Lecture 1: Measuring Fund Performance

Institutional Investors

Calvin J. Chiou

2026-09-16

Table of contents

1 Overview

How do we know whether a fund manager adds value? This lecture introduces the canonical **returns-based** framework for fund performance evaluation, following the survey by Wermers (2011). The central takeaway from Wermers: the **Carhart (1997) four-factor model is the baseline model** for evaluating actively managed equity portfolios.

We will:

1. Motivate the need for multi-factor performance models
2. Build from CAPM to Fama-French 3-factor to Carhart 4-factor, step by step
3. Download price data for 10 popular US ETFs using `tidyquant`
4. Obtain factor returns directly from **Kenneth French's Data Library** using the `frenchdata` package
5. Estimate and interpret Carhart alphas and factor loadings

---

2 1. Motivation: Why Not Just Compare Raw Returns?

A fund that returned 15% last year sounds impressive—until you learn it was a small-cap value fund during a year when small-cap value stocks returned 18%. The fund actually *destroyed* value relative to the passive alternative.

Wermers (2011) states the core principle clearly: "Managers should be rewarded for bets not easily replicated by uninformed investors." In other words, we need to control for the passive risk factors that investors can access cheaply, and only then can we attribute the residual—*alpha*—to manager skill.

2.1 The Evolution of Baseline Models

Table 1: Evolution of performance benchmarks

Year	Model	Factors Controlled
Jensen (1968)	CAPM (1-factor)	Market (systematic) risk
Fama and French (1993)	Fama-French 3-factor	Market + Size (SMB) + Value (HML)
Carhart (1997)	4-factor	FF3 + Momentum (UMD)

Wermers (2011) (Section 3.1) notes that a regression of a diversified long-only portfolio on the Carhart model typically yields $R^2 > 90\%$, meaning nearly all return variation is explained by these four systematic factors. This is precisely why it serves as the *baseline*: it leaves little room for data mining.

---

3 2. Setup

```
library(tidyverse)
library(tidyquant)
library(frenchdata)   # Kenneth French's Data Library
library(broom)
library(gt)
library(scales)
```

---

4 3. The Ten US ETFs

We use ten ETFs spanning broad market, style (size/value), sector, and active strategies. This cross-section lets students observe how factor loadings vary systematically across fund types—a key pedagogical goal.

Table 2: Ten US ETFs used in this lecture

Ticker	Name	Category	Expected Factor Tilt
SPY	SPDR S&P 500	Broad market	1, others 0
QQQ	Invesco Nasdaq-100	Tech/growth	Negative HML
IWM	iShares Russell 2000	Small cap	Positive SMB
VTV	Vanguard Value	Large-cap value	Positive HML
VUG	Vanguard Growth	Large-cap growth	Negative HML
ARKK	ARK Innovation	Active growth	High , negative HML
XLE	Energy Select SPDR	Energy sector	Positive HML (value)
XLF	Financial Select SPDR	Financial sector	High
VNQ	Vanguard Real Estate	REITs	Positive HML
IWD	iShares Russell 1000 Value	Large-cap value	Positive HML

```r
etf_tickers <- c(
  "SPY",   # S&P 500
  "QQQ",   # Nasdaq-100
  "IWM",   # Russell 2000
  "VTV",   # Vanguard Value
  "VUG",   # Vanguard Growth
  "ARKK",  # ARK Innovation (active)
  "XLE",   # Energy sector
  "XLF",   # Financials sector
  "VNQ",   # Real estate
  "IWD"    # Russell 1000 Value
)

etf_labels <- tibble(
  symbol   = etf_tickers,
  name     = c("S&P 500", "Nasdaq-100", "Russell 2000",
               "Vanguard Value", "Vanguard Growth", "ARK Innovation",
               "Energy Sector", "Financials Sector",
               "Real Estate (REIT)", "Russell 1000 Value"),
  category = c("Broad market", "Tech/Growth", "Small cap",
               "Large value", "Large growth", "Active growth",
               "Sector", "Sector", "REIT", "Large value")
)
```

