# The Anatomy of Machine Learning-Based Portfolio Performance

Philippe Goulet Coulombe[1] Dave Rapach[2] Erik Christian Montes Schütte[3] Sander Schwenk-Nebbe[4]

[1]UQAM [2]Federal Reserve Bank of Atlanta [3]Aarhus University [4]LogicBanker

EDHEC Speaker Series—The Future of Finance—Season 5

February 3, 2026

- **Disclaimer** $\Rightarrow$ The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any errors are the authors' responsibility.

- Return predictability is a leading topic in empirical asset pricing
- Out-of-sample tests are now routinely employed
  - Most rigorous/informative tests in the era of big data and ML (Nagel 2021; Martin & Nagel 2022)
- In addition to statistical accuracy, it is now commonplace to analyze the **economic value** of return predictability via asset allocation exercises
  - Return forecasts based on a (large) set of predictors are generated and serve as inputs for constructing a portfolio
  - Portfolio performance metrics are computed over a forecast evaluation period (and often compared to a benchmark) to measure the economic value of return predictability

- Spate of recent studies employs a multitude of firm characteristics and ML to forecast out-of-sample cross-sectional stock returns (eg, Freyberger, Neuhierl & Weber 2020; Gu, Kelly & Xiu 2020; Avramov, Cheng & Metzker 2023; Han et al 2024)
    - Construct a long-short portfolio by sorting stocks according to their return forecasts for the next month
        - Go long (short) stocks with the highest (lowest) return forecasts
    - Long-short portfolios based on ML provide substantive economic value to investors
        - Strong evidence of cross-sectional stock return predictability
- However, the existing literature does not provide a general methodology for measuring exactly **how** individual or groups of predictors in fitted ML models contribute to economic value

- Fill this gap by developing a method based on Shapley (1953) values to directly measure the contributions of individual or groups of predictors to portfolio performance
  - Use the logic of Shapley values to fairly allocate the contributions of the predictors in fitted prediction models to portfolio performance
    - Shapley values have a number of attractive properties, making them one of, if not the most, popular model interpretation tools in machine learning
  - Exactly decompose portfolio performance in terms of the underlying predictors $\Rightarrow$ **anatomize** economic value
- New measure $\Rightarrow$ **Shapley-based portfolio performance contribution** ($\text{SPPC}_p$ for predictor $p$)
  - ML model interpretation tool for finance to peer inside the "black box" and understand the roles of individual or groups of predictors in determining the economic value of return predictability

- Extend conventional Shapley values to measure the contributions of predictors to the following:
  1. Out-of-sample return forecast
  2. Portfolio return
  3. Portfolio performance metric $\Rightarrow$ resulting in the SPPC
- SPPC is very flexible
  - Model agnostic (ie, it can be applied to any prediction model)
  - Can be used for any allocation strategy for mapping return forecasts to portfolio weights
  - Can be computed for any portfolio performance metric

- Illustrate the use of the SPPC in an empirical application investigating the economic value of cross-sectional stock return predictability
- Generate monthly individual stock return forecasts using 207 firm characteristics from Chen & Zimmermann (2022) and the XGBoost ML algorithm (Chen & Guestrin 2016)
  - Sort stocks into quintiles based on the XGBoost return forecasts and go long (short) the fifth (first) quintile, where each leg is value weighted
- Long-short portfolio performs impressively
  - Annualized Sharpe ratio of 1.80
  - Large alphas in the context of leading multifactor models

- Place the individual characteristics into 20 groups and use the $\text{SPPC}_p$ to measure the contributions of the predictor groups to portfolio performance
- **Risk**, **Earnings**, **Seasonal momentum**, **Momentum**
  - Make the largest positive contributions to portfolio performance over the full 1973:01–2021:12 out-of-sample period
- **Earnings**, **Seasonal momentum**, **Investment**
  - Make positive/sizable contributions on a consistent basis over time
- **Sales**, **Ownership**
  - Make negative contributions consistently over time

- Shapley values ⇒ canonical solution method for fairly allocating the payoff to the players in a coalitional game
- In the context of machine learning, we want to fairly allocate the contributions of the predictors to a model's prediction (nontrivial task)
    - Exploit the analogy between the players in a coalitional game earning a payoff and the predictors in a model making a prediction
- Štrumbelj & Kononenko (2010, 2014) and Lundberg & Lee (2017) show how Shapley values can be used to allocate the contributions of the predictors to a model's prediction
    - Adapt Štrumbelj & Kononenko (2014) to a panel setting where a model generates individual stock return predictions based on a set of firm characteristics (but our methodology can be applied to any prediction model)

- Index set of predictors/features $\Rightarrow S = \{1, \ldots, P\}$
  - Index individual predictors by $p$
- Index set of cross-sectional units $\Rightarrow C = \{1, \ldots, N\}$
  - Index individual cross-sectional units by $i$
- $P$-vector of predictors for stock $i$ in period $t \Rightarrow \boldsymbol{x}_{i,t} = \begin{bmatrix} x_{1,i,t} & \cdots & x_{P,i,t} \end{bmatrix}'$
- Return on stock $i$ in period $t \Rightarrow r_{i,t}$
- Prediction model $\Rightarrow r_{i,t+1} = \underbrace{f(\boldsymbol{x}_{i,t})}_{\substack{\text{conditional} \\ \text{expectation} \\ \text{function}}} + \underbrace{\varepsilon_{i,t+1}}_{\substack{\text{additive} \\ \text{error} \\ \text{term}}}, \quad \hat{f} \Rightarrow$ fitted model
- Training window of panel data observations $\Rightarrow W_j = \{t_{j,\text{start}}, \ldots, t_{j,\text{end}} - 1\}$
- Fitted prediction model evaluated at $\boldsymbol{x}_{i,t}$ trained using $W_j \Rightarrow \hat{f}(\boldsymbol{x}_{i,t}\,; W_j)$

