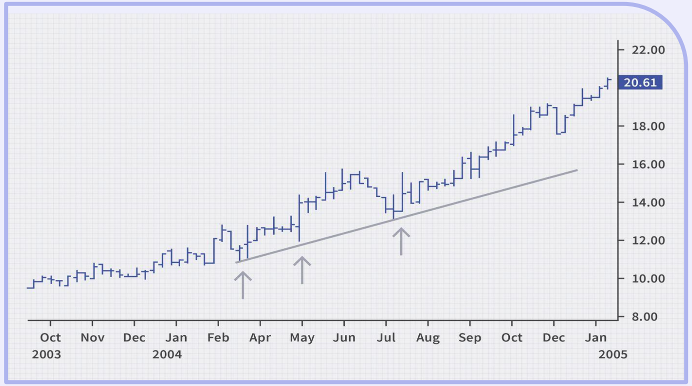
Systematic Trading Strategy Design (feat. Momentum)

Paragon Global Investments

Harrison Wang and Nick Klatsky

The Goal

Design a very simple (momentum) signal and discuss facets of general algorithmic trading strategy design.



What is Momentum?

"An object in motion stays in motion"

- **Fundamental Human Instincts**
- Claim: market trends persist
- Buy stocks which historically outperform
- Sell stocks which historically underperform
- Is this strategy profitable?

Constructing the Momentum Signal

- How do we assign a value to momentum?
  - For asset i and time t, let $P_{i,t}$ be the price of the asset at time t
  - Then, define momentum to be the returns over some lookback window

$$Mom_{i,t} = \frac{P_{i,t}}{P_{i,t-n}} - 1$$

- Modified signals
  - Short term vs. long-term momentum
    - We backtest long-term momentum (t represents time in months)
  - Lookback window
    - Test n = 3, 6, 9, 12 (months)
  - Remove reversal effect of previous month
    - Returns excluding past month

$$R_{i,t}^{t-12:t-1} = \frac{P_{i,t-1}}{P_{i,t-12}} - 1$$

Preliminary Signal Strength Testing

- Each month, sort stocks into 10 deciles depending on momentum value, calculate average returns

Panel B: Value-Weighted Portfolio Returns

Sort Variable	1	2	3	4	5	6	7	8	9	10	10-1	CAPM α	FF α
Mom	−0.76	−0.12	0.04	0.35	0.39	0.37	0.53	0.67	0.75	1.18	1.95	2.13	2.37
											(5.39)	(6.48)	(7.54)
R^{12M}	−0.36	0.01	0.26	0.39	0.37	0.43	0.54	0.62	0.73	1.05	1.42	1.66	1.93
											(3.64)	(4.81)	(5.82)
R^{9M}	0.02	0.40	0.40	0.40	0.43	0.44	0.44	0.58	0.65	0.97	0.95	1.24	1.45
											(2.51)	(3.67)	(4.48)
R^{6M}	0.21	0.48	0.48	0.50	0.53	0.47	0.51	0.47	0.53	0.78	0.58	0.85	0.98
											(1.71)	(2.76)	(3.30)
R^{3M}	0.33	0.54	0.60	0.60	0.61	0.51	0.46	0.48	0.47	0.63	0.31	0.57	0.69
											(1.07)	(2.08)	(2.40)
$R^{t-12:t-1}$	−0.71	−0.12	0.19	0.31	0.42	0.40	0.48	0.66	0.74	1.08	1.79	1.96	2.23
											(5.12)	(6.21)	(7.33)
$R^{t-6:t-1}$	−0.52	0.22	0.38	0.40	0.51	0.48	0.52	0.51	0.54	0.97	1.50	1.71	1.86
											(4.11)	(5.14)	(5.91)

Source: Bali et al., *Empirical Asset Pricing: The Cross Section of Stock Returns*, 2016