4.1 Download ETF Price Data

```r
start_date <- "2015-01-01"
end_date   <- "2024-12-31"

prices_raw <- tq_get(
  etf_tickers,
  from = start_date,
  to   = end_date,
  get  = "stock.prices"
)

# Quick check
prices_raw |>
  select(symbol, date, adjusted) |>
  slice_head(n = 5)
```

```
# A tibble: 5 x 3
  symbol date       adjusted
  <chr>  <date>        <dbl>
1 SPY    2015-01-02     170.
2 SPY    2015-01-05     167.
3 SPY    2015-01-06     165.
4 SPY    2015-01-07     167.
5 SPY    2015-01-08     170.
```

4.2 Compute Monthly Returns

The Carhart model is estimated on **monthly** returns—the same frequency as the French factor data.

```r
returns_monthly <- prices_raw |>
  group_by(symbol) |>
  tq_transmute(
    select     = adjusted,
    mutate_fun = periodReturn,
    period     = "monthly",
    type       = "arithmetic",   # simple returns to match French data convention
    col_rename = "ret"
  ) |>
```

```
ungroup() |>
left_join(etf_labels, by = "symbol")
```

---

5 4. Kenneth French's Factor Data

Kenneth French maintains a publicly accessible data library at mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. The `frenchdata` package downloads and parses these files directly.

5.1 The Four Factors

The Carhart model uses four factors:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i \cdot RMRF_t + s_i \cdot SMB_t + h_i \cdot HML_t + u_i \cdot UMD_t + \varepsilon_{i,t} \tag{1}$$

Table 3: Carhart four-factor model components

Factor	Full name	Economic rationale
$RMRF$	Market excess return	Compensation for bearing systematic risk
SMB	Small-minus-big	Size premium: small stocks earn higher returns on average (Fama and French 1993)
HML	High-minus-low	Value premium: high book-to-market (cheap) stocks outperform growth stocks (Fama and French 1993)
UMD	Up-minus-down (momentum)	Past winners continue to outperform past losers (Jegadeesh and Titman 1993; Carhart 1997)

5.2 Download and Prepare Factor Data

```r
# Fama-French 3 factors: Mkt-RF, SMB, HML, RF (monthly)
ff3_raw <- download_french_data("Fama/French 3 Factors")

# Momentum factor: UMD (monthly)
mom_raw <- download_french_data("Momentum Factor (Mom)")

# Extract monthly tables and convert from % to decimal
ff3_monthly <- ff3_raw$subsets$data[[1]] |>
  mutate(
    date   = lubridate::ym(date),          # "YYYYMM" → date
    across(c(`Mkt-RF`, SMB, HML, RF), ~ . / 100)
  ) |>
  rename(RMRF = `Mkt-RF`) |>
  filter(date >= as.Date(start_date), date <= as.Date(end_date))

mom_monthly <- mom_raw$subsets$data[[1]] |>
  mutate(
    date = lubridate::ym(date),
    Mom  = Mom / 100
  ) |>
  rename(UMD = Mom) |>
  filter(date >= as.Date(start_date), date <= as.Date(end_date))

# Combine into a single factor tibble
factors <- ff3_monthly |>
  left_join(mom_monthly, by = "date")

factors |> slice_head(n = 6)
```

```
# A tibble: 6 x 6
  date          RMRF     SMB     HML    RF     UMD
  <date>       <dbl>   <dbl>   <dbl> <dbl>   <dbl>
1 2015-01-01 -0.031  -0.0061 -0.0348     0  0.0372
2 2015-02-01  0.0613  0.0064 -0.0175     0 -0.0287
3 2015-03-01 -0.0111  0.0306 -0.0038     0  0.0272
4 2015-04-01  0.0059 -0.0298  0.0181     0 -0.0726
5 2015-05-01  0.0137  0.0096 -0.011      0  0.0576
6 2015-06-01 -0.0152  0.0295 -0.008      0  0.0301
```