- Shapley value for model interpretation
  - Marginal contribution of $x_{p,i,t}$ to $\hat{f}(\boldsymbol{x}_{i,t}\,;W_j)$ given $S\setminus\{p\} \Rightarrow$

$$\phi_p(\boldsymbol{x}_{i,t}\,;W_j) = \sum_{Q\subseteq S\setminus\{p\}} \underbrace{\frac{|Q|!(P-|Q|-1)!}{P!}}_{\text{coalition } Q\text{'s weight}} \underbrace{\left[\xi_{Q\cup\{p\}}(\boldsymbol{x}_{i,t}\,;W_j) - \xi_Q(\boldsymbol{x}_{i,t}\,;W_j)\right]}_{\text{predictor } p\text{'s marginal contribution}}$$

  for $p\in S$, $i\in C$, $t\in W_j$

    - $Q \Rightarrow$ subset of predictors (ie, coalition)
    - $Q\subseteq S\setminus\{p\} \Rightarrow$ set of all possible coalitions of $P-1$ predictors in $S$ that exclude $p$
    - $|Q| \Rightarrow$ cardinality of $Q$
    - $\xi_Q(\boldsymbol{x}_{i,t}\,;W_j) \Rightarrow$ value function

- Different value functions in terms of predictor removal define different Shapley values (eg, Chen et al 2020; Chen et al 2023)

- Use the **interventional** Shapley value (Datta, Sen & Zick 2016; Janzing, Minorics & Blöbaum 2020; Sundararajan & Najmi 2020) $\Rightarrow$

$$\xi_Q(\boldsymbol{x}_{i,t}\,;W_j) = \mathbb{E}_{\mathbb{P}(x_{p,i,t}\,:\,p\notin Q)}\left[\hat{f}(x_{p,i,t}\ :\ p\in Q, x_{p,i,t}\ :\ p\notin Q\,;W_j)\right]$$

  - Remove the predictors via the marginal distribution of the predictors not in the coalition $Q$
  - "True to the model" (Chen et al 2020)

- Like Shapley values from coalitional game theory, interventional Shapley values have a number of attractive properties, including **efficiency** (aka **local accuracy**) $\Rightarrow$

$$\sum_{p\in S}\phi_p(\boldsymbol{x}_{i,t}\,;W_j) = \hat{f}(\boldsymbol{x}_{i,t}\,;W_j) - \underbrace{\mathbb{E}\left[\hat{f}\,;W_j\right]}_{\text{baseline prediction}} \quad \text{for } i\in C,\ t\in W_j$$

  - Shapley values provide an **exact** decomposition of $\hat{f}(\boldsymbol{x}_{i,t}\,;W_j)$

- Becomes practically infeasible to exactly compute the Shapley value as the number of predictors increases
  - Štrumbelj & Kononenko (2014) propose an algorithm using the sampling-based approach of Castro et al (2009)
  - Develop a refined version of their algorithm and then extend it to measure the contributions of individual predictors to **portfolio performance**
- Express the Shapley value in an equivalent form $\Rightarrow$

$$\phi_p(\boldsymbol{x}_{i,t}\,;W_j) = \frac{1}{P!}\sum_{\mathcal{O}\in\pi(P)}\left[\xi_{\mathrm{Pre}_p(\mathcal{O})\cup\{p\}}(\boldsymbol{x}_{i,t}\,;W_j) - \xi_{\mathrm{Pre}_p(\mathcal{O})}(\boldsymbol{x}_{i,t}\,;W_j)\right]$$

$$\text{for } p\in S,\ i\in C,\ t\in W_j$$

  - $\mathcal{O} \Rightarrow$ ordered permutation for the predictor indices in $S$
  - $\pi(P) \Rightarrow$ set of all ordered permutations for $S$
  - $\mathrm{Pre}_p(\mathcal{O}) \Rightarrow$ set of indices that precede $p$ in $\mathcal{O}$

- Make a random draw $m$ (with replacement) from $\pi(P)$, denoted by $\mathcal{O}_m$, and compute

$$\hat{\theta}_{p,m}(\boldsymbol{x}_{i,t}\,;W_j) =$$

$$\underbrace{\frac{1}{|C||W_j|}\sum_{u\in C}\sum_{s\in W_j}\hat{f}(x_{k,i,t}: k\in \text{Pre}_p(\mathcal{O}_m)\cup\{p\},\ x_{l,u,s}: l\in \text{Post}_p(\mathcal{O}_m)\,;W_j)}_{\hat{\xi}_{\text{Pre}_p(\mathcal{O})\cup\{p\}}(\boldsymbol{x}_{i,t}\,;W_j)} -$$

$$\underbrace{\frac{1}{|C||W_j|}\sum_{u\in C}\sum_{s\in W_j}\hat{f}(x_{k,i,t}: k\in \text{Pre}_p(\mathcal{O}_m),\ x_{l,u,s}: l\in \text{Post}_p(\mathcal{O}_m)\cup\{p\}\,;W_j)}_{\hat{\xi}_{\text{Pre}_p(\mathcal{O})}(\boldsymbol{x}_{i,t}\,;W_j)}$$

  - $\text{Post}_p(\mathcal{O}) \Rightarrow$ set of indices that follow $p$ in $\mathcal{O}$

- Use "background data" from the training sample $W_j$ to integrate out the predictors not in a coalition $\Rightarrow$ Monte Carlo integration

- Sampling approximation $\Rightarrow \hat{\phi}_p(\boldsymbol{x}_{i,t}\,;W_j) = \frac{1}{2M}\sum_{m=1}^{2M}\hat{\theta}_{p,m}(\boldsymbol{x}_{i,t}\,;W_j)$
  - $M \Rightarrow$ number of draws
  - Antithetic sampling $\Rightarrow$ compute $\hat{\theta}_{p,m}(\boldsymbol{x}_{i,t}\,;W_j)$ for the original order of the randomly drawn ordered permutation and when the order is reversed (Mitchell et al 2022)
- Compute the Shapley value for each predictor $p \in S$ for each random draw $m$ (Castro et al 2009), which ensures local accuracy for $\hat{\phi}_p(\boldsymbol{x}_{i,t}\,;W_j) \Rightarrow$

$$\sum_{p \in S}\hat{\phi}_p(\boldsymbol{x}_{i,t}\,;W_j) = \hat{f}(\boldsymbol{x}_{i,t}\,;W_j) - \underbrace{\bar{\hat{f}}(W_j)}_{\hat{\phi}_{\emptyset}(W_j)} \quad \text{for } i \in C,\ t \in W_j$$

