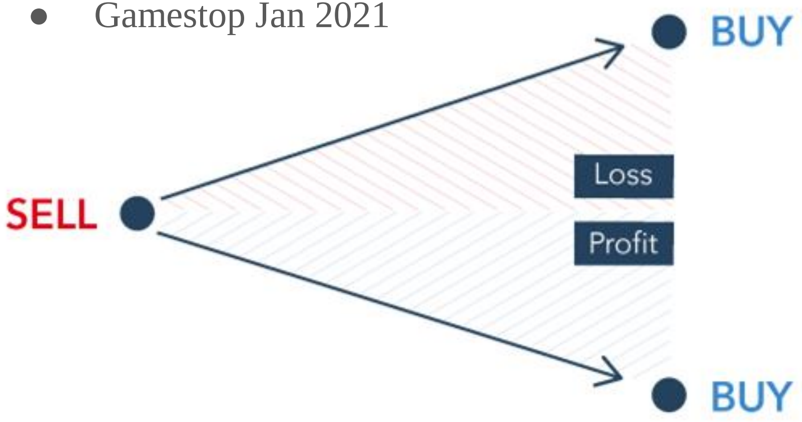
Long

- Buy-and-hold, bet on the stock to rise
- Stock A @ $1,000/share, drops to $0/share
- Maximum Loss is 100%



Short

- Opposite of Long, bet on the stock to fall
- Maximum Loss is infinite
- Gamestop Jan 2021



"*Risk premium:* Long-only exposure on equities or bonds"

Source: [IG Australia]

What is Factor Investing?

Factor investing:

- Decompose excess return into multiple factors
  - Macroeconomic factors
  - Fundamental factors
  - Statistical factors
- Cross-sectional regression to explain excess returns
- Long-only strategy
- **Idea**: Offset risk by exposing assets to other factors of return

Alternative Risk Premia (ARP):

- Extension of factor investing, Long/Short strategy
- All systematic risk factors that generates returns through other factors

What is Alternative Risk Premia?

ARP Strategies are Long/Short strategies. Five implementations of ARP strategies in equity markets:

1. Carry
    - Well known in currency markets, investors borrow currency with a low interest rate to buy a currency with a high interest rate, with the difference in rates the investor returns
    - Koijen et al (2013) abstracts the principles behind this to identify equities
    - Costs: Risk-free rate (to finance its purchase) Benefits: Trailing dividend yield (to predict future dividend yield)
    - Go long on stocks with net benefit (positive carry), short on stocks with net cost (negative carry)
2. Size
    - Assets with smaller market cap generate higher returns
    - Gain additional exposure through the illiquidity of the asset: (Smaller size, lower trading volume, higher bid-ask spread, slippage, etc.)
3. Momentum
    - Jegadeesh and Titman (1993) Rank stocks based on returns over the past J months
    - Go long on past winners (top deciles), short on past lowers (bottom deciles)
4. Reversal
    - Construct autocorrelation functions to establish if a mean-reverting pattern exists, buy the stock, because on average a negative return is followed by a positive return
    - Otherwise known as Mean-Reverting Strategies, capitalize based on investor’s short-term overreaction to information and fads
5. Value
    - Value strategies tilt asset allocation towards under-valuated stocks over over-valuated ones, in expectation greater rates of return for the former as the market acknowledges its value.

How can Machine Learning augment ARP strategies:

1. Better predictive power against existing factor models for better stock selections
2. Model non-linear relationships between factors and complex interactions between returns and factors
3. With strong predictive power, Long/Short enables you to profit from both upside and downside price movement of the stock.

The Investor’s Problem

Portfolio: A mixture of finance investment held by investor

Chris has 100% of his portfolio invested in stocks
Chris has 40% of his portfolio invested in stocks, 40% in bonds, 20% in real estate

The Investor’s Problem: How to best distribute one’s money to a portfolio to maximize/minimize your investment goal

Portfolio A: 5% in Tesla, 10% in Microsoft, …
Portfolio B: 10% in Southwest Airlines, 10% in Boeing, …

Portfolio Optimization: Selecting the combination of asset allocation such that the portfolio maximizes/minimizes the objective function out of the set of all possible portfolios

Portfolio Optimization as a Statistics Problem

- Portfolio optimization as a Quadratic Program (QP) optimization
  - Maximize/minimize quadratic objective function
  - Leaves out realities such as Transaction Costs from the optimization

- Portfolio optimization with Quadratic Transaction Costs as a Quadratically Constrained Quadratic Program (QCQP) optimization
  - Maximize/minimize Quadratic objective function
  - Quadratic constraints
    - Transaction Costs (broker commissions, bid-ask spread, market impact)

The aim of FYP

What is the practical implication of the problem?

- Transaction costs erodes profitability
- Portfolio optimization with quadratic transaction costs assesses portfolio performances realistically

The aim of FYP

Potential of Machine Learning in Stock Markets

We assess whether using forecasts from machine learning models, the subsequent portfolio constructed from these predictions provide a reasonable level of performance, after taking into **account transaction costs and other realistic constraints**.

Current and novel works in empirical asset pricing

1. Cross-sectional approach
    a. Capital Asset Pricing Model (CAPM) - Sharpe (1963)
    b. Fama-French Three-Factor Model - Fama & French (1993)
    c. Carhart Four-Factor Model - Carhart (1997)
    d. Machine Learning Models - Gu et al (2018)
        i. Linear Regression, ElasticNet (L1 + L2 Regularization), Principal Components Regression, Partial Least Squares Regression, Random Forest, Gradient-boosted Regression Trees, Fully-connected feed-forward Neural Networks

Current and novel works in empirical asset pricing

2. Time-series approach
    a. Linear Models - Welch & Goyal (2008)
    b. Deep Learning Models - Nakagawa et al (2019)
        i. Recurrent Neural Networks (RNNs), Long-Short Term Memory Networks (LSTMs), Gated Recurrent Units (GRUs)
    c. Achal (2020)
        i. Gu et al (2018) and Nakagawa et al (2019) models

Modelling

Investable Universe:

- 622 publicly traded US companies
- Market capitalization of companies:
  - Small, medium, large cap stocks
  - Worth of a company determined by stock market (Outstanding Shares * Current Market Value)
  - Each have taken turns leading the market as they react differently to economic developments

2-step modelling:

1. ARP factor returns with RNNs

$$\hat{f}_{it} = RNN(X_{it})$$

2. Excess returns with Linear model

$$\hat{y}_{it} = \hat{f}_{1t}(X_{1t}) + \hat{f}_{2t}(X_{2t}) + \ldots + \hat{f}_{nt}(X_{nt})$$

Results: 12.2% out-of-sample R2 (excess returns vs predicted excess returns) from 2017-2019 Q3

Experimental Setup

US Equities:

- Stock-level indicators, daily frequency from 2005 - 2019-Q3
- 240 predictors, quarterly and/or yearly frequency:
    - Macroeconomic, market, industry-related indicators
- 60+ million rows of data:
    - 15,000+ publicly listed US companies

Azure Virtual Machine:

- General Purpose Compute (D2-Series – D64s v4)
- 64 vCPU (Xeon Platinum 8272CL, 2.5Ghz Base, 3.4Ghz Turbo)
- 256GB RAM, 512GB SSD

Solvers:

- **MOSEK**, GUROBI, XPRESS, SCS, CVXOPT

Problem Statement

Portfolios analyzed in the literature are not designed to optimize or pay attention to transactions costs in any way

Only a handful of Machine Learning literature paid attention to the effect of transaction costs on portfolio performance

- Borghi & Giuliano (2020), Ma, Han & Wang (2021)

Portfolio optimization with quadratic constraints to minimize transaction costs can be computationally intractable