5.3 Align Dates and Compute Excess Returns

French data uses end-of-month dates; `tq_transmute()` returns the last trading day of each month. We align on year-month.

```
# Standardize ETF returns to year-month for merging
returns_ym <- returns_monthly |>
  mutate(ym = lubridate::floor_date(date, "month")) |>
  select(symbol, name, category, ym, ret)

factors_ym <- factors |>
  mutate(ym = lubridate::floor_date(date, "month")) |>
  select(ym, RMRF, SMB, HML, UMD, RF)

# Merge and compute excess return
data_merged <- returns_ym |>
  left_join(factors_ym, by = "ym") |>
  mutate(excess_ret = ret - RF) |>
  filter(!is.na(RMRF))   # drop months outside factor data range
```

---

6 5. Cumulative Return Plot

```
data_merged |>
  group_by(symbol, name) |>
  arrange(ym) |>
  mutate(cum_ret = cumprod(1 + ret) - 1) |>
  ggplot(aes(x = ym, y = cum_ret, color = name)) +
  geom_line(linewidth = 0.6) +
  labs(
    title = "Cumulative Returns",
    x = NULL, y = "Cumulative Return",
    color = NULL
  ) +
  scale_y_continuous(labels = percent_format()) +
  theme_minimal() +
  theme(legend.position = "bottom",
        legend.text = element_text(size = 7)) +
  guides(color = guide_legend(nrow = 3))
```

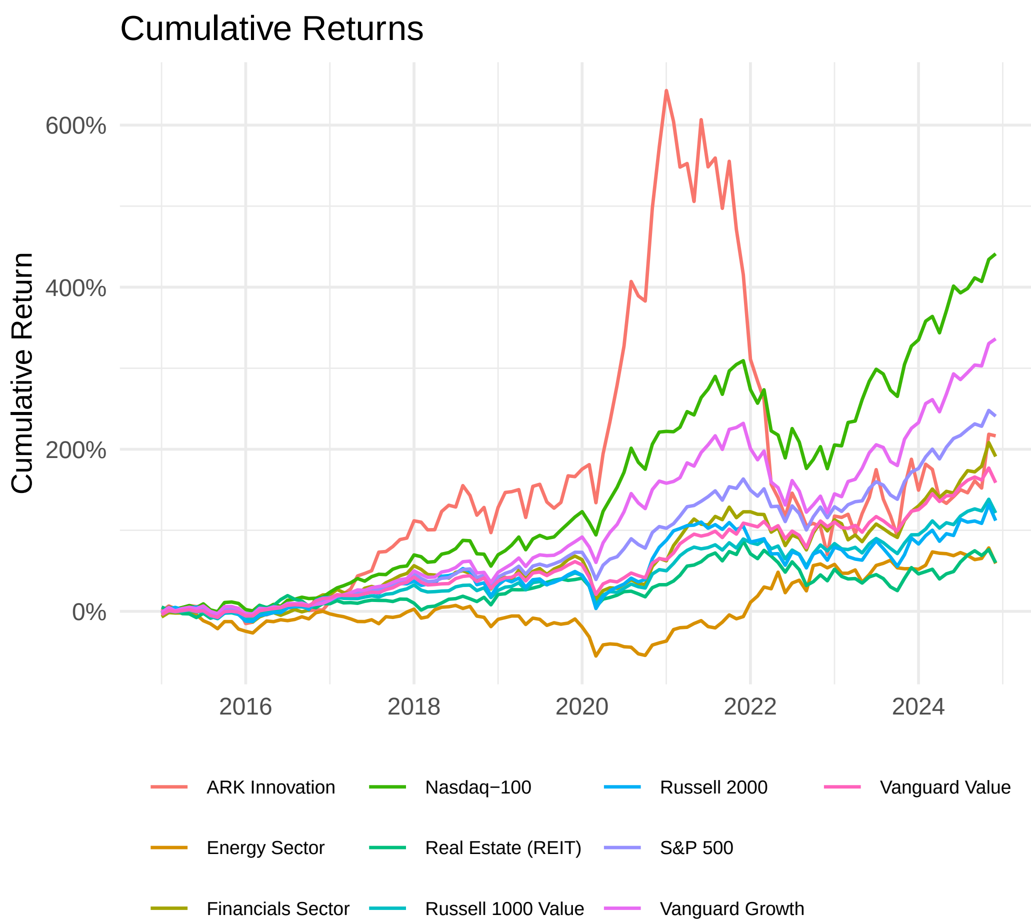

Figure 1: Cumulative simple returns of 10 US ETFs (2015–2024)