Sort Variable	1	10	10-1	CAPM α	FF α
Mom	−0.76	1.18	1.95	2.13	2.37
			(5.39)	(6.48)	(7.54)
R^{12M}	−0.36	1.05	1.42	1.66	1.93
			(3.64)	(4.81)	(5.82)
R^{9M}	0.02	0.97	0.95	1.24	1.45
			(2.51)	(3.67)	(4.48)
R^{6M}	0.21	0.78	0.58	0.85	0.98
			(1.71)	(2.76)	(3.30)
R^{3M}	0.33	0.63	0.31	0.57	0.69
			(1.07)	(2.08)	(2.40)
$R^{t-12:t-1}$	−0.71	1.08	1.79	1.96	2.23
			(5.12)	(6.21)	(7.33)
$R^{t-6:t-1}$	−0.52	0.97	1.50	1.71	1.86
			(4.11)	(5.14)	(5.91)

Takeaways

Mom = $R^{t-11:t-1}$

- Sorting by longer-term momentum signal generates profitable portfolios
- Excluding the previous month increases signal strength
- Best signal outperforms market by 2.13% on average

Backtesting Trading Strategy

Momentum-specific parameters:

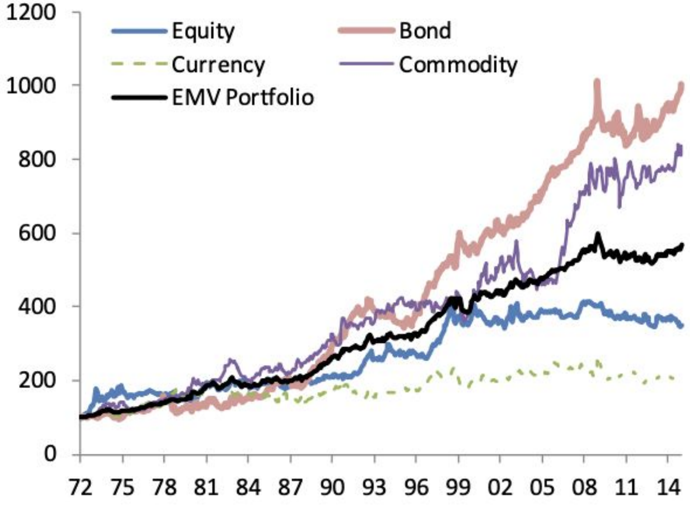
- Signal characteristics: lookback window, exclusion of past month, frequency
- Relative (Cross-sectional) vs. Absolute (Time Series) Momentum
  - Compare stocks to each other or to risk-free rate?

General algorithmic design parameters:

- Asset class
- Long-only vs. long-short
- Rebalancing frequency
  - Balance t-costs with signal capture
- Allocation of portfolio
  - Equal-weighted, rank-weighted, value-weighted, mean-variance optimal, risk-parity
- Other add-ons
  - Hedging, signal overlay, discussed later…

Example: Asset Class

Figure 1: Performance of Prototype Momentum Risk Factors by Asset Class



Source: J.P. Morgan Quantitative and Derivatives Strategy.
* For comparison purpose, Equity, Bond, Currency and Commodity prototype Momentum strategies are scaled to have an ex-post volatility of 10% per annum in the chart.

Table 1: Performance and Risk Statistics for Prototype Momentum Factors

	Equity	*Bond*	*Curncy*	*Comdty*	*EMV*
Excess Ret (%)	4.6	3.5	1.7	7.3	4.1
STDev (%)	17.7	6.2	9.0	15.1	5.3
MaxDD (%)	-37.5	-19.1	-27.9	-33.3	-13.2
MaxDDur (yrs)	16.4	6.0	15.2	5.7	6.0
***t*-Statistic**	**2.2**	**3.9**	**1.5**	**3.6**	**5.2**
Sharpe Ratio	**0.26**	**0.57**	**0.18**	**0.49**	**0.78**
Hit Rate (%)	51.7	63.4	56.8	56.2	63.8
Skewness	0.51	-0.10	-0.33	0.01	-0.32
Kurtosis	6.17	3.67	1.87	1.45	1.13

Source: J.P. Morgan Quantitative and Derivatives Strategy.