  - Baseline prediction $\Rightarrow \bar{\hat{f}}(W_j) = \frac{1}{|C||W_j|}\sum_{i \in C}\sum_{t \in W_j}\hat{f}(\boldsymbol{x}_{i,t}\,;W_j)$

- To this point, we have computed conventional Shapley values for the in-sample model predictions corresponding to the training sample observations
  - To derive the SPPC, we begin by defining the Shapley value corresponding to an **out-of-sample prediction**
- Suppose that we train a model using window $W_j$ and generate an out-of-sample return forecast for stock $i$ and period $t_{j,\text{end}} + 1$ based on the fitted model $\Rightarrow$

$$\hat{r}_{i,t_{j,\text{end}}+1} = \hat{f}\left(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j\right) \quad \text{for } i \in C$$

- Define the out-of-sample Shapley value for the out-of-sample prediction $\Rightarrow$

$$\phi_p^{OS}\left(\boldsymbol{x}_{i,t_{j,\text{end}}};W_j\right)=\frac{1}{P!}\sum_{\mathcal{O}\in\pi(P)}\left[\xi^{OS}_{\text{Pre}_p(\mathcal{O})\cup\{p\}}\left(\boldsymbol{x}_{i,t_{j,\text{end}}};W_j\right)-\xi^{OS}_{\text{Pre}_p(\mathcal{O})}\left(\boldsymbol{x}_{i,t_{j,\text{end}}};W_j\right)\right]\quad \text{for } p\in S,\ i\in C$$

- Use the sampling algorithm to make a random draw $m$ $\Rightarrow$

$$\hat{\theta}^{OS}_{p,m}\left(\boldsymbol{x}_{i,t_{j,\text{end}}};W_j\right)=\underbrace{\frac{1}{|C||W_j|}\sum_{u\in C}\sum_{s\in W_j}\hat{f}\left(x_{k,i,t_{j,\text{end}}}:k\in\text{Pre}_p(\mathcal{O}_m)\cup\{p\},\ x_{l,u,s}:l\in\text{Post}_p(\mathcal{O}_m);W_j\right)}_{\hat{r}_{i,t_{j,\text{end}}+1,m,p}(\boldsymbol{x}_{i,t_{j,\text{end}}};W_j):=\hat{\xi}^{OS}_{\text{Pre}_p(\mathcal{O})\cup\{p\}}(\boldsymbol{x}_{i,t_{j,\text{end}}};W_j)}-\underbrace{\frac{1}{|C||W_j|}\sum_{u\in C}\sum_{s\in W_j}\hat{f}\left(x_{k,i,t_{j,\text{end}}}:k\in\text{Pre}_p(\mathcal{O}_m),\ x_{l,u,s}:l\in\text{Post}_p(\mathcal{O}_m)\cup\{p\};W_j\right)}_{\hat{r}_{i,t_{j,\text{end}}+1,m,\setminus p}(\boldsymbol{x}_{i,t_{j,\text{end}}};W_j):=\xi^{OS}_{\text{Pre}_p(\mathcal{O})}(\boldsymbol{x}_{i,t_{j,\text{end}}};W_j)}$$

- Sampling approximation $\Rightarrow$

$$\hat{\phi}_p^{OS}(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j) = \frac{1}{2M}\sum_{m=1}^{2M}\hat{\theta}_{p,m}^{OS}(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j) \quad \text{for } p \in S,\ i \in C$$

- Local accuracy holds for $\hat{\phi}_p^{OS}(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j) \Rightarrow$

$$\sum_{p \in S}\hat{\phi}_p^{OS}(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j) = \underbrace{\hat{f}(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j)}_{\hat{r}_{i,t_{j,\text{end}}+1}} - \hat{\phi}_{\emptyset}(W_j) \quad \text{for } i \in C$$

- Exactly decompose $\hat{r}_{i,t_{j,\text{end}}+1}$ (in terms of the deviation from the empty coalition or baseline forecast) into the contributions made by each of the $P$ predictors

- Consider an investor who decides on their allocations across the $N$ stocks for period $t_{j,\text{end}} + 1$ based on the set of return forecasts formed using data through $t_{j,\text{end}}$
  - Allocation to stock $i$ generally depends on the entire set of forecasts for period $t_{j,\text{end}} + 1 \Rightarrow$

$$w_{i,t_{j,\text{end}}+1}\Big(\big\{\underbrace{\hat{f}\big(\mathbf{x}_{i,t_{j,\text{end}}};W_j\big)}_{\hat{r}_{i,t_{j,\text{end}}+1}}\big\}_{i\in C}\Big) \quad \text{for } i \in C$$

  - Methodology is general, so it can be used for any allocation strategy for mapping the return forecasts to the portfolio weights

- Portfolio return for period $t_{j,\text{end}} + 1 \Rightarrow$

$$r^{\text{Port}}_{t_{j,\text{end}}+1}\Big(\big\{\hat{f}\big(\mathbf{x}_{i,t_{j,\text{end}}};W_j\big)\big\}_{i\in C}\Big) = \sum_{i\in C} w_{i,t_{j,\text{end}}+1}\Big(\big\{\hat{f}\big(\mathbf{x}_{i,t_{j,\text{end}}};W_j\big)\big\}_{i\in C}\Big)\, r_{i,t_{j,\text{end}}+1}$$

- Define the Shapley value corresponding to the **portfolio return** for period $t_{j,\text{end}} + 1 \Rightarrow$

$$\phi_p^{\text{PR}}\left(\left\{\boldsymbol{x}_{i,t_{j,\text{end}}}\right\}_{i\in C}; W_j\right) = \frac{1}{P!}\sum_{\mathcal{O}\in\pi(P)}\left[\xi^{\text{PR}}_{\text{Pre}_p(\mathcal{O})\cup\{p\}}\left(\left\{\boldsymbol{x}_{i,t_{j,\text{end}}}\right\}_{i\in C}; W_j\right) - \xi^{\text{PR}}_{\text{Pre}_p(\mathcal{O})}\left(\left\{\boldsymbol{x}_{i,t_{j,\text{end}}}\right\}_{i\in C}; W_j\right)\right] \quad \text{for } p \in S$$