---

7 6. Step 1 — CAPM / Jensen's Alpha (Single Factor)

Before moving to Carhart, we start with the **CAPM** as a baseline. Jensen (1968) showed that fund performance can be measured as the regression intercept α from:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i \cdot RMRF_t + \varepsilon_{i,t} \tag{2}$$

$\alpha > 0$ means the fund earned more than its CAPM-implied expected return.

```r
capm_results <- data_merged |>
  group_by(symbol, name) |>
  do(tidy(lm(excess_ret ~ RMRF, data = .))) |>
  select(symbol, name, term, estimate, std.error, statistic, p.value) |>
  filter(term %in% c("(Intercept)", "RMRF")) |>
  mutate(term = recode(term, "(Intercept)" = "alpha", "RMRF" = "beta_mkt")) |>
  pivot_wider(
    names_from  = term,
    values_from = c(estimate, std.error, statistic, p.value),
    names_glue  = "{term}_{.value}"
  ) |>
  mutate(alpha_ann = alpha_estimate * 12) |>
  ungroup()

capm_results |>
  arrange(desc(alpha_ann)) |>
  select(name, alpha_ann, beta_mkt_estimate, alpha_statistic, alpha_p.value) |>
  gt() |>
  fmt_number(columns = c(alpha_ann, beta_mkt_estimate, alpha_statistic), decimals = 3) |>
  fmt(columns = alpha_p.value, fns = \(x) sprintf("%.3f", x)) |>
  cols_label(
    name              = "ETF",
    alpha_ann         = "Alpha (Ann.)",
    beta_mkt_estimate = "Beta",
    alpha_statistic   = "t( )",
    alpha_p.value     = "p( )"
  ) |>
  tab_header(
    title    = "CAPM / Jensen's Alpha",
    subtitle = "Monthly excess returns regressed on market factor only"
  )
```

Limitation of CAPM alpha: If IWM (small cap) has $\alpha > 0$ in the CAPM, it might simply be capturing the *size premium* that any investor could access by buying small stocks passively—not manager skill. This is why we need additional factors.

CAPM / Jensen's Alpha

Monthly excess returns regressed on market factor only

ETF	Alpha (Ann.)	Beta	t( )	p( )
Nasdaq-100	0.045	1.065	1.847	0.067
Vanguard Growth	0.023	1.058	1.200	0.232
S&P 500	0.005	0.957	1.019	0.311
Vanguard Value	-0.013	0.870	-0.668	0.505
Financials Sector	-0.015	1.065	-0.451	0.653
Russell 1000 Value	-0.034	0.928	-1.886	0.062
ARK Innovation	-0.042	1.740	-0.553	0.581
Real Estate (REIT)	-0.054	0.860	-1.388	0.168
Russell 2000	-0.058	1.173	-1.962	0.052
Energy Sector	-0.067	1.200	-0.892	0.374

8 7. Step 2 — Fama-French 3-Factor Model

Fama and French (1993) document that two additional risk factors—size (SMB) and value (HML)—explain cross-sectional variation in average returns beyond the market. Adding these factors purges the CAPM alpha of style-related biases:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i \cdot RMRF_t + s_i \cdot SMB_t + h_i \cdot HML_t + \varepsilon_{i,t} \tag{3}$$

```r
ff3_results <- data_merged |>
  group_by(symbol, name) |>
  do(tidy(lm(excess_ret ~ RMRF + SMB + HML, data = .))) |>
  filter(term %in% c("(Intercept)", "RMRF", "SMB", "HML")) |>
  mutate(term = recode(term,
    "(Intercept)" = "alpha", "RMRF" = "b_mkt",
    "SMB" = "b_smb", "HML" = "b_hml"
  )) |>
  pivot_wider(
    names_from  = term,
    values_from = c(estimate, std.error, statistic, p.value),
    names_glue  = "{term}_{.value}"
  ) |>
  mutate(alpha_ann = alpha_estimate * 12) |>
  ungroup()

ff3_results |>
```