Performance Evaluation

Return Characteristics

- Excess returns (over benchmark)
- Factor return contribution
- Leftover alpha

Other metrics

- Sharpe ratio = excess return / risk
- Information ratio = alpha / stdev($e_{i,t}$)
- Turnover/transaction costs
- Stress-testing (2008, 2020)

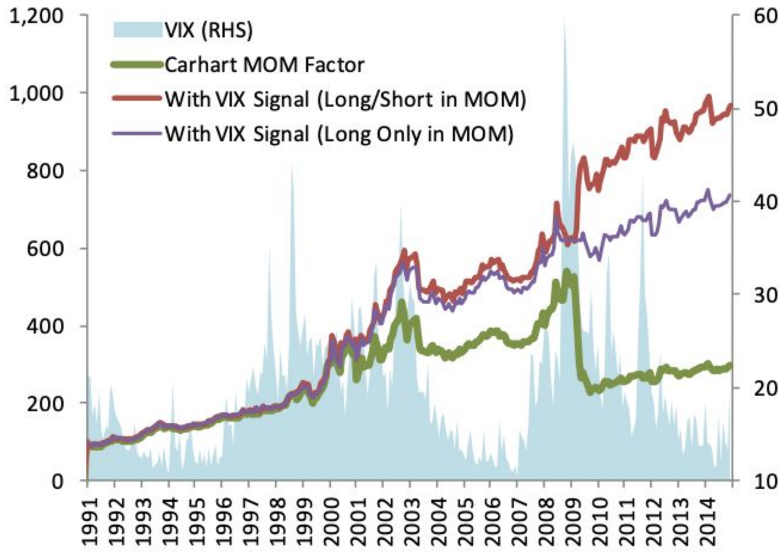
Risk Characteristics

- Volatility (downside vol)
- Factor betas
- Idiosyncratic risk
- Max drawdown
- Max monthly loss
- VaR
- Skewness, Kurtosis

$$R_{i,t} - R_{ft} = \alpha_{(5\text{-factor})} + \beta_{i,MKT}(R_{m,t} - R_{ft}) + \beta_{i,SMB} SMB + \beta_{i,HML} HML + \beta_{i,RMW} RMW + \beta_{i,CMA} CMA + \varepsilon_{i,t}$$

Improving the Strategy (Signal Overlay)

Figure 42: Enhanced Carhart Momentum strategy with VIX signal during 1991-2014



Source: J.P. Morgan Quantitative and Derivatives Strategy.

When does momentum fail?

- Reversals
  - Due to the persistence of beta, highest outperformers become worst underperformers
  - Large market reversal in 2009 leads to huge losses (~550 → ~200)

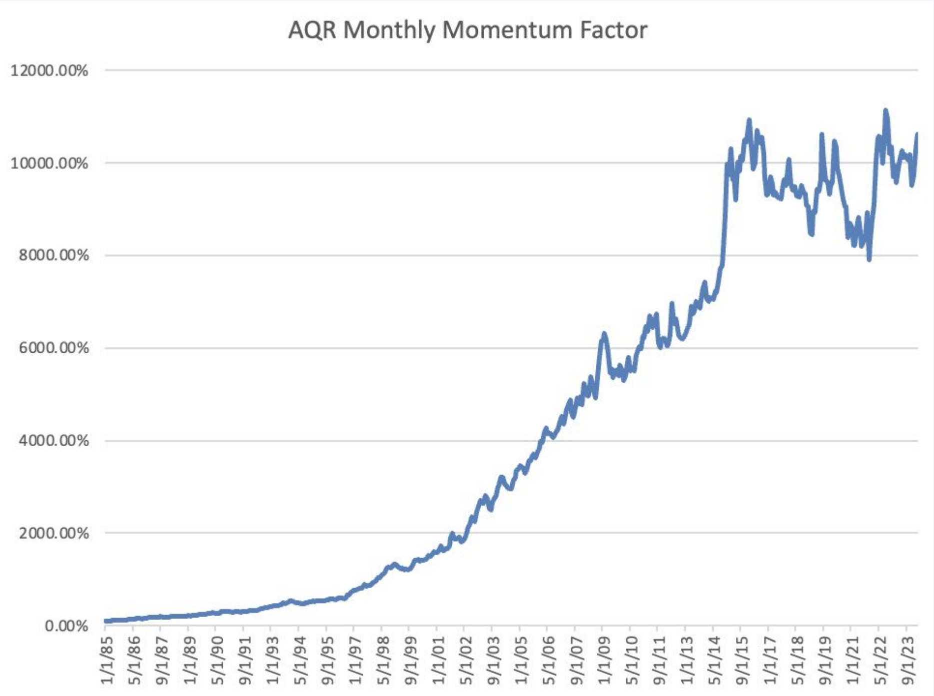
How do we address this?