- Use the sampling algorithm to make a random draw $m \Rightarrow$

$$\hat{\theta}^{\text{PR}}_{p,m}\left(\left\{\boldsymbol{x}_{i,t_{j,\text{end}}}\right\}_{i\in C}; W_j\right) = \underbrace{\sum_{i\in C}\left[w_{i,t_{j,\text{end}}+1}\left(\left\{\hat{r}_{i,t_{j,\text{end}}+1,m,p}\left(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j\right)\right\}_{i\in C}\right) r_{i,t_{j,\text{end}}+1}\right]}_{\hat{\xi}^{\text{PR}}_{\text{Pre}_p(\mathcal{O})\cup\{p\}}(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i\in C}; W_j)} - \underbrace{\sum_{i\in C}\left[w_{i,t_{j,\text{end}}+1}\left(\left\{\hat{r}_{i,t_{j,\text{end}}+1,m,\backslash p}\left(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j\right)\right\}_{i\in C}\right) r_{i,t_{j,\text{end}}+1}\right]}_{\hat{\xi}^{\text{PR}}_{\text{Pre}_p(\mathcal{O})}(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i\in C}; W_j)}$$

- Sampling approximation $\Rightarrow$

$$\hat{\phi}_p^{\text{PR}}\left(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i \in C}\,;W_j\right) = \frac{1}{2M}\sum_{m=1}^{2M}\hat{\theta}_{p,m}^{\text{PR}}\left(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i \in C}\,;W_j\right) \quad \text{for } p \in S$$

- Need to decide on the baseline portfolio return ($r_{t_{j,\text{end}}+1}^{\text{Base}}$) corresponding to the empty coalition set
    - Sensible to ask, "If I had an empty set of predictors—and so no predictor information—how would I form a portfolio?"
    - Relevant baseline depends on the context (eg, CRSP value-weighted market portfolio for a portfolio that broadly invests in equities)
- Local accuracy holds for $\hat{\phi}_p^{\text{PR}}\left(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i \in C}\,;W_j\right) \Rightarrow$

$$\sum_{p \in S}\hat{\phi}_p^{\text{PR}}\left(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i \in C}\,;W_j\right) = r_{t_{j,\text{end}}+1}^{\text{Port}} - r_{t_{j,\text{end}}+1}^{\text{Base}}$$

- To derive the SPPC, we need to take into account the entire series of out-of-sample return forecasts and corresponding portfolio returns over the forecast evaluation period
- Sample of panel data spans $T$ periods
- Initial in-sample period ends in $T_{\text{in}}$
- Generate return forecasts for $T_{\text{in}+1}$ through $T$
  - $D = T - T_{\text{in}}$ sets of return forecasts (ie, length of forecast evaluation period)
- Index set of training windows for the sequence of fitted prediction models $\Rightarrow$ $W = \{1, \ldots, D\}$
  - $t_{j,\text{end}}$ corresponds to $T_{\text{in}}, T_{\text{in}} + 1, \ldots, T - 1$ for $j = 1, 2, \ldots, D$

- $\mathcal{M}(\cdot) \Rightarrow$ performance metric function
    - Methodology is general, so it accommodates any performance metric
- Define the Shapley value corresponding to the **portfolio performance metric** $\Rightarrow$

$$\phi_p^{\text{Perf}}\left(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i\in C}; W, \mathcal{M}\right) = \frac{1}{P!}\sum_{\mathcal{O}\in\pi(P)}\left[\xi^{\text{Perf}}_{\text{Pre}_p(\mathcal{O})\cup\{p\}}\left(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i\in C}; W, \mathcal{M}\right) - \xi^{\text{Perf}}_{\text{Pre}_p(\mathcal{O})}\left(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i\in C}; W, \mathcal{M}\right)\right] \quad \text{for } p \in S$$

- Intuition $\Rightarrow$ wrap a function around the sequence of portfolio returns

- Use the sampling algorithm to make a random draw $m \Rightarrow$

$$\hat{\theta}^{\text{Perf}}_{p,m}\left(\left\{\boldsymbol{x}_{i,t_{j,\text{end}}}\right\}_{i\in C}; W, \mathcal{M}\right) =$$

$$\underbrace{\mathcal{M}\left(\left\{\sum_{i\in C}\left[w_{i,t_{j,\text{end}}+1}\left(\left\{\hat{r}_{i,t_{j,\text{end}}+1,m,p}\left(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j\right)\right\}_{i\in C}\right) r_{i,t_{j,\text{end}}+1}\right]\right\}_{j\in W}\right)}_{\hat{\xi}^{\text{Perf}}_{\text{Prep}(\mathcal{O})\cup\{p\}}\left(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i\in C}; W, \mathcal{M}\right)} -$$

$$\underbrace{\mathcal{M}\left(\left\{\sum_{i\in C}\left[w_{i,t_{j,\text{end}}+1}\left(\left\{\hat{r}_{i,t_{j,\text{end}}+1,m,\backslash p}\left(\boldsymbol{x}_{i,t_{j,\text{end}}}; W_j\right)\right\}_{i\in C}\right) r_{i,t_{j,\text{end}}+1}\right]\right\}_{j\in W}\right)}_{\hat{\xi}^{\text{Perf}}_{\text{Prep}(\mathcal{O})}\left(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i\in C}; W, \mathcal{M}\right)}$$

- Sampling approximation $\Rightarrow$

$$\underbrace{\hat{\phi}^{\text{Perf}}_{p}\left(\left\{\boldsymbol{x}_{i,t_{j,\text{end}}}\right\}_{i\in C}; W, \mathcal{M}\right)}_{\text{SPPC}_p} = \frac{1}{2M}\sum_{m=1}^{2M}\hat{\theta}^{\text{Perf}}_{p,m}\left(\left\{\boldsymbol{x}_{i,t_{j,\text{end}}}\right\}_{i\in C}; W, \mathcal{M}\right)$$

for $p \in S$

- $\text{SPPC}_p$ allows a researcher to measure how an individual predictor contributes to portfolio performance
  - Can also be computed for a group of predictors (as in the empirical application)
- Local accuracy holds for $\hat{\phi}_p^{\text{Perf}}(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i\in C}; W, \mathcal{M}) \Rightarrow$

$$\sum_{p\in S} \underbrace{\hat{\phi}_p^{\text{Perf}}(\{\boldsymbol{x}_{i,t_{j,\text{end}}}\}_{i\in C}; W, \mathcal{M})}_{\text{SPPC}_p} = \underbrace{\mathcal{M}(\{r^{\text{Port}}_{t_{j,\text{end}}+1}\}_{j\in W})}_{\substack{\text{portfolio performance}\\ \text{over the forecast}\\ \text{evaluation period}}} - \mathcal{M}(\{r^{\text{Base}}_{t_{j,\text{end}}+1}\}_{j\in W})$$