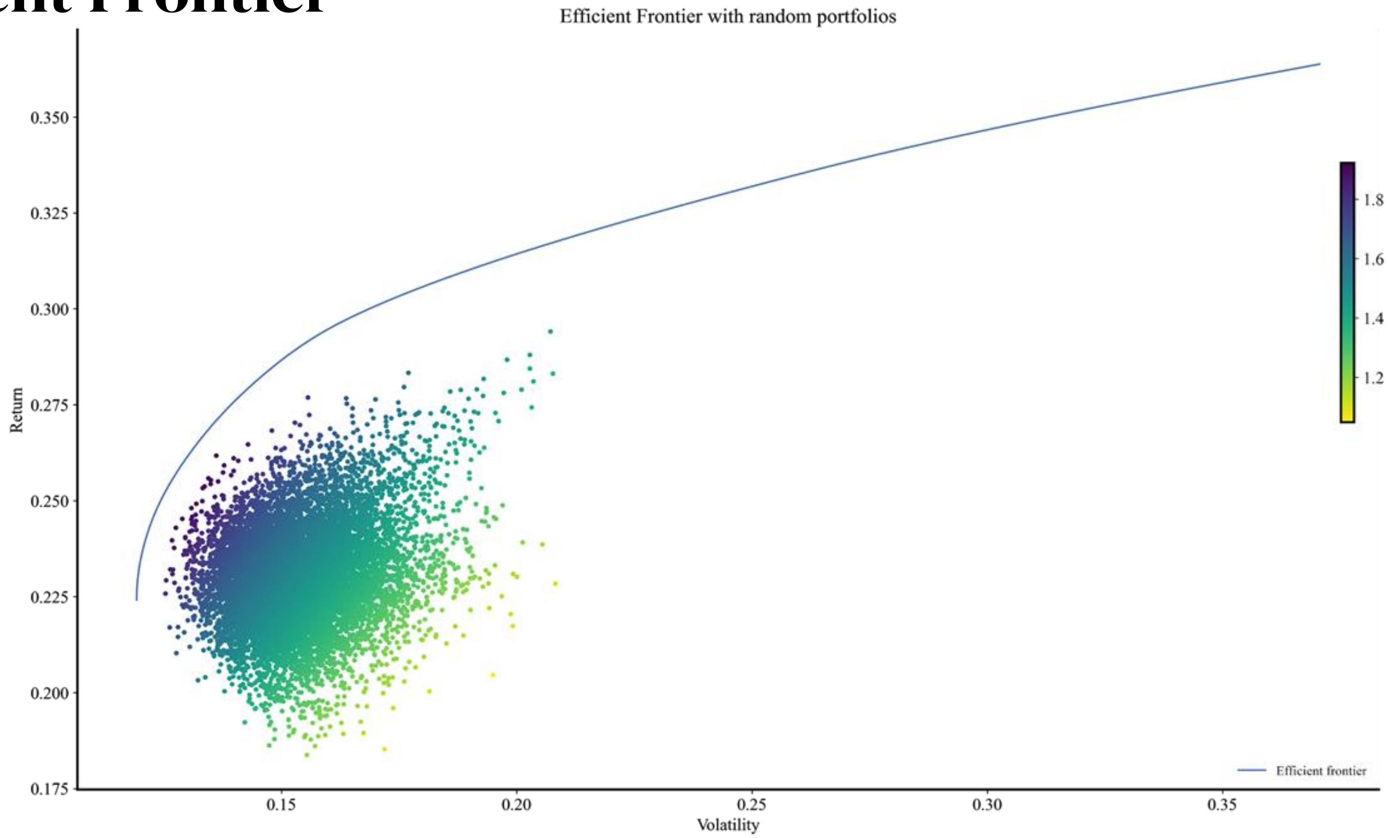
Contributions

1. Transaction cost modelling to estimate the market impact of any arbitrary trade of an institutional investor before it is executed

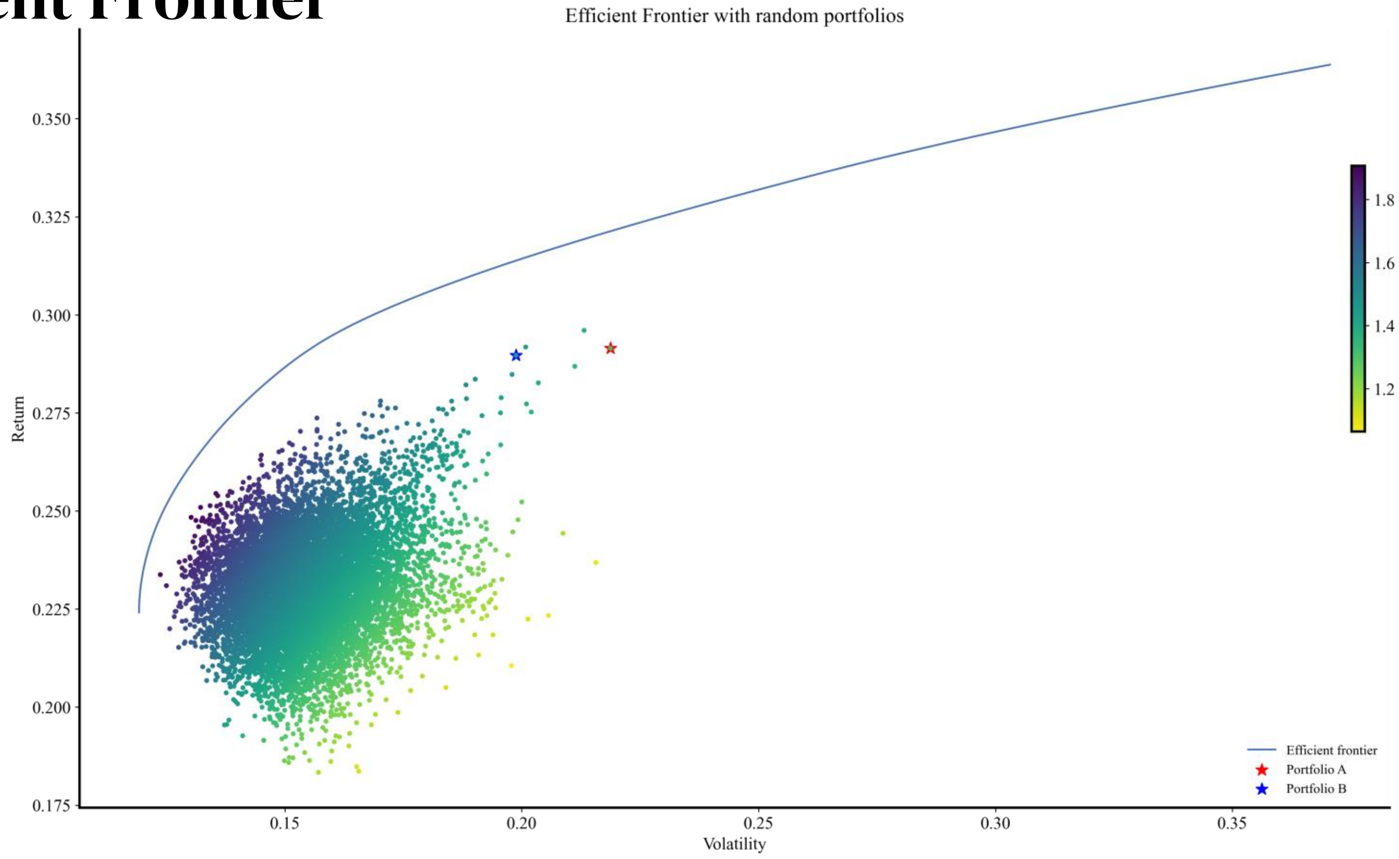
2. Novel way of portfolio optimization with quadratic transaction costs to produce computationally tractable asset allocations

3. Backtest optimized strategies that minimizes transaction costs, and assess performance of machine learning models under realistic conditions