Fama-French 3-Factor Model

Controlling for market, size, and value

ETF	Alpha (Ann.)	t( )	(MKT)	(SMB)	(HML)
Nasdaq-100	0.029	1.772	1.103	-0.150	-0.397
Vanguard Growth	0.009	0.846	1.091	-0.132	-0.341
Financials Sector	0.001	0.028	1.045	0.021	0.599
S&P 500	0.000	0.150	0.982	-0.142	0.012
Vanguard Value	-0.008	-0.677	0.880	-0.108	0.357
ARK Innovation	-0.018	-0.309	1.540	1.236	-0.867
Russell 2000	-0.021	-2.523	1.012	0.859	0.226
Russell 1000 Value	-0.026	-2.448	0.922	-0.015	0.345
Energy Sector	-0.032	-0.571	1.142	0.152	1.162
Real Estate (REIT)	-0.048	-1.228	0.836	0.127	0.046

```r
  arrange(desc(alpha_ann)) |>
  select(name, alpha_ann, alpha_statistic, b_mkt_estimate,
         b_smb_estimate, b_hml_estimate) |>
  gt() |>
  fmt_number(columns = where(is.numeric), decimals = 3) |>
  cols_label(
    name              = "ETF",
    alpha_ann         = "Alpha (Ann.)",
    alpha_statistic   = "t( )",
    b_mkt_estimate    = " (MKT)",
    b_smb_estimate    = " (SMB)",
    b_hml_estimate    = " (HML)"
  ) |>
  tab_header(
    title    = "Fama-French 3-Factor Model",
    subtitle = "Controlling for market, size, and value"
  )
```

Notice how alpha estimates change relative to CAPM—this is the style-correction at work. IWM's alpha shrinks once we control for SMB; VTV's alpha changes once we control for HML.

---

9 8. Step 3 — Carhart 4-Factor Model (Baseline)

Carhart (1997) adds the **momentum factor** (UMD) to the Fama-French model. Stocks that have outperformed over the prior 12 months tend to continue outperforming over the next few months (Jegadeesh and Titman 1993). If a fund merely holds past winners, its FF3 alpha is overstated—Carhart strips this out.

$$\boxed{R_{i,t} - R_{f,t} = \alpha_i + \beta_i \cdot RMRF_t + s_i \cdot SMB_t + h_i \cdot HML_t + u_i \cdot UMD_t + \varepsilon_{i,t}} \quad (4)$$

As Wermers (2011) (p. 542) states: *“The returns-based model most widely used among academics in analyzing equity managers is the four-factor model of Carhart (1997).”*

```
carhart_results <- data_merged |>
  group_by(symbol, name) |>
  do({
    fit <- lm(excess_ret ~ RMRF + SMB + HML + UMD, data = .)
    bind_cols(
      tidy(fit)  |> select(term, estimate, std.error, statistic, p.value),
      glance(fit)|> select(r.squared, adj.r.squared) |> slice(1)
    )
  }) |>
  filter(term %in% c("(Intercept)", "RMRF", "SMB", "HML", "UMD")) |>
  mutate(term = recode(term,
    "(Intercept)" = "alpha", "RMRF" = "b_mkt",
    "SMB" = "b_smb", "HML" = "b_hml", "UMD" = "b_umd"
  )) |>
  pivot_wider(
    names_from  = term,
    values_from = c(estimate, std.error, statistic, p.value),
    names_glue  = "{term}_{.value}"
  ) |>
  mutate(alpha_ann = alpha_estimate * 12) |>
  ungroup()

carhart_results |>
  arrange(desc(alpha_ann)) |>
  select(name,
         alpha_ann, alpha_statistic, alpha_p.value,
         b_mkt_estimate, b_smb_estimate,
         b_hml_estimate, b_umd_estimate,
         r.squared) |>
```

```r
  gt() |>
  fmt_number(columns = c(alpha_ann, alpha_statistic,
                         b_mkt_estimate, b_smb_estimate,
                         b_hml_estimate, b_umd_estimate,
                         r.squared),
             decimals = 3) |>
  fmt(columns = alpha_p.value, fns = \(x) sprintf("%.3f", x)) |>
  cols_label(
    name             = "ETF",
    alpha_ann        = "  (Ann.)",
    alpha_statistic  = "t( )",
    alpha_p.value    = "p( )",
    b_mkt_estimate   = " (MKT)",
    b_smb_estimate   = " (SMB)",
    b_hml_estimate   = " (HML)",
    b_umd_estimate   = " (UMD)",
    r.squared        = "R²"
  ) |>
  tab_header(
    title    = "Carhart Four-Factor Model",
    subtitle = "Wermers (2011) baseline model - sorted by annualized alpha"
  ) |>
  tab_style(
    style     = cell_fill(color = "#f0f4ff"),
    locations = cells_body(rows = alpha_p.value < 0.10)
  )
```