- Sum of the $\text{SPPC}_p$ measures provides an exact decomposition of portfolio performance (relative to the baseline portfolio) in terms of the contributions made by each of the $P$ predictors

- Develop a computationally efficient algorithm for computing the $SPPC_p$
  - In the process of creating a Python package
- 2 dimensions along which to limit computational cost
  1. Number of randomly drawn ordered permutations ($M$)
  2. Proportion of training sample observations used to integrate out the predictors not in a coalition
- Empirical application $\Rightarrow$ $M = 50$ and 1% of the training sample observations
  - Results are similar when we use 10% of the training sample observations
- Less computationally costly alternatives $\Rightarrow$ "permutation" / "scramble" methods (eg, Breiman 2001; Jensen et al 2024)
  - But the alternatives generally don't have the attractive properties of Shapley values (including local accuracy, so an exact decomposition isn't furnished)

- 207 firm characteristics from Chen & Zimmermann (2022)
  - Available at the Open Source Asset Pricing website
  - Transform each characteristic each month by cross-sectionally ranking the characteristics and mapping the ranks into the $[-1, 1]$ interval (Freyberger, Neuhierl & Weber 2020; Gu, Kelly & Xiu 2020)
- Monthly firm-level stock return data from CRSP
  - All firms listed on the NYSE/AMEX/NASDAQ with a market value on CRSP at the end of the previous month and a non-missing value for common equity in the firm's annual financial statement
  - Compute the excess return for each stock in a given month using the CRSP risk-free return
- Total sample $\Rightarrow$ 1960:01–2021:12 (744 months)

Predictor groups

| | |
|---|---|
| Earnings (9) | Profitability (14) |
| Earnings forecast (10) | R&D (8) |
| Financing (10) | Reversal (7) |
| Financing alt (7) | Risk (12) |
| Investment (14) | Risk alt (12) |
| Investment alt (12) | Sales (10) |
| Lead lag (9) | Seasonal momentum (10) |
| Liquidity (11) | Valuation (12) |
| Momentum (11) | Valuation ratio (11) |
| Ownership (11) | Volume (6) |

- Construct a zero-investment long-short portfolio that goes long (short) stocks that are predicted to have returns in the highest (lowest) quintile
  - Drop stocks with market cap below the NYSE 20th percentile
  - Long/short legs are value weighted
  - Scale the weights in the long (short) leg to sum to 1 ($-1$)
- Initial in-sample period $\Rightarrow$ 1960:01–1972:12 (156 months)
- Out-of-sample period $\Rightarrow$ 1973:01–2021:12 (588 months)
- Generate out-of-sample classification forecasts using the powerful XGBoost algorithm (Chen & Guestrin 2016)
  - Retrain the prediction model each month as additional data become available using a rolling window
  - Tune the hyperparameters each month using a walk-forward procedure (based on the Sharpe ratio) that respects the time-series dimension of the panel data

Portfolio performance (1973:01–2021:12)

| Model | Ann mean | Ann vol | Ann Sharpe ratio | Ann FF6 alpha | Ann Q5 alpha |
|---|---|---|---|---|---|
| XGBoost | 22.58% | 12.53% | 1.80 | 19.44%*** | 16.29%*** |
| Market | 7.48% | 15.85% | 0.47 | — | — |

FF6 $\Rightarrow$ Fama & French (2015) 5-factor model + momentum

Q5 $\Rightarrow$ Hou et al (2021) augmented q-factor model

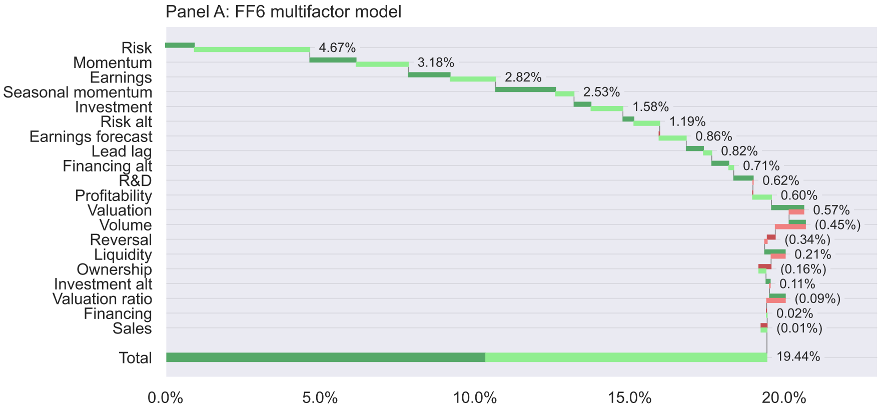
Portfolio performance contributions based on $\text{SPPC}_p$

| Predictor group | Ann mean | Ann vol | Ann Sharpe ratio | Ann FF6 alpha | Ann Q5 alpha |
|---|---|---|---|---|---|
| Baseline | 7.48% | 15.85% | 0.47 | 0% | 0% |
| Risk | **5.01** | −0.09 | **0.38** | **4.67** | **4.59** |
| Earnings | **2.67** | −0.21 | **0.20** | **2.82** | **2.18** |
| Seas momentum | **1.72** | **−0.91** | **0.17** | **2.53** | **2.65** |
| Momentum | **4.38** | 2.24 | **0.14** | **3.18** | **1.85** |
| Lead lag | 0.90 | −0.75 | 0.09 | 0.82 | 0.40 |
| Investment | 0.96 | −0.45 | 0.09 | 1.58 | 0.75 |
| Valuation ratio | 0.00 | **−1.01** | 0.08 | −0.09 | 0.16 |
| Risk alt | 0.42 | −0.66 | 0.08 | 1.19 | 0.71 |
| Profitability | 1.00 | 0.08 | 0.05 | 0.60 | −0.88 |
| Earnings forecast | 0.92 | 0.05 | 0.06 | 0.86 | 1.01 |