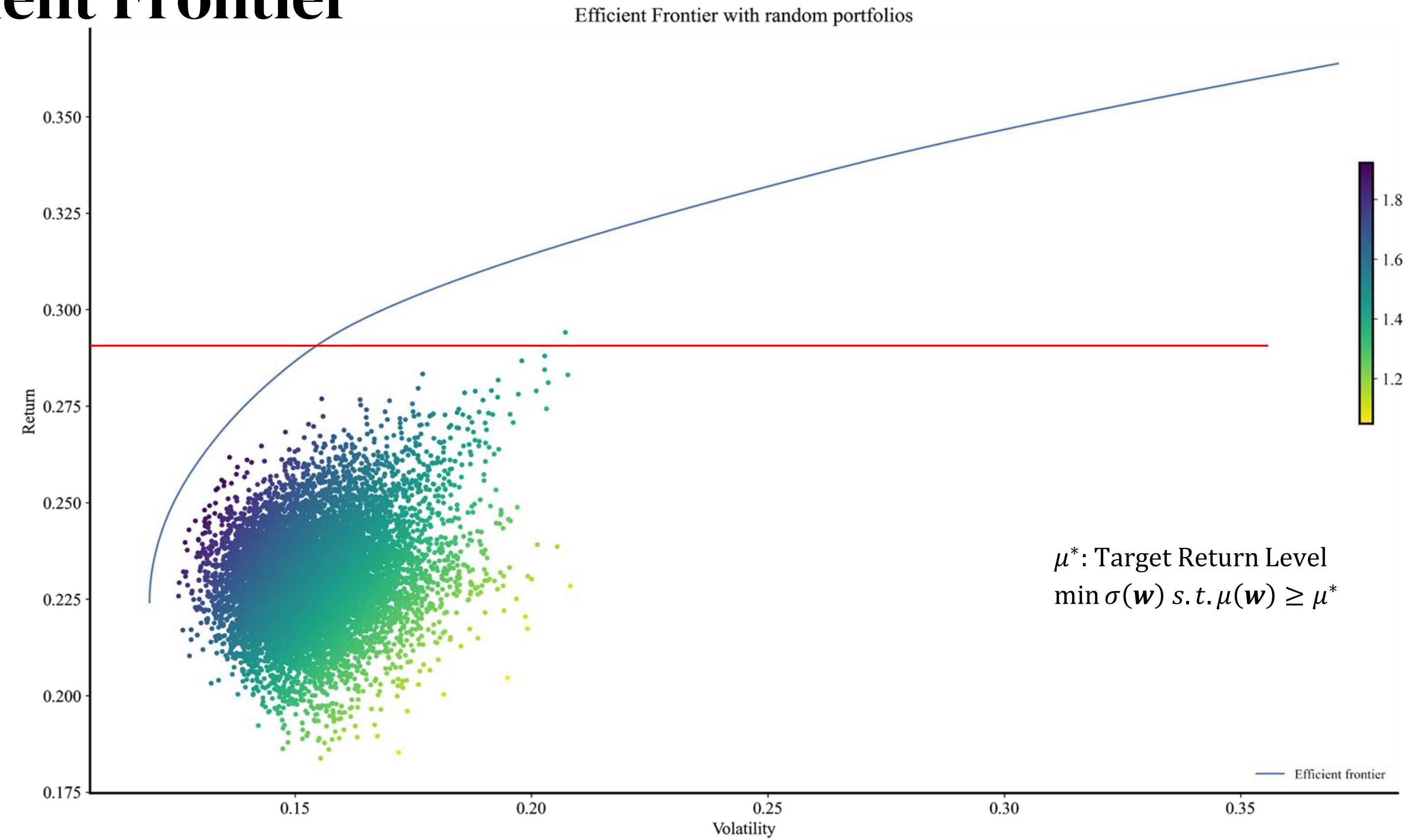
Efficient Frontier



Efficient Frontier



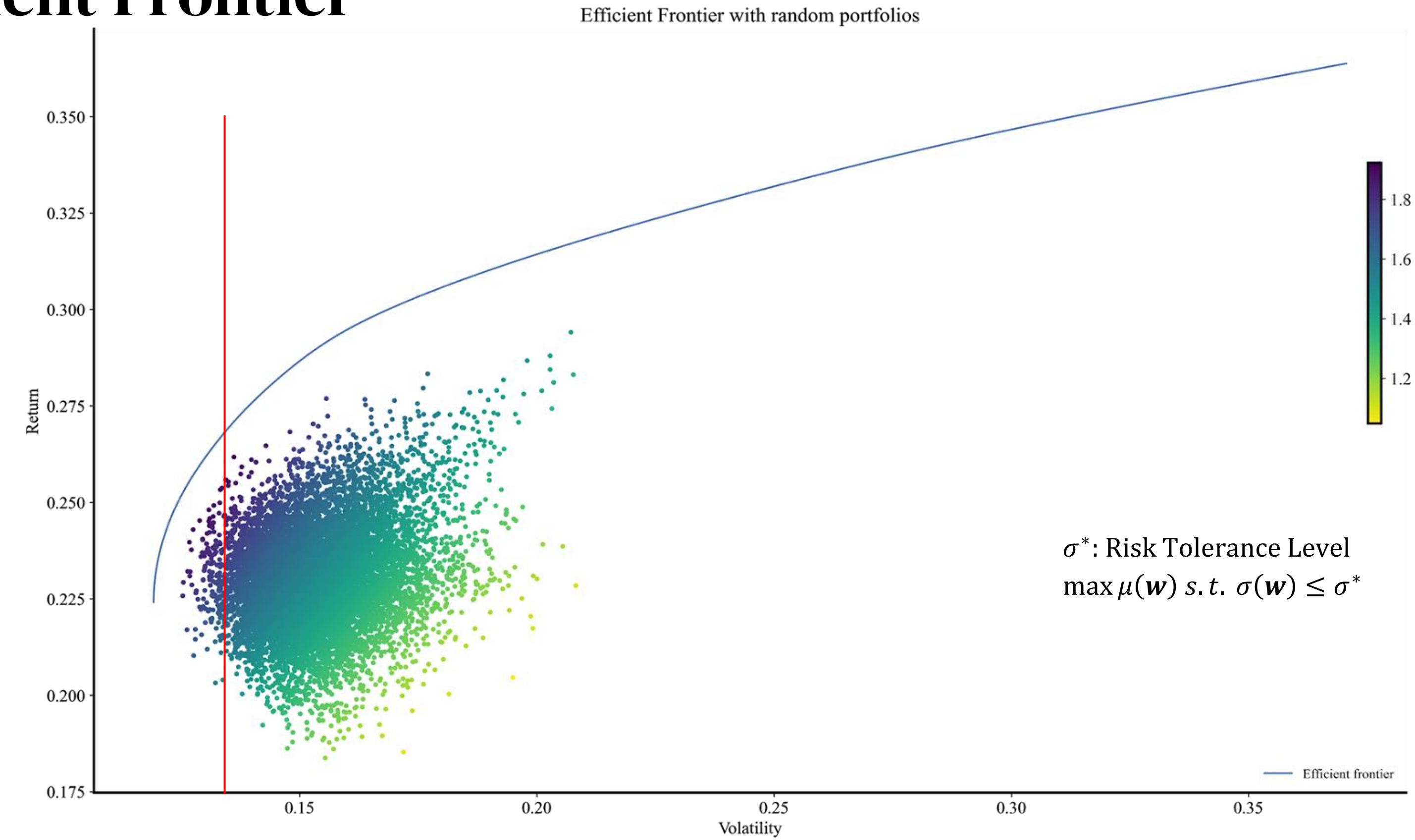
Efficient Frontier



μ^*: Target Return Level

$\min \sigma(\boldsymbol{w})\ s.t.\ \mu(\boldsymbol{w}) \geq \mu^*$

Efficient Frontier



σ^*: Risk Tolerance Level

$\max \mu(\boldsymbol{w}) \ s.t. \ \sigma(\boldsymbol{w}) \leq \sigma^*$

Mean-Variance Optimization

The most popular way for portfolio optimization is the Mean-Variance Optimization by Markowitz (1952)

Using a Quadratic utility Function:

$$U(w) = w^\top \mu - \frac{\phi}{2} w^\top \Sigma\, w$$

turns the Investor's Problem into a QP:

$$\min \mathrm{w}^\top \mu - \frac{\phi}{2} w^\top \Sigma w$$

Subjected to:

$$\mathbf{1} w = 1$$
$$-1 \leq w \leq 1$$

Where:

$\boldsymbol{w}$: $(w_1, \dots, w_n)$ is the vector of weights in a portfolio of assets
μ: vector of expected returns of assets in portfolio
Σ: covariance matrix of returns on asset in portfolio
ϕ: risk aversion, $\gamma = \phi^{-1}$ is the given risk level

Explanation:

$\mathbf{1}w = 1$**:** Budget constrained to the size of the original investment
$-1 \leq w \leq 1$: Asset allocation

Mean-Variance Optimization

- **Mean-**Variance Optimization
  - Expected Returns is intuitively your model's forecast of the future stocks

$$\min \mathrm{w}^{\top} \underbrace{\mu}_{\text{Expected returns}} - \frac{\phi}{2} w^{\top} \Sigma w$$

Mean-Variance Optimization

- Mean-**Variance** Optimization
  - Covariance Matrix is a Risk Model to quantify asset risk and co-dependencies
  - Risk-reward tradeoff

$$\min \mathrm{w}^\top \mu - \frac{\phi}{2} w^\top \underbrace{\Sigma}_{\text{Covariance Matrix}} w$$

Mean-Variance Optimization

- Diversification
  - Risk can be reduced by taking many uncorrelated bets

$$\min \mathrm{w}^\top \underbrace{\mu}_{\text{Expected returns}} - \frac{\phi}{2} w^\top \underbrace{\Sigma w}_{\text{Covariance Matrix}}$$

Limtations

- Sample Covariance Matrix
  - Assumes multivariate normal distributed returns

In-sample covariance matrix is calculated by:

$$S = \frac{1}{n-1} \sum \left(x_{ij} - \bar{x}\right)\left(x_{ij} - \bar{x}\right)^{\top}$$

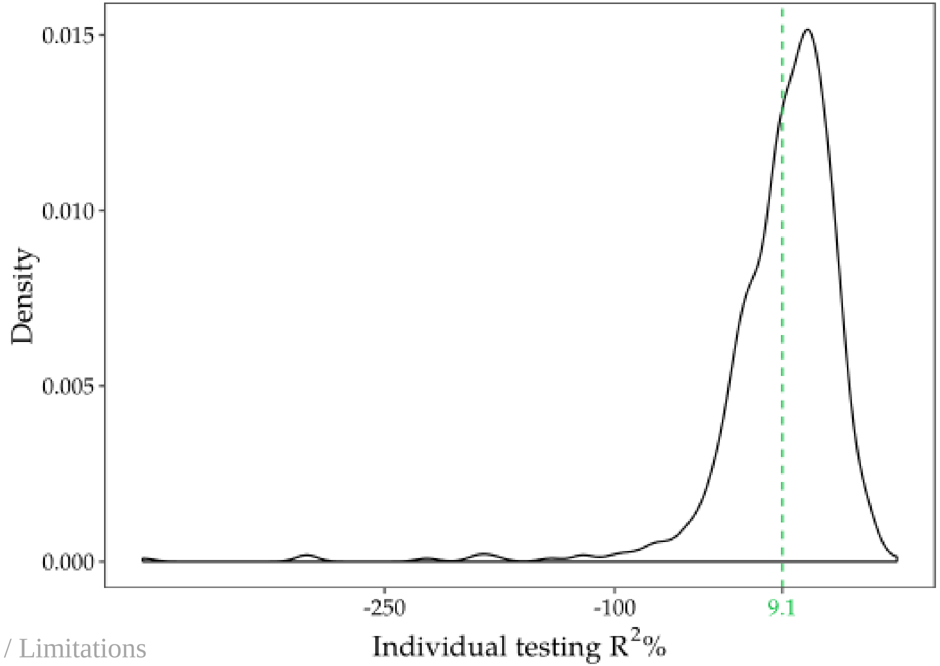
Is an unbiased estimator of the Covariance Matrix:

$$\sum = \mathbb{E}[(X - \mu_x)(X - \mu_x)^{\top}]$$

where
x_{ij} is the return of asset i with respect to asset j
$\bar{x}$ is the sample historical returns of asset i
μ_x is the mean of historical returns of asset i

Limitations

- Empirical returns are not normally distributed
  - Sample skewness and sample kurtosis exhibit fat tails
  - Non-symmetric distribution
  - Multivariate tail dependence



Limitations

- Mean-Variance Optimization
  - Risk Model: In-sample covariance matrix (S)
    - Low bias, large estimation errors
    - Mathematically optimized, but financially intractable positions

	Sample Covariance
A	-0.999999
AAN	-0.999999
	…
ZBH	-0.999999
ZBRA	0.999997
ZEUS	-0.999999

Table 4 – Weight allocations from mean-variance optimization

Solutions

- Companies (APT, BARRA) have researched into generating covariance matrices claimed to be better suited for Mean-Variance optimization.
  - The companies do not participate in the risk itself
  - Proprietary: Unable to independently inspect such risk models
- Sharpe (1963) – One Factor Model
  - Highly structured estimator (F)
  - Small estimation errors, but high bias
- Ledoit & Wolf (2013)
  - Developed specifically for financial portfolio allocations
  - Combine both a highly structured estimator F as well as an unstructured estimator S using a shrinkage coefficient