- Signal Overlay
  - Exit all positions when VIX high
- Hedging
  - Purchase options to hedge downside risk

Momentum Losing Momentum

Efficient Market Hypothesis

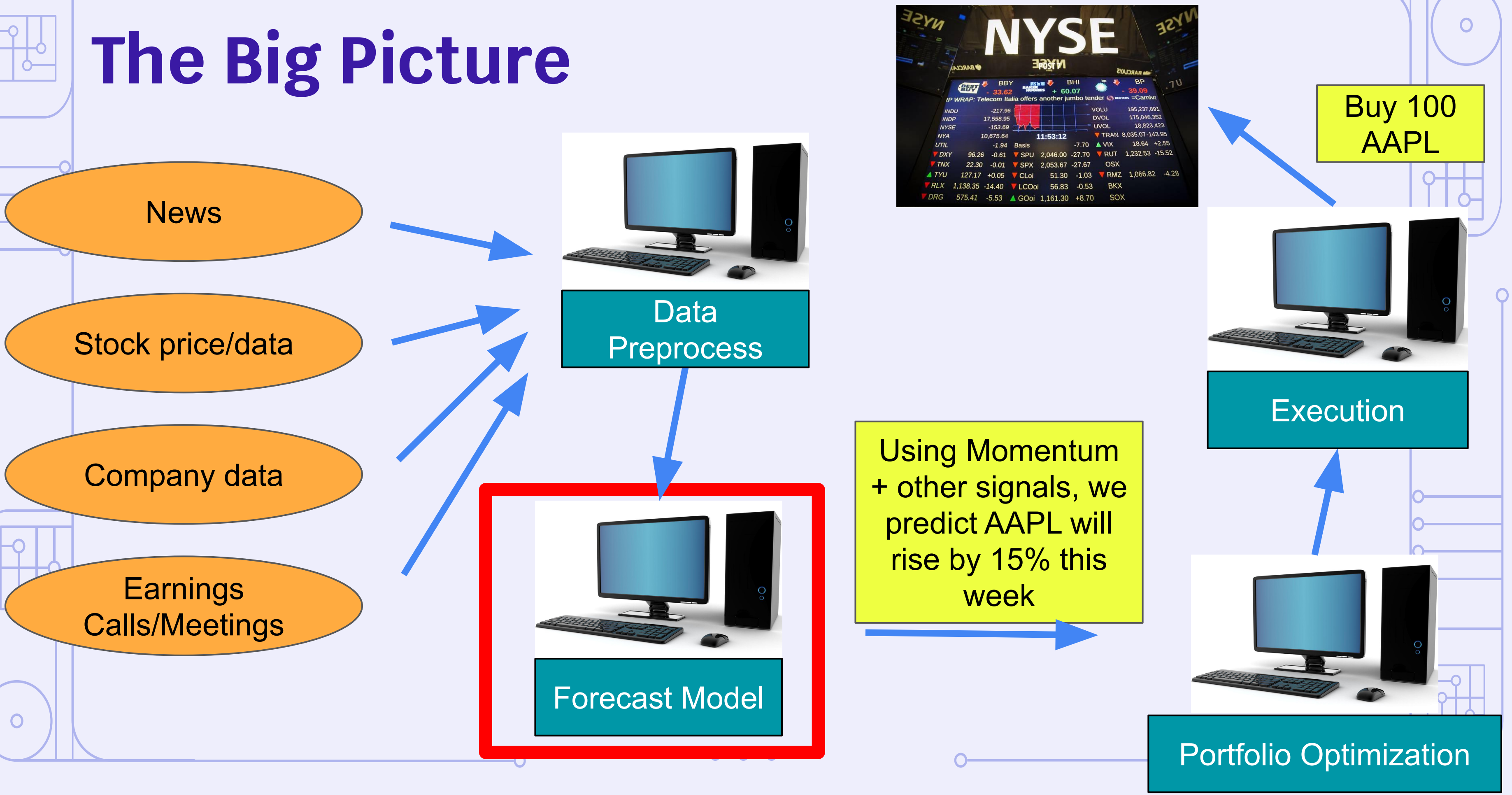
- Every profitable signal loses profitability as information becomes more widely accessible
- Simple momentum loses profitability around 2015