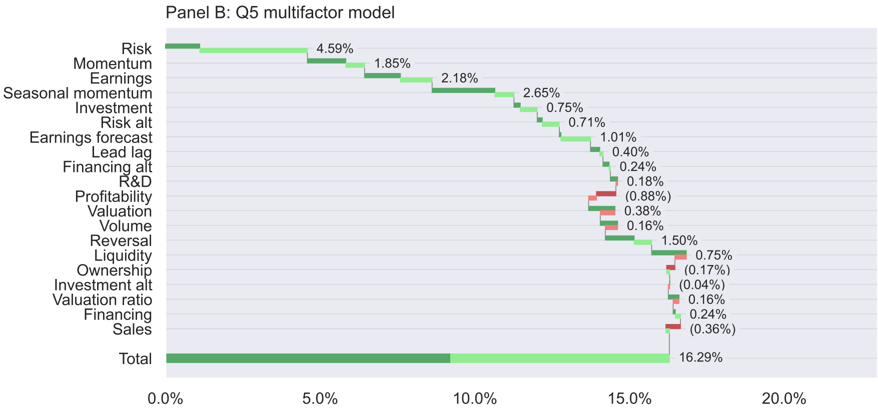
## Portfolio performance contributions based on $SPPC_p$ (cont'd)

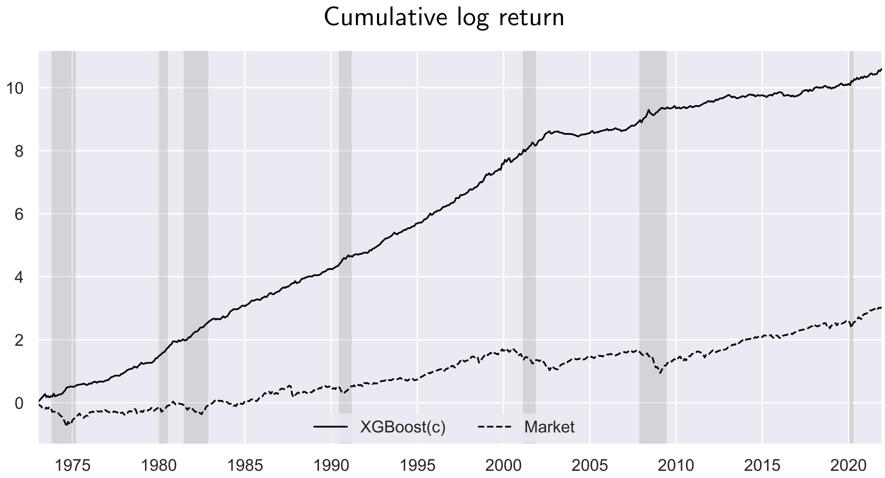
| Predictor group | Ann mean | Ann vol | Ann Sharpe ratio | Ann FF6 alpha | Ann Q5 alpha |
|---|---|---|---|---|---|
| Valuation | 0.24 | **−0.78** | 0.06 | 0.57 | 0.38 |
| Financing | −0.04 | 0.07 | 0.00 | 0.02 | 0.24 |
| Financing alt | 0.05 | 0.10 | 0.00 | 0.71 | 0.24 |
| Volume | −0.87 | **−1.19** | 0.00 | −0.45 | 0.16 |
| Liquidity | 0.28 | 0.98 | 0.04 | 0.21 | 0.75 |
| Investment alt | −0.04 | 0.06 | 0.00 | 0.11 | −0.04 |
| R&D | −0.04 | 0.16 | −0.01 | 0.62 | 0.18 |
| Reversal | −1.50 | −0.20 | −0.04 | −0.34 | 1.50 |
| Sales | −0.35 | −0.19 | −0.02 | −0.01 | −0.36 |
| Ownership | −0.61 | −0.60 | −0.03 | −0.16 | −0.17 |
| Total | 22.58% | 12.53% | 1.80 | 19.44% | 16.29% |

# Alpha long- and short-leg contributions

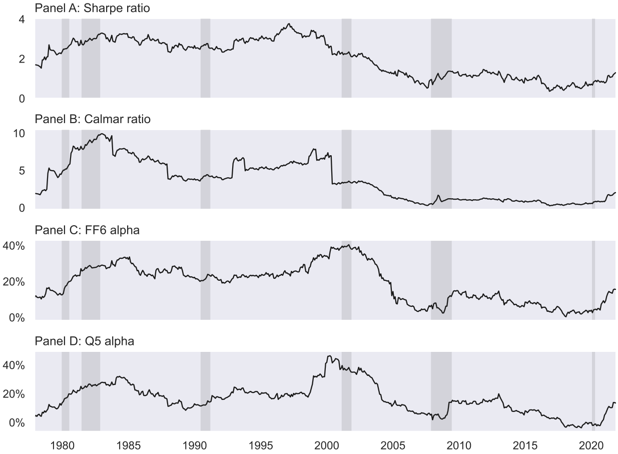


# Alpha long- and short-leg contributions (cont'd)

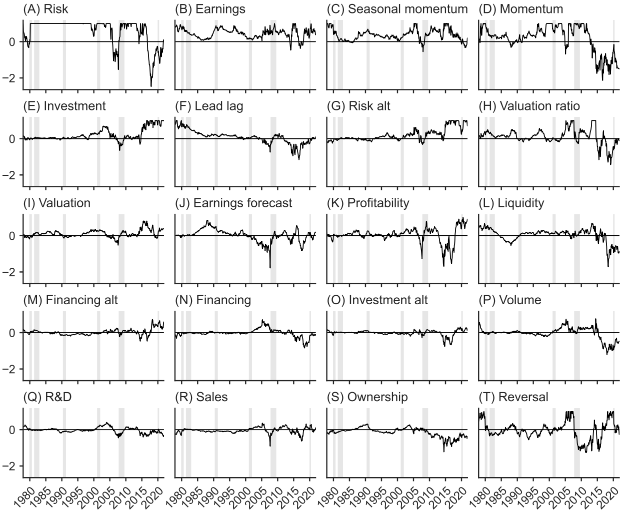



Cumulative log return
10
8
6
4
2
0
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
XGBoost(c)
Market