Solutions

Ledoit-Wolf Shrinkage

Ledoit-Wolf Shrinkage

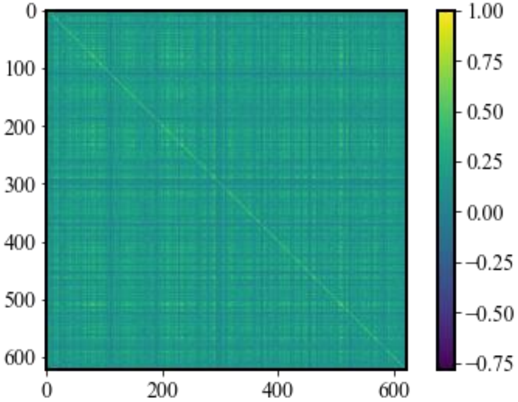
$$\delta F + (1-\delta)S$$

Where F is Sharpe's One-Factor Model

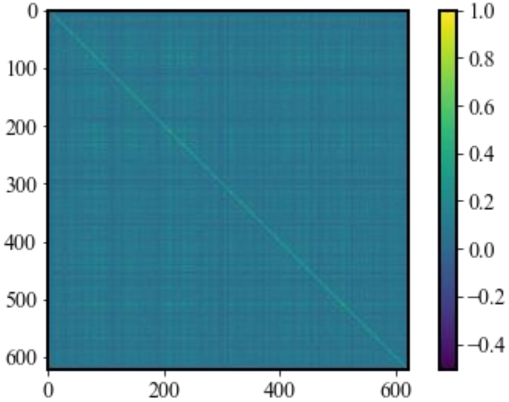
$$F_{ij} = \beta_i * Var(R_j)$$

Where S is Sample Covariance Matrix

$$S = \frac{1}{n-1}\sum(x_{ij}-\bar{x})(x_{ij}-\bar{x})^{\top}$$

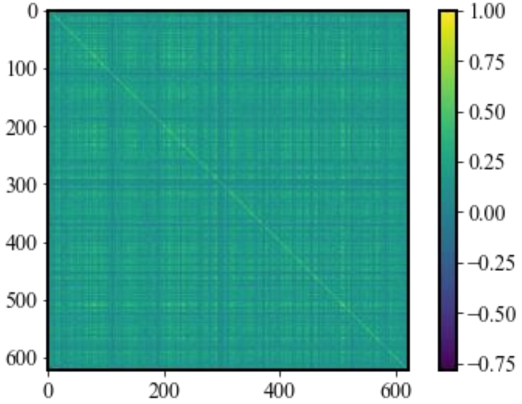


Sample Covariance Matrix

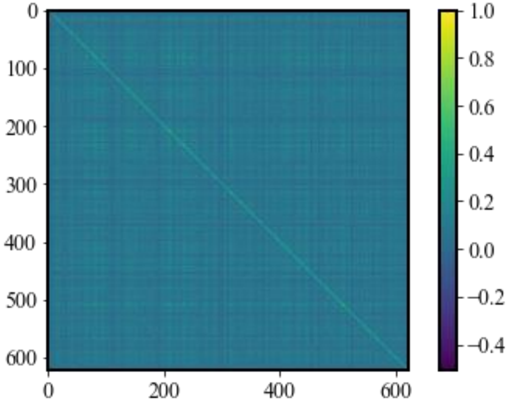


Ledoit-Wolf Shrinkage

Solutions



Sample Covariance Matrix



Ledoit-Wolf Shrinkage

	Sample Covariance	Ledoit-Wolf Shrinkage
A	-0.999999	-0.00363
AAN	-0.999999	0.00927
	…	…
ZBH	-0.999999	-0.00523
ZBRA	0.999997	0.00796
ZEUS	-0.999999	-0.01642

Table 4 – Weight allocations from mean-variance optimization

Transaction Cost

1. Bid-ask spread
   a. When you want to buy/sell an asset, the order to do that doesn’t execute at the price you see, because someone has to fulfill the other side of the trade
   b. The difference between the highest price that a buyer is willing to pay for an asset and the lowest price that a seller is willing to accept.
2. Commission
   a. Service charge assessed by a broker forhandling purchases and sales of securities
3. Market Impact
   a. Institutional investors with large trade sizes, fund's sizable purchase/selling of stock artificially drives the price higher/lower as the market responds to the order
   b. If an institutional investor wishes to purchase 1 million shares of a stock, there will not enough shares in the market available at the current price (liquidity), and the act of buying or selling moves the price against the institutional investor.

(Frazzini et al – 2018) For an institutional investor, the transaction costs of its trade is dominated by its market impact, since bid-ask spread, commission and other types of transaction costs do not effectively scale

Transaction Cost

Transaction costs model in literature:

1. Frazzini, A., Israel, R., & Moskowitz, T. J. (2018) - Trading Costs

2. Scherer (2007) - Portfolio Construction and Risk Budgeting

3. Boyd S. et al (2007) - Multi period trading via convex optimization

Frazzini et al (2018) Market Impact Model:

$$MI = a + b * c + c * sign(x)\sqrt{|X|}$$

Where:

$a = \theta_2$time-trend $+ \theta_3 log$ (1+ market cap) $+ \theta_6 \sigma^{IV} + \theta_7$VIX,
a controls the general level of the trading and the market characteristics
$b = \theta_4$ (fraction of daily average volume)
$c = \theta_5$ (sqrt of fraction of daily average volume)
a and b controls the characteristics of the trade with respect to the stock
x: 100*m/dtv signed dollar volume of the trade (m) as a fraction of stock's average one-year dollar volume

Scherer (2007)'s Transaction Cost Model:

$$tc = Commision + \frac{Bid}{Ask} - Spread + \theta\sqrt{\frac{TradeVolume}{Dailyvolume}}$$

Boyd (2007)'s Transaction Cost Model:

$$tc = a|x| + b\sigma\frac{|x|^{3/2}}{V^{1/2}} + cx$$

a: one-half bid-ask spread ex-ante
V: daily average volume traded
σ: idiosyncratic volatility
x: signed dollar of the trade

Market Impact Model

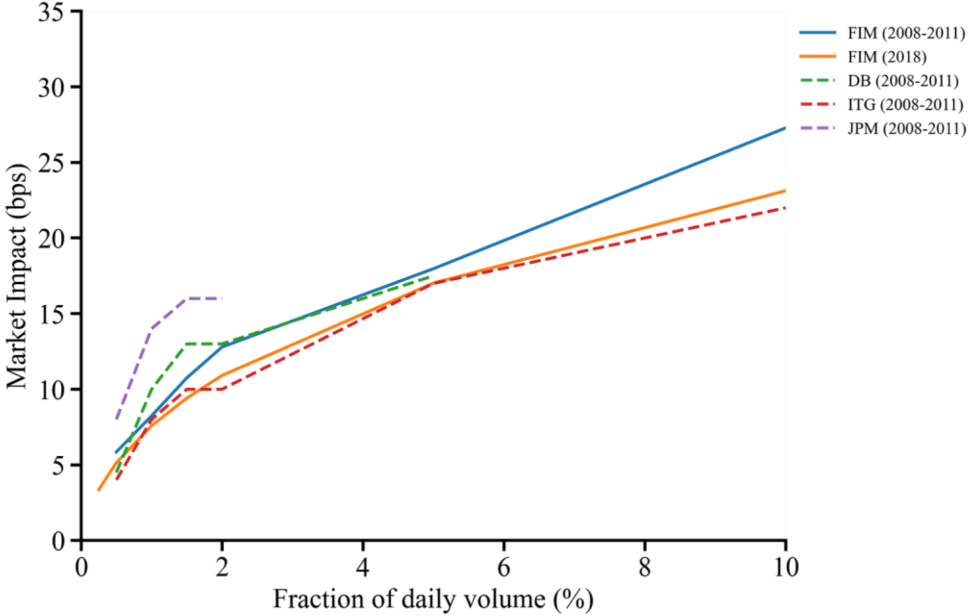
Trained on \$1.7 trillion of a large institutional investor's (AQR) live executed trades

19-year period from August 1998 to June 2016 across 21 equity markets

The model is based on market, trade and stock characteristics

- Matches economic theory
- Matches statistics measures of market impact such as Kyle's Lambda (Kyle, 1985) and other empirical measures (Hasbrouck, 2009)
- Matches other market impact data from other brokerage houses

Results from Market Impact Model



	Market Impact (bps) at 2.4% DTV
ANcerno	24.5
AQR	18.2
Model Estimates	23.1

Table 3 – Reported realized market impact

Brokerage Houses:

1. Investment Technology Group (ITG)
2. Deutsche Bank (DB)
3. JP Morgan (JPM)

Consulting Firm:

1. ANcerno

Economic Theory (Hashbrouck - 1991), Empirical evidence from institutional transaction costs (Almgren et al - 2005) both suggests concavity of market impact costs with respect to trade size

Concavity of market impact size suggests that transaction costs are quadratic, rather than linear in nature.

MVO with Quadratic Constraints

Scherer (2007): Fitting a quadratic transaction cost

Considering a Quadratic Function to model transaction cost:

$\boxed{C(w|\vec{w})\text{: Cost to rebalance current portfolio } \vec{w} \text{ to portfolio } w}$

$$\boxed{\min \mathrm{w}^\top \mu - \frac{\phi}{2} w^\top \Sigma w \; - C(w|\vec{w})}$$

Subjected to:

$$\underbrace{\boxed{\mathbf{1}w + C(w|\vec{w}) = 1}}_{\text{Quadratic Term}}$$

Where:

$\boldsymbol{w}$: $(w_1, \ldots, w_n)$ is the vector of weights in a portfolio of assets

$\mathbf{1}w + C(w|\vec{w}) = 1$: Budget Constraint (The cost of rebalancing + reallocating weights must not exceed the budget)

Mean-Variance Optimization

Using a Quadratic utility Function:

$$U(w) = w^\top \mu - \frac{\phi}{2} w^\top \Sigma\, w$$

turns the Investor's Problem into a QP:

$$\boxed{\min \mathrm{w}^\top \mu - \frac{\phi}{2} w^\top \Sigma w}$$

Subjected to:

$$\boxed{\mathbf{1}w = 1}$$

Where:

$\boldsymbol{w}$: $(w_1, \ldots, w_n)$ is the vector of weights in a portfolio of assets

μ: vector of expected returns of assets in portfolio

Σ: covariance matrix of returns on asset in portfolio

ϕ: risk aversion, $\gamma = \phi^{-1}$ is the given risk level

$\mathbf{1}w = 1$: Budget constrained to the size of the original investment

Optimized Market Impact Portfolio

Our contribution:

- Solve Quadratic Transaction Costs with an Augmented QP instead of QCQP
- Idea inspired by usage of ex-ante tracking error constraint in index-replicating strategies

1. Calculate the weights w_b of the unconstrainted portfolio:

$$\mathrm{w_b}:\ \min \mathrm{w}_b^\top \mu - \frac{\phi}{2} w_b^\top \Sigma w_b$$

2. Find w such that it minimizes Transaction Cost:

$$w:\ \min \text{Transaction Cost } (w)$$

Subjected to:

1. Tracking Error Constraint: $\sqrt{(w - w_b)^\top \Sigma (w - w_b)} \leq 1\%$

2. Fraction of Daily Average Volume $\leq 10\%$

Optimized Market Impact Portfolio

1. Unconstrainted Mean-Variance Portfolio

Ticker	Price	Previous Weight	Previous Share	New Weights	New Shares	Shares Difference	1-year DAV
CAMP	16.6	-0.00839	-57863	0.0995	718807	776670	439010

Table 5 – Allocation of CAMP using unconstrainted mean-variance optimization

- Estimated Market Impact: 44bps (1 basis point is hundredth of a percent 0.01%)
- Estimated Execution Price: $P^{ex} = \left(1 + \frac{MI}{100}\right) * P^{start} = \left(1 + \frac{44}{100}\right) * 16.6 = 23.9$
- Profitability eroded: $(23.9 - 16.6) * 776{,}670 = \$5{,}708{,}525$

2. Market Impact Portfolio

Ticker	Price	Previous Weight	Previous Share	New Weights	New Shares	Shares Difference	1-year DAV
CAMP	16.6	-0.015	-15329	0.015	16760	32089	439010

Table 6 – Allocation of CAMP using market impact optimized strategy

Optimized Mean-Variance Portfolio

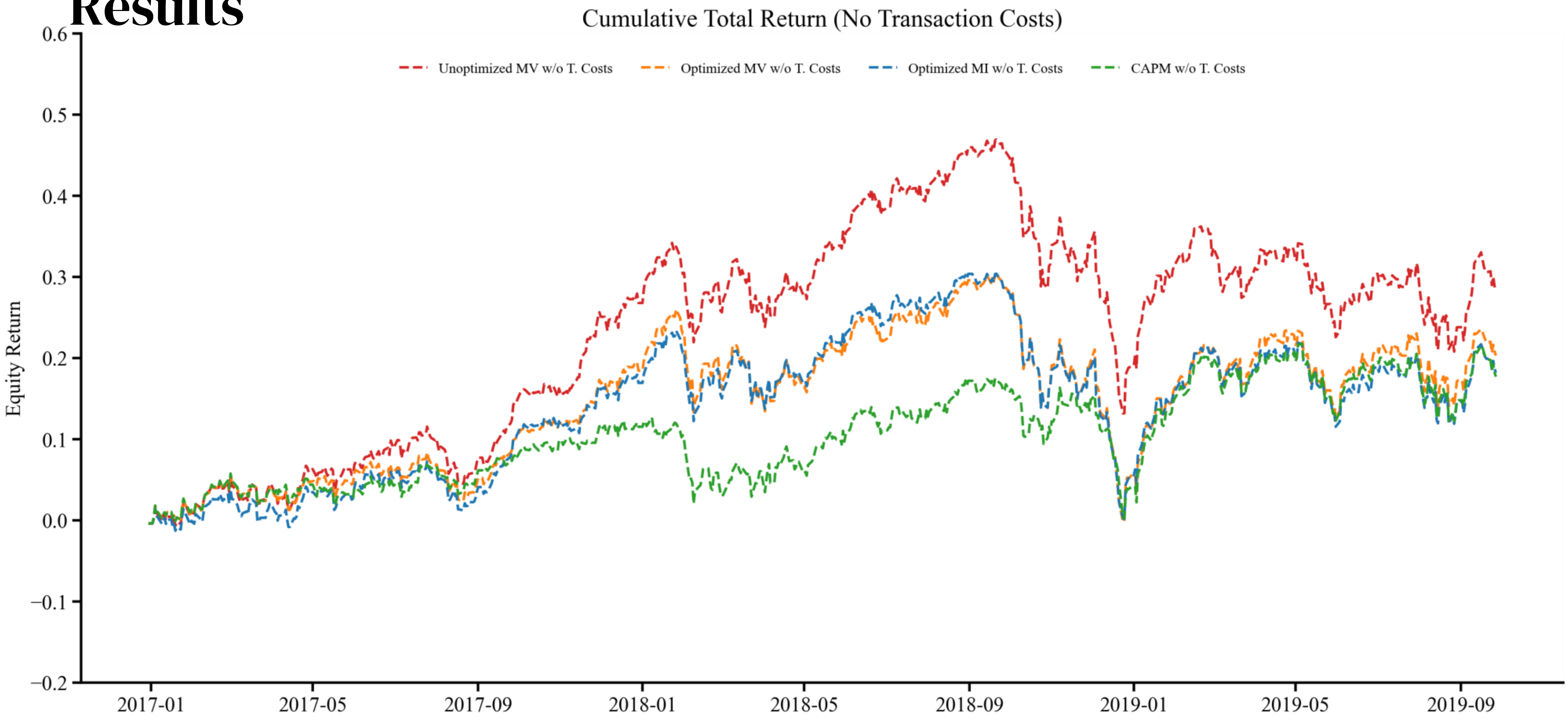
Commonly used constraints in quant strategies in practice (Borghi & Giuliano 2020):

1. Net sector exposure is limited within -10% and 10% of net portfolio value using Level 1 GICS sectors
2. Individual stock weights are capped at 1.5%, with a threshold of 0.5%
3. Limit amount that can traded at rebalancing to 10% of its 1-year Daily Average Volume

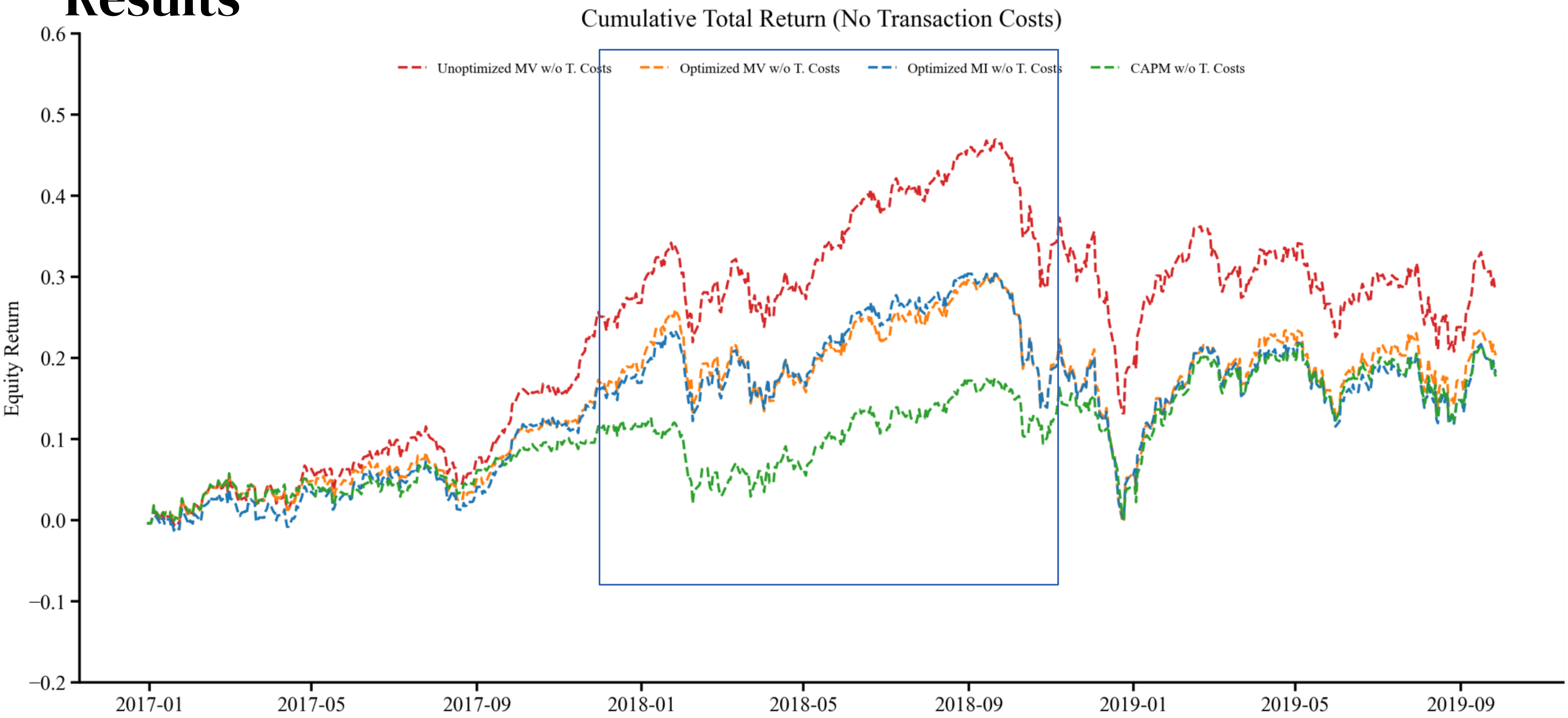
Simulation

Backtesting Setup	
Backtest period	Q4-2016 to Q3-2019
Rebalance Frequency	Quarterly
Observations	11
Strategy	Long/Short
Maximum Risk Level	5%
Starting Capital	100 Million USD
Trading Cost	One-way 10bps
Fully collateralized portfolio	