Summary (and Improvements)

1. Identify some hypothesis about market motivated by experience
2. Construct signal from data sources
    a. Use alternative/non-public data sources
    b. Can use ML/Deep Learning in this step to generate signal
        i. Ex: Sentiment signals from Reddit post data using NLP
3. Conduct a grid-search through various parameters
    a. Asset universe, type of signal, portfolio allocation, etc.
4. Evaluate each strategy backtest using quantitative metrics
5. Implement strategy (with transaction cost modeling)

The Big Picture



Ex: Signals as Inputs to Factor Models

In general, we might not be interested in a signal solely for the purpose of directly using it to predict asset prices – we can also use our signals to enrich other models

Examples:

- Use momentum as an input to a factor model, then use the factor returns:
  - For dimensionality reduction
  - As covariates to condition on
  - Something else … be creative!

Example 1: Portfolio Optimization + Covariance Matrix Forecasting

Mean-Variance Optimization Problem at Time t

Given a portfolio of n assets with expected returns $\mathbf{E}[r_t]$ and a covariance matrix of returns Σ_t at time t, the mean-variance optimization aims to find the portfolio weights $\mathbf{w}$ that minimize the portfolio variance subject to a target expected return, or maximize the expected return for a given level of risk.

Objective: Minimize the portfolio variance for a given expected return R_{target},

$$\min_{\mathbf{w}} \mathbf{w}^T \Sigma_t \mathbf{w}$$

Subject to:

1. The portfolio weights sum to one:

$$\sum_{i=1}^{n} w_i = 1$$

2. The expected return of the portfolio equals the target return:

$$\mathbf{w}^T \mathbf{E}[r_t] = R_{target}$$

Example 1: Portfolio Optimization + Covariance Matrix Forecasting

We need the conditional covariance matrix in order to find the optimal portfolio– however, forecasting it via a time series regression (ex: AR(p), M-GARCH) can be difficult because the covariance matrix is high-dimensional (d(d+1)/2 unique features for d assets). Luckily, we can shrink the dimensionality!

Factor Model

$$r_{i,t} = \beta_0 + \beta_{i1,t} f_{1,t} + \beta_{i2,t} f_{2,t} + \cdots + \beta_{iK,t} f_{K,t} + \epsilon_{i,t}$$

$$\Sigma_t = B_t \Sigma_{f,t} B_t' + \Sigma_{\epsilon,t}$$

Then, forecast $\Sigma_{f,t}$ via an autoregressive (or other) method (with k(k+1)/2 features!)

Example 2: Bayesian Linear Regression with AR–GARCH Errors

Model the return of currency y with an AR-GARCH error process:

$$y_t = x_t \gamma + u_t, \quad t = 1, \ldots, T;$$

where x_t is a $kx1$ vector of factor returns (or other covariates) at time t, γ is a vector of regression coefficients (factor loadings), and u_t is an error term:

$$u_t = \sum_{j=1}^{p} \phi_j u_{t-j} + \varepsilon_t, \quad \varepsilon_t | \mathcal{F}_{t-1} \sim N(0, \sigma_t^2);$$

and σ_t^2 follows a GARCH process:

$$\sigma_t^2 = \omega + \sum_{j=1}^{r} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2;$$

Define:

$$\Phi = (\phi_1, \ldots, \phi_p)^\top,$$

$$\Omega = (\omega, \alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s)^\top,$$

Example 2: Bayesian Linear Regression with AR-GARCH Errors

Lastly, define δ as the vector of all unknown parameters:

$$\delta = (\gamma, \Phi, \Omega)$$

The joint posterior density of unknown parameters δ is proportional to the likelihood of observing δ given data times the prior distribution on δ:

$$p(\delta \mid Y, X) \propto \ell(\delta \mid Y, X) p(\delta)$$

where $\ell(\cdot)$ is the likelihood function given by:

$$\ell(\delta \mid Y, X) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{(y_t - x_t\gamma - \Phi(L)(y_t - x_t\gamma))^2}{2\sigma_t^2}\right);$$

Then simulate from the posterior distribution via MCMC to estimate the model parameters!

Thanks for watching!