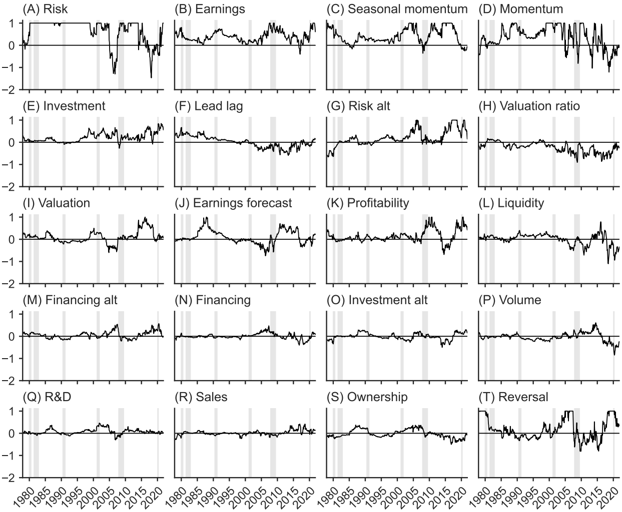
# Portfolio performance for 60-month rolling windows



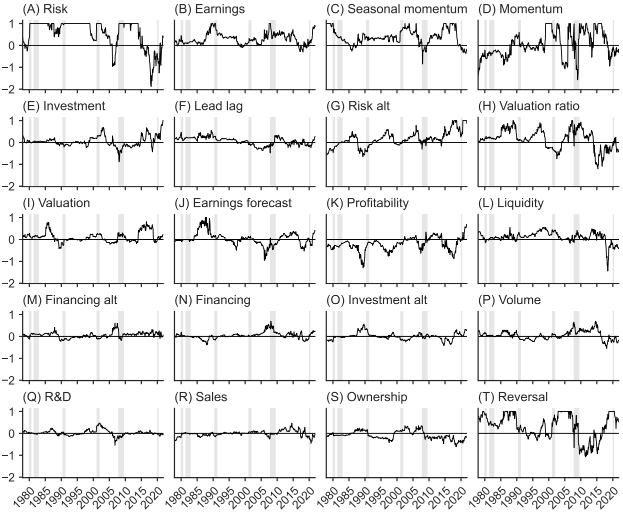
# Sharpe ratio contributions for 60-month rolling windows



# FF6 alpha contributions for 60-month rolling windows



# Q5 alpha contributions for 60-month rolling windows



## Alpha contributions using permutation method

| | Ann FF6 alpha | | Ann Q5 alpha | |
|---|---|---|---|---|
| Predictor group | $SPPC_p$ | Permutation | $SPPC_p$ | Permutation |
| Risk | 4.67 | 7.74 | 4.59 | 7.31 |
| Earnings | 2.82 | 4.82 | 2.18 | 4.00 |
| Seasonal Momentum | 2.53 | 1.93 | 2.65 | 2.69 |
| Momentum | 3.18 | 6.08 | 1.85 | 5.32 |
| Lead lag | 0.82 | 1.83 | 0.40 | 0.94 |
| Investment | 1.58 | 2.00 | 0.75 | 0.46 |
| Valuation Ratio | −0.09 | 1.41 | 0.16 | 1.37 |
| Risk alt | 1.19 | 1.94 | 0.71 | 1.55 |
| Profitability | 0.60 | 2.28 | −0.88 | 0.46 |
| Earnings Forecast | 0.86 | 1.15 | 1.01 | 0.60 |

# Alpha contributions using permutation method (cont'd)

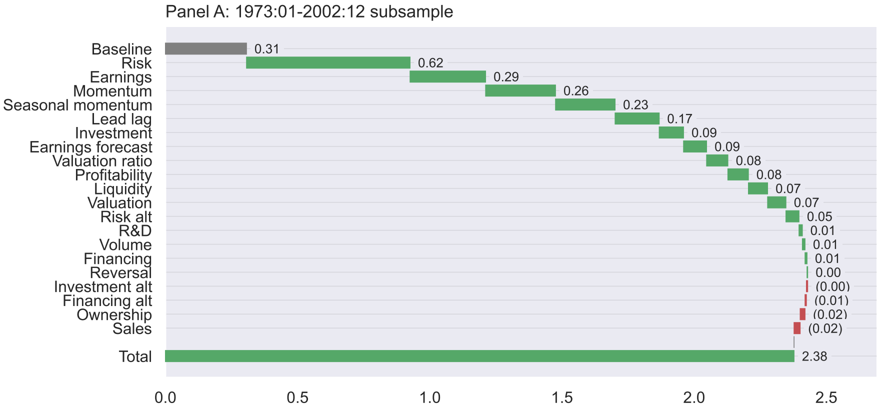
| | Ann FF6 alpha | | Ann Q5 alpha | |
|---|---|---|---|---|
| Predictor group | $SPPC_p$ | Permutation | $SPPC_p$ | Permutation |
| Valuation | 0.57 | 1.65 | 0.38 | 0.73 |
| Financing | 0.02 | 1.67 | 0.24 | 1.64 |
| Financing alt | 0.71 | 0.59 | 0.24 | −0.36 |
| Volume | −0.45 | −1.08 | 0.16 | −0.80 |
| Liquidity | 0.21 | 0.42 | 0.75 | 1.35 |
| Investment alt | 0.11 | 0.97 | −0.04 | 1.28 |
| R&D | 0.62 | −0.19 | 0.18 | −0.32 |
| Reversal | −0.34 | 0.12 | 1.50 | 0.94 |
| Sales | −0.01 | 0.56 | −0.36 | −0.90 |
| Ownership | −0.16 | 0.31 | −0.17 | 0.23 |
| Total | 19.44% | 36.20% | 16.29% | 28.49% |