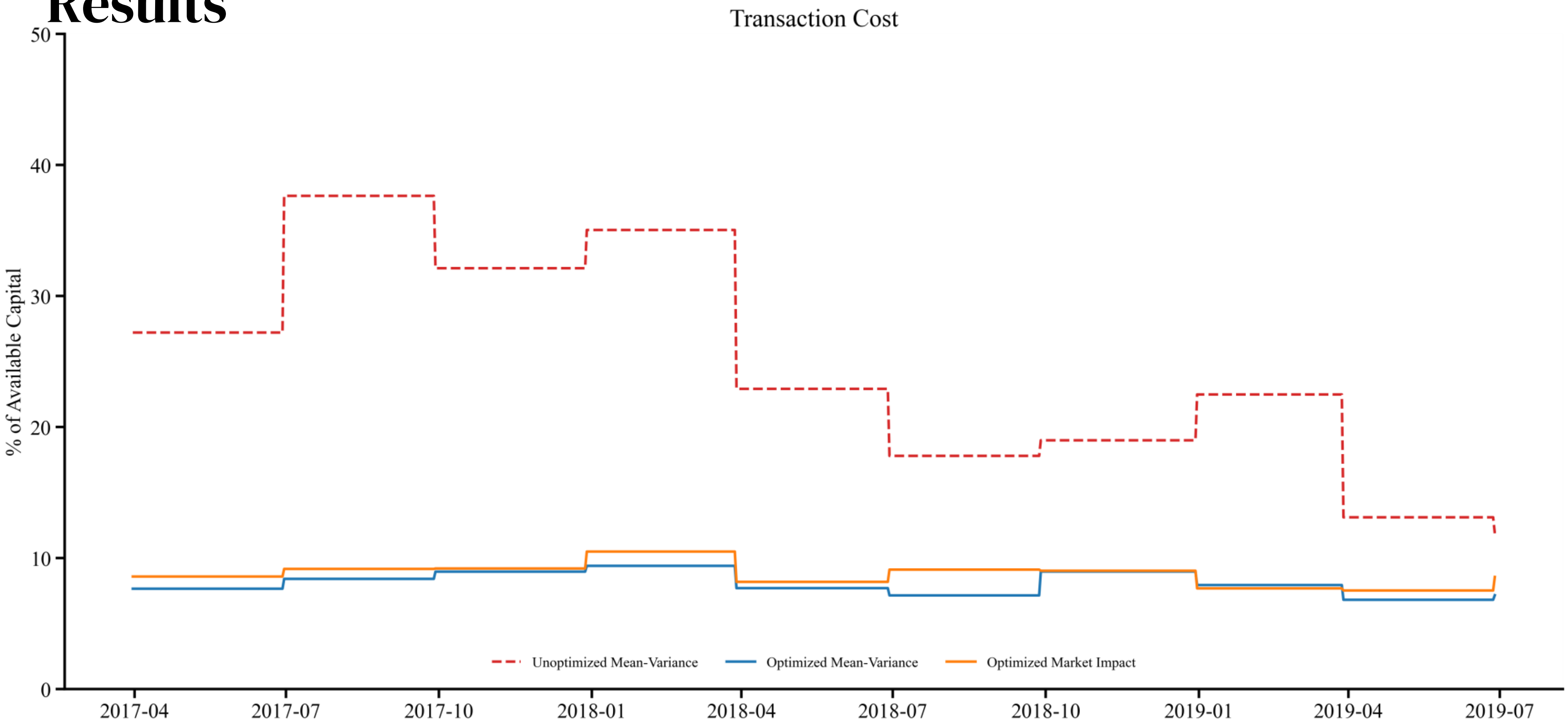
Results


Cumulative Total Return (No Transaction Costs)
Unoptimized MV w/o T. Costs
Optimized MV w/o T. Costs
Optimized MI w/o T. Costs
CAPM w/o T. Costs
Equity Return
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
2017-01
2017-05
2017-09
2018-01
2018-05
2018-09
2019-01
2019-05
2019-09

Results


Cumulative Total Return (No Transaction Costs)
Unoptimized MV w/o T. Costs
Optimized MV w/o T. Costs
Optimized MI w/o T. Costs
CAPM w/o T. Costs
Equity Return
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.2
2017-01
2017-05
2017-09
2018-01
2018-05
2018-09
2019-01
2019-05
2019-09

Results


Transaction Cost
% of Available Capital
50
40
30
20
10
0
2017-04
2017-07
2017-10
2018-01
2018-04
2018-07
2018-10
2019-01
2019-04
2019-07
Unoptimized Mean-Variance
Optimized Mean-Variance
Optimized Market Impact

Evaluation

Alpha Model: Macquarie Quant Alpha Model (Borghi & Giuliano - 2020)

- Backtesting period from January 2011 to December 2018 (95 months)
- Portfolios rebalanced monthly, rather than quarterly
- Other aspects of methodologies identical to our backtest setup

	CAPM	Mean-Variance	Market Impact	Alpha
Net Return (Geom.)	11.7%	23.3%	23.4%	7.9%
Net Return (Geom.) - Optimized	9.59%	14.3%	13.9%	6.2%
Net Information Ratio	0.789	0.962	0.962	1.11
Net Information Ratio - Optimized	0.801	0.797	0.802	0.870
Avg. Turnover	9.51	2.47	2.47	
Avg. Turnover - Optimized	9.01	0.725	1.63	1.39
Max Drawdown	14.4%	23.5%	23.2%	11.9%
No. of Observations	11	11	$\boxed{11}$	$\boxed{95}$

Table 7 – Backtest results

Evaluation

Turnover:

- Calculated as a total of long and short sides and expressed as a percentage of the collateral value
- Transaction cost impacts a portfolio with higher turnover more
- Higher turnover means greater expenses

	CAPM	Mean-Variance	Market Impact	Alpha
Net Return (Geom.)	11.7%	23.3%	23.4%	7.9%
Net Return (Geom.) - Optimized	9.59%	14.3%	13.9%	6.2%
Net Information Ratio	0.789	0.962	0.962	1.11
Net Information Ratio - Optimized	0.801	0.797	0.802	0.870
Avg. Turnover	9.51	2.47	2.47	
Avg. Turnover - Optimized	9.01	0.725	1.63	1.39
Max Drawdown	14.4%	23.5%	23.2%	11.9%
No. of Observations	11	11	11	95

Table 7 – Backtest results

Evaluation

Turnover:

- Calculated as a total of long and short sides and expressed as a percentage of the collateral value
- Transaction cost impacts a portfolio with higher turnover more
- Net Returns 9% lower vs Net Returns (Alpha) 1.7% lower

	CAPM	Mean-Variance	Market Impact	Alpha
Net Return (Geom.)	11.7%	23.3%	23.4%	7.9%
Net Return (Geom.) - Optimized	9.59%	14.3%	13.9%	6.2%
Net Information Ratio	0.789	0.962	0.962	1.11
Net Information Ratio - Optimized	0.801	0.797	0.802	0.870
Avg. Turnover	9.51	2.47	2.47	
Avg. Turnover - Optimized	9.01	0.725	1.63	1.39
Max Drawdown	14.4%	23.5%	23.2%	11.9%
No. of Observations	11	11	11	95

Table 7 – Backtest results

Evaluation

Max Drawdown

- Calculates maximum observed loss from peak to trough of a portfolio
- Measures the downside risk over a specific time period
- Key concern for investors to measure the relatively riskiness of one strategy over another

	CAPM	Mean-Variance	Market Impact	Alpha
Net Return (Geom.)	11.7%	23.3%	23.4%	7.9%
Net Return (Geom.) - Optimized	9.59%	14.3%	13.9%	6.2%
Net Information Ratio	0.789	0.962	0.962	1.11
Net Information Ratio - Optimized	0.801	0.797	0.802	0.870
Avg. Turnover	9.51	2.47	2.47	
Avg. Turnover - Optimized	9.01	0.725	1.63	1.39
Max Drawdown	14.4%	$\boxed{23.5\%}$	$\boxed{23.2\%}$	$\boxed{11.9\%}$
No. of Observations	11	11	11	95

Table 7 – Backtest results

Evaluation

Information Ratio (IR)

- Daily outperformance of fund vs index / Annualized std deviation of daily outperformance
- Metric of a portfolio's ability to generate excess returns against a benchmark index
- (Informa) IR 0.4 – 0.6 considered already good, IR > 1.0 is considered rare

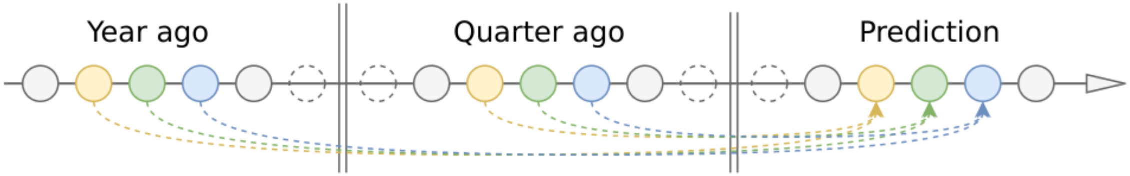
	CAPM	Mean-Variance	Market Impact	Alpha
Net Return (Geom.)	11.7%	23.3%	23.4%	7.9%
Net Return (Geom.) - Optimized	9.59%	14.3%	13.9%	6.2%
Net Information Ratio	0.789	0.962	0.962	1.11
Net Information Ratio - Optimized	0.801	$\boxed{0.797}$	$\boxed{0.802}$	$\boxed{0.870}$
Avg. Turnover	9.51	2.47	2.47	
Avg. Turnover - Optimized	9.01	0.725	1.63	1.39
Max Drawdown	14.4%	23.5%	23.2%	11.9%
No. of Observations	11	11	11	95

Table 7 – Backtest results

Further work

1. Using Attention mechanisms to augment long time-series

- Even LSTM/GRU gradually forget information from oldest states for long time-series
- Classical Bahdanau (Bahdanau, Cho & Bengio - 2014) or Luong (Luong, Pham & Manning - 2015) Attention methods would be too computationally intensive
- Fixed-weight sliding-window Attention (Sullin - 2018)



Further work

2. Different Risk Models & Portfolio Construction Methodologies

- (Rockafellar & Stanislav, 2000) Conditional Value at Risk (CVaR) – for more effective downside risk management
- Additional constraints: portfolio turnover constraint, etc.
- Other portfolio construction methodologies

Conclusion

Key points

- Strong predictions does not mean stable positions
- The importance of transaction cost, risk management and other realities

Our contributions

- Implemented Market Impact Model to measure market impact of any arbitrary trade
- Two ways to solve quadratic transaction costs in portfolio optimization

Link to code: github.com/mingboi95

Thank You

Appendix

Why not incorporate the higher moments to adjust for downside risk?

It has been done before. I'm not sure on the current market practice for this but literature published in 2011 have looked into different solutions used by hedge funds for portfolio optimization, one of which is optimizing for the third and higher order moments

- Bruder et al (2011) Lyxor Asset Management - Portfolio Allocation of Hedge Funds
  - Inclusion of higher-order moments at the objective function level does not solve problem but destroyed value in terms of loss due to the estimation errors when using higher-order moment estimators
  - Loss of value persists even when robustness techniques are used at estimation level

Factor Investing

Fama and French highlighted that investors must be able to ride out the extra short-term volatility and periodic underperformance that could occur in a short time. Investors with a long-term time horizon of 15 years or more will be rewarded for losses suffered in the short term. Using thousands of random stock portfolios, Fama and French conducted studies to test their model and found that when size and value factors are combined with the beta factor, they could then explain as much as 95% of the return in a diversified stock portfolio.

1. Market (Systematic risk of the stock as compared to the market as a whole sensitivity)
2. Size (Smaller cap stocks generate higher returns)
3. Value (Undervalued stock generates higher returns as market acknowledges its value, because investors are currently paying for company relative to its network worth)

Why is CAPM a Factor Investing approach

You can look at it as a Single Factor model, because Sharpe modelled returns solely on the market risk factor.

Fama-French expands on the Capital Asset Pricing Model by incorporating size and value factors

Carhart added momentum factor into Fama-French Three Factor Model

Why R2

Using other metrics such as MSE or MAPE will lead to optimistic results over the data we are fitted over. In addition, in the Gu et al paper, the R2 metric was used because it provides a convenient way of finding out important features by calculating the reduction in panel predictive R2 setting all values of predictor j to zero, while holding remaining estimates fixed to estimate the marginal relationship of expected return and the factor. For consistency, we adopt the R2 as an evaluation metric as well.

Anatomy of a live trade

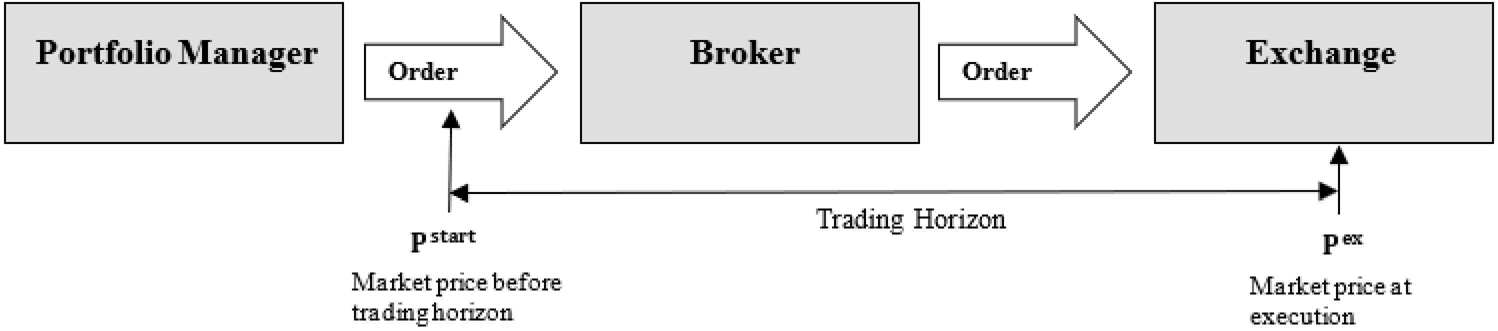


Figure 1 – Typical institutional investor trading process

- The trade process starts with a portfolio manager signally the execution of a trade as a result of their investment process: which stocks to buy and sell, the number of the stock to be traded, and the trade horizon: the duration they want the trade to be executed by, before the trade gets executed by proprietary trading algorithms.

Trading horizon

- Transaction costs can be mitigated by executing your trades over a longer time horizon.
- (Frazzini et al – 2018) AQR's average ex-ante expected trade horizon is 2.7 days, with the median trade taking place over 1.7 days and the maximum taking 9.8 days
- Most trades are executed within a day and 99% are executed within three days

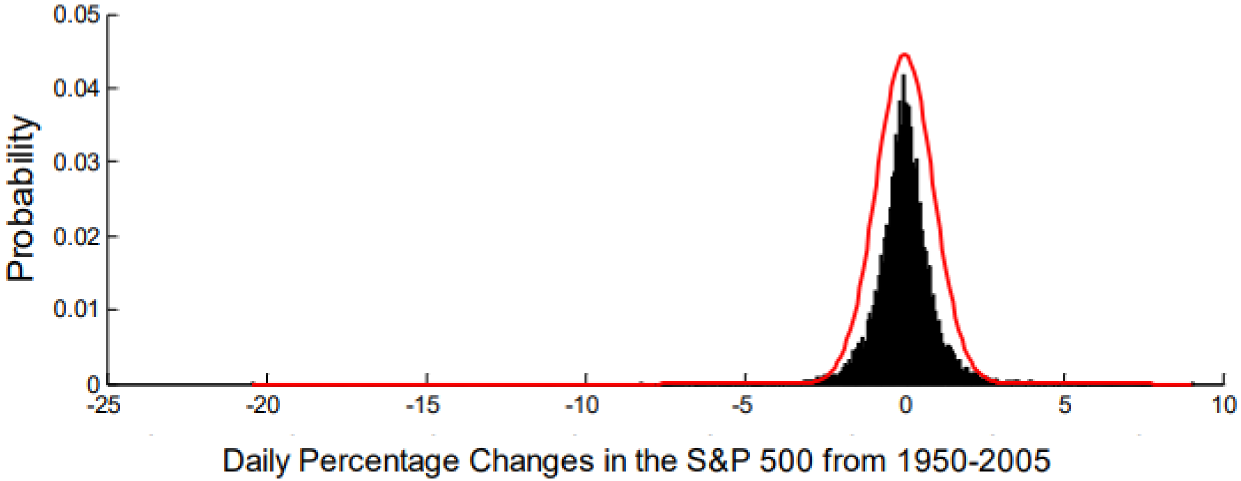
ARP in Covid

- ARP strategies did not do well in Covid, in fact the quantitative funds running the strategies suffered heavy losses. You can see it from the ARP strategies crashing in Q1-2019 as well.
- While our models are relatively quite simple, we contend that machine learning models is not able to augment investment strategies to the point it can be can beat the market consistently.

Literature on the distribution of returns

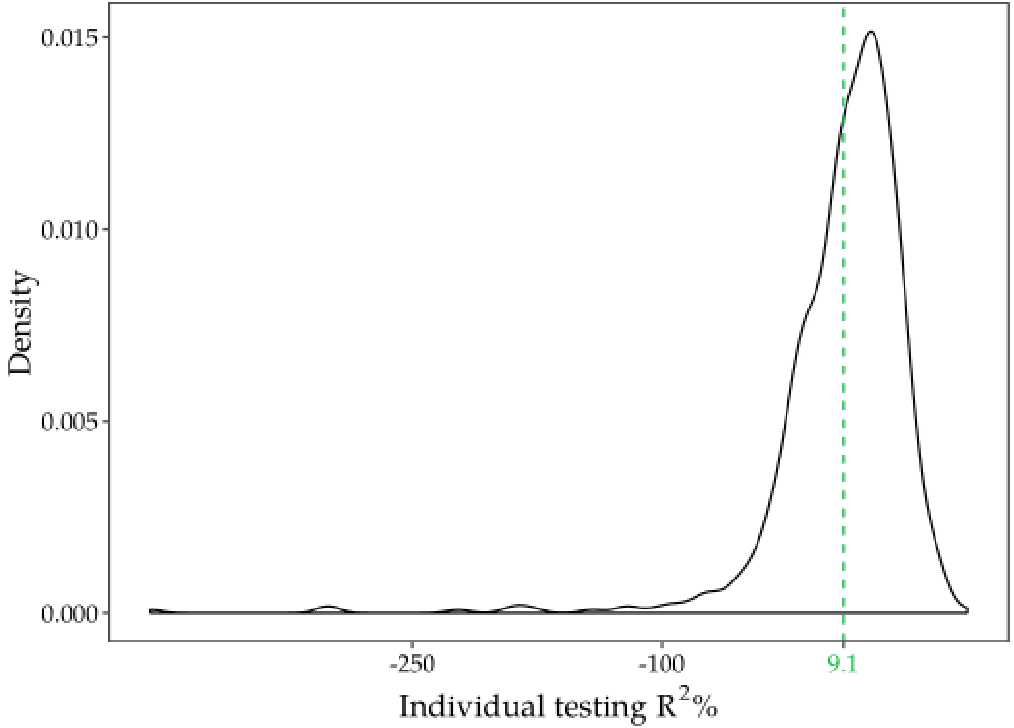
Lauprete, G. J., Samarov, A. M., & Welsch, R. E. (2002). Robust portfolio optimization. Metrika, 55(1), 139-149. https://doi.org/10.1007/s001840200193

Egan, William J., The Distribution of S&P 500 Index Returns (January 6, 2007).



Empirical evidence on distribution of returns

Achal (2020)



Ledoit-Wolf Covariance Shrinkage Estimator

(Ledoit & Michael, 2003) set all pairwise correlations to the mean of all sample correlations, this means that the shrinkage target matrix F is given by: $f_{ii} = s_{ii}, f_{ij} = \bar{r}\sqrt{s_{ii}s_{jj}}$, where s_{ij}^2 is the sample covariance between two stocks, of the sample covariance matrix S

Let y_{it} be the returns of the i^{th} equity at time t, and T be the number of price datapoints.