- Employ the SPPC in real time to improve portfolio performance
  - Becomes a **monitoring** as well as a model interpretation tool
- Based on a look-back period, don't use information for predictors that contribute negatively on average to portfolio performance
  - Integrate out the "bad" predictors in the most recent fitted model (instead of fitting a different model that excludes the predictors)
- Intel Xeon Platinum 8260 with AVX-512 enabled $\Rightarrow \approx 7$ minutes for monthly update using the algorithm
- Results will be available soon in the next revision of the paper

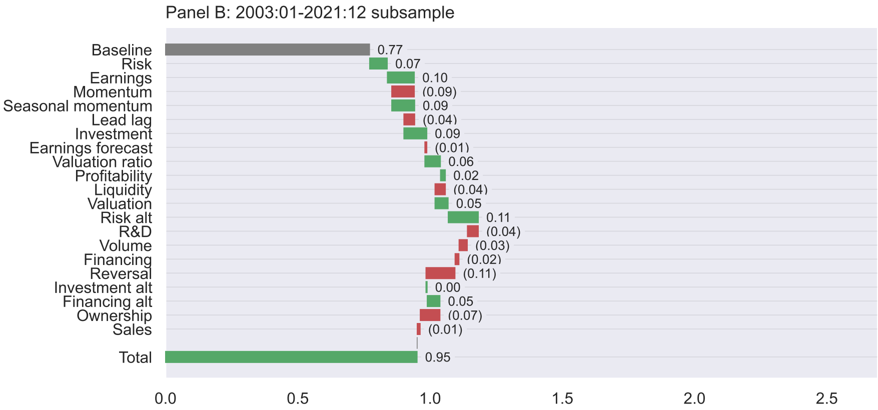
## Portfolio performance for subsamples

| Model | Ann mean | Ann vol | Ann Sharpe ratio | Ann FF6 alpha | Ann Q5 alpha |
|---|---|---|---|---|---|
| **Panel A: 1973:01–2002:12 subsample** | | | | | |
| XGBoost | 29.74% | 12.51% | 2.38 | 24.57%*** | 22.25%*** |
| Market | 5.08% | 16.54% | 0.31 | — | — |
| **Panel B: 2003:01–2021:12 subsample** | | | | | |
| XGBoost | 11.29% | 11.86% | 0.95 | 9.67%*** | 8.53%*** |
| Market | 11.28% | 14.62% | 0.77 | — | — |

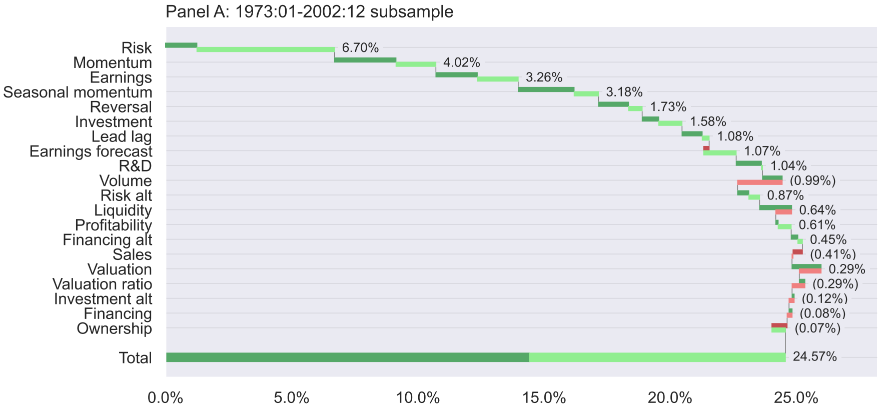
# Sharpe ratio contributions for subsamples



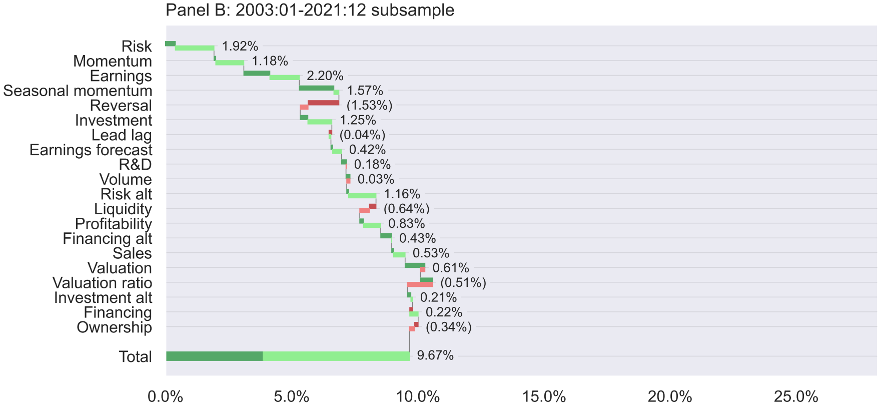
# Sharpe ratio contributions for subsamples (cont'd)



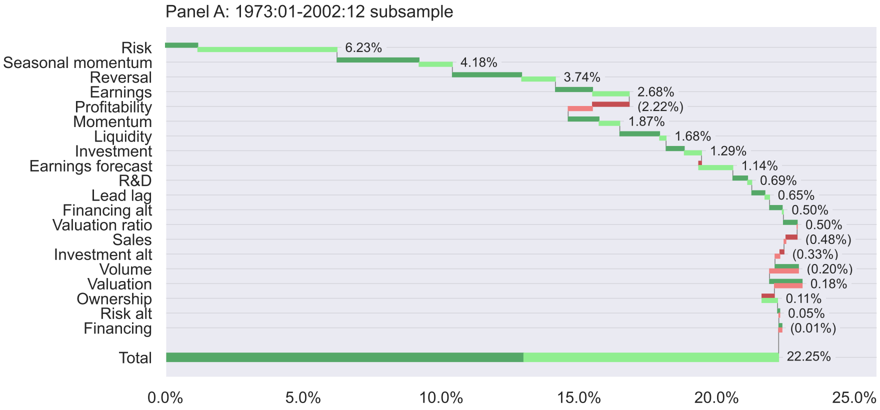
# FF6 alpha long- and short-leg contributions for subsamples



# FF6 alpha long- and short-leg contributions for subsamples (cont'd)



# Q5 alpha long- and short-leg contributions for subsamples



# Q5 alpha long- and short-leg contributions for subsamples (cont'd)