$\hat{\pi}$ estimates the sum of asymptotic variances of S, scaled by $\sqrt{T}$

$$\hat{\pi} = \sum_{i}^{N}\sum_{j}^{N}\widehat{\pi_{ij}} \text{ with} \approx \widehat{\pi_{ij}} = \frac{1}{T}\sum_{t=1}^{T}\{(y_{it} - \bar{y}_i)(y_{jt} - \bar{y}_j) - s_{ij}\}^2$$

$\hat{\rho}$ estimates the sum of asymptotic covariance of entries of F with entries of S, scaled by $\sqrt{T}$

$$\hat{\rho} = \sum_{i=1}^{N}\widehat{\pi_{ij}} + \sum_{i=1}^{N}\sum_{j=1, j\neq i}^{N}\frac{\bar{r}}{2}\left(\sqrt{\frac{s_{jj}}{s_{ii}}}\widehat{\theta_{ii,ij}} + \sqrt{\frac{s_{ii}}{s_{jj}}}\widehat{\theta_{jj,ij}}\right)$$

Where $\widehat{\theta_{ii,ij}} = \frac{1}{T}\sum_{t=1}^{T}((y_{it} - \bar{y}_i)^2 - s_{ii})((y_{it} - \bar{y}_i)(y_{jt} - \bar{y}_j) - s_{ij})$

$\hat{\gamma}$ estimates misspecification of the F

$$\hat{\gamma} = \sum_{i=1}^{N}\sum_{j=1}^{N}(f_{ij} - s_{ij})^2$$

Then the optimal shrinkage coefficient δ is then estimated by the estimator $\hat{\kappa} = \frac{\hat{\pi} - \hat{\rho}}{\hat{\gamma}}$

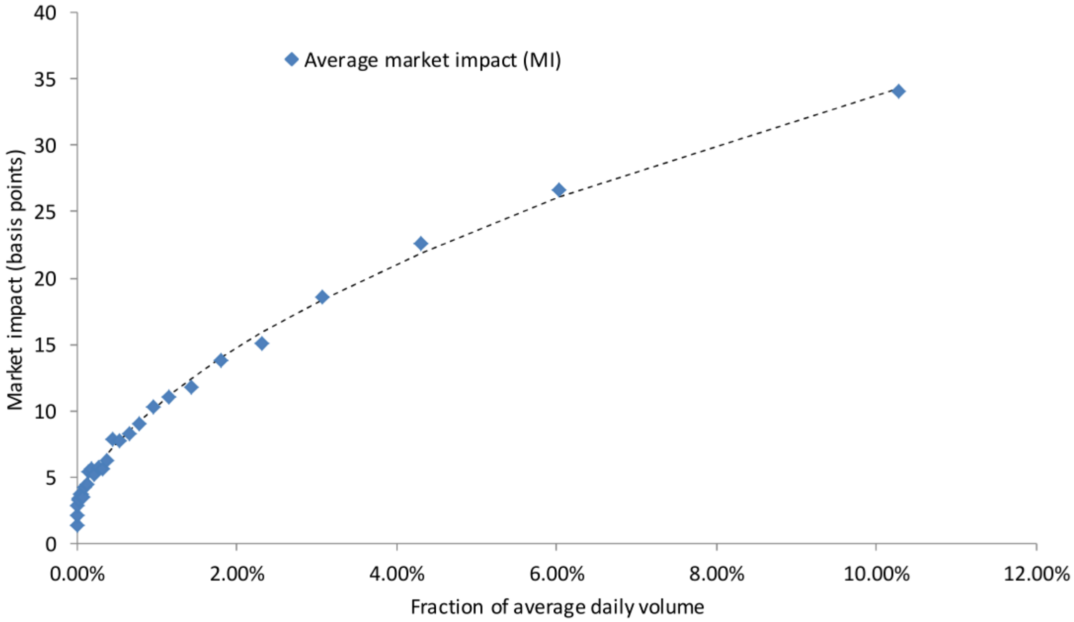
$$\hat{\delta} = \max\left\{0, \min\left\{\frac{\hat{\kappa}}{T}, 1\right\}\right\}$$

The truncation is necessary because the values can be greater than 1 or less than 0. (Ledoit & Michael, 2003) provided empirical evaluation of their Ledoit-Wolf Shrinkage method, as compared with other risk models, with promising results. To this end, we also provide our own empirical findings with the Sample Covariance and Ledoit-Wolf Shrinkage Risk Model.

Live Trade Database

Figure 5. Market Impact by Fraction of Trading Volume

This figure plots the average market impact (MI) for actual live trades from our execution database. We sort all trades into 30 bins based on their fraction of daily volume and compute average market impact for each bucket. This table includes all available developed market equity transactions (cash equities and equity swaps) in our data between August 1998 and June 2016. Market impact is in basis points.



Won’t using AQR’s database not generalize well

The question remains, however, which cost estimates are a better reflection of the real-world costs facing a generic large institutional trader? On the one hand, our data comes from a large institutional trader who executes trades patiently as opposed to TAQ data that is an average of all trades. This suggests our cost estimates may more closely resemble the real-world costs facing a large trader. On the other hand, our data comes from a single institutional manager. So, if our cost estimates are unique to our specific manager (e.g., if our manager simply experiences lower trading costs than the average trader), then they may not be a good estimate of costs more generally. To answer this question, Frazzini et al did out-of-sample testing using both models on passive funds that closely track an index (S&P and FTSE Russell 100) to see if models can produce expected trading cost that approximate the actual cost the funds incur.

So what they found is that using other models in literature actually estimates an inflated trading cost, as compared to their model, which delivers cost estimates cost to actual fund costs

Market Impact Model

Theta	Name	Coef	Economic theory
θ_2	time-trend: June 1926 is 1, June 1927 is 2, ...	-0.01	Transaction costs have declined over time, largely driven by technological events such as moving to decimalization in traded places. (Frazzini, Ronen, & Moskowitz, 2018)
θ_3	market capitalization: log(1 + Market Capitalization)	-0.14	Larger stocks face lower price impact costs, which is also intuitive as the liquidity large stocks is typically much greater than for smaller stocks. A large trade size in illiquid stocks will result in much larger transaction costs, consistent with both literature (Kyle, 1985) and many empirical estimates (Hasbrouck, 2009)
	fraction of daily volume	-0.53	Larger trades have higher transaction costs, consistent with price impact and transaction cost models (Kyle, 1985). The relationship between price impact and trade size shows a concave shape, with costs growing significantly at very large sizes
	sqrt of fraction of daily volume	11.21	
	idiosyncratic volatility: the standard deviation of the residuals of one-year daily stock returns regressed against the benchmark index	0.31	Idiosyncratic Volatility measures the given stock's fluctuations as being larger or smaller than that of the benchmark
θ_7	VIX: monthly variance of the CRSP value weighted index representing market level of volatility	0.12	VIX represents the market's expectations for volatility over the coming 30 days, a measure for level of risk, fear, or stress in the market when making investment decisions. More volatile firms have higher transaction costs, consistent with models of market maker inventory risk, and more volatile market environments are also associated with larger price impact costs, consistent with market makers needing to be compensated more in more volatile markets.

Frazzini et al (2018) Market Impact Model:

$$MI = a + b * c + c * sign(x)\sqrt{|X|}$$

Where:

a = $\theta_2$time-trend + $\theta_3 log$ (1+ market cap) + $\theta_6 \sigma^{IV}$ + θ_7VIX,
a controls the general level of the trading and the market characteristics
b = θ_4 (fraction of daily average volume)
c = θ_5 (sqrt of fraction of daily average volume)
a and b controls the characteristics of the trade with respect to the stock
x: 100*m/dtv signed dollar volume of the trade (m) as a fraction of stock's average one-year dollar volume

Why the constraints?

- Net sector exposure <= +-10% using Level 1 GICS: to diversify the portfolio and to avoid the optimizer from over tilting allocation to any particular.
- 0.5% <= | Individual Asset Weights |<= 1.5%With trade sizes being a factor in transaction cost, the second constraint is there as a heuristic to avoid very small or very large positions in any asset.
- DAV <= 10%: to ensure that the portfolio does not overly dominate a market for very small stocks, and it is also consistent with reports of an institutional investor's own trades (Frazzini, Ronen, & Moskowitz, 2018), where most of their reported trades do not exceed this threshold either, either as a direct result of the investor's asset allocation strategies, or as an indirect result of their optimization strategies.

Why 10bps of one-way trade?

- Why not use the Market Impact model to estimate trading costs?
- If you use market impact model to estimate the trading cost, and then your optimization is to minimize the same function as the objective functions, obviously the results from that optimization will be the best result.

Single-Period Optimization

- Mean-Variance Optimization is a single-period optimizer, mean that you balance it at time i and time $i+1$ and that it doesn't automatically adjust to new information.
- You will need to decide on some period over which views will be aggregated and return vectors constructed, then use mean-variance to make an allocation.
- At the end of this period, you will have to repeat the process (including your new data) and rebalance positions.
- A multi-period optimization allows us to consider if the asset allocation we make in current period put us in good/bad position to make trade in future periods.
- Suppose: Our optimization tell us go long in a rarely-traded asset, thus unwinding the position will incur large transaction costs.
- Even if it is optimized after consider transaction cost, optimizing for a single period models cost of rebalancing the portfolio to move into that position, but not the cost of moving out of the position when you rebalance the portfolio at the next time horizon

Market Regime and Style Timing

Regime-switching Dynamic Correlation

- Hedge funds that uses feature allocation models which integrate the correlation dynamics between the returns perform better, because they are able to capture changes in market regime and react accordingly
- https://github.com/microprediction/mkalgo
- Which historical segment best matches the current k-length period, generalized to find sets of motifs rather than pairs efficiently
- Moving further away from Statistics and more into economic structure of the market