9.1 Interpreting the Results

Factor loadings tell us about passive style exposures:

- $\beta(MKT) > 1$: more market-sensitive than the index (e.g., ARKK)
- $\beta(SMB) > 0$: tilts toward small-cap stocks (e.g., IWM)
- $\beta(HML) > 0$: value tilt (e.g., IWD, VTV, XLE); $\beta(HML) < 0$: growth tilt (e.g., QQQ, VUG)
- $\beta(UMD) > 0$: momentum tilt; < 0: contrarian or mean-reverting strategy

Alpha after controlling for all four factors measures what is *not* explained by these passive risk exposures. For passive index ETFs, we expect:

$$\alpha \approx 0 - \text{expense ratio}$$

Carhart Four-Factor Model

Wermers (2011) baseline model — sorted by annualized alpha

ETF	(Ann.)	t( )	p( )	(MKT)	(SMB)	(HML)	(UMD)	R^2
Nasdaq-100	0.033	1.980	0.050	1.082	-0.174	-0.416	-0.070	0.928
Vanguard Growth	0.012	1.117	0.266	1.074	-0.152	-0.358	-0.059	0.967
S&P 500	0.001	0.214	0.831	0.981	-0.144	0.011	-0.004	0.996
Financials Sector	-0.002	-0.088	0.930	1.061	0.040	0.615	0.053	0.888
Vanguard Value	-0.009	-0.768	0.444	0.886	-0.100	0.364	0.023	0.948
ARK Innovation	-0.014	-0.245	0.807	1.518	1.210	-0.888	-0.075	0.761
Russell 2000	-0.024	-3.049	0.003	1.032	0.883	0.245	0.068	0.987
Energy Sector	-0.024	-0.424	0.672	1.093	0.094	1.114	-0.171	0.675
Russell 1000 Value	-0.025	-2.327	0.022	0.915	-0.023	0.339	-0.023	0.959
Real Estate (REIT)	-0.052	-1.306	0.194	0.857	0.151	0.066	0.072	0.571

A passive ETF like SPY should produce $\alpha \approx -0.09\%$/year (its expense ratio) once we properly control for all factors.

---

10 9. Comparing Alpha Across the Three Models

A key lesson: alpha estimates are *model-dependent*. Adding factors changes (and usually shrinks) alpha.

```
alpha_compare <- capm_results |>
  select(name, capm_alpha = alpha_ann) |>
  left_join(
    ff3_results |> select(name, ff3_alpha = alpha_ann),
    by = "name"
  ) |>
  left_join(
    carhart_results |> select(name, carhart_alpha = alpha_ann),
    by = "name"
  ) |>
  arrange(desc(carhart_alpha))

alpha_compare |>
  gt() |>
```

```
  fmt_number(columns = c(capm_alpha, ff3_alpha, carhart_alpha), decimals = 3) |>
  cols_label(
    name           = "ETF",
    capm_alpha     = "CAPM  ",
    ff3_alpha      = "FF3  ",
    carhart_alpha  = "Carhart  "
  ) |>
  tab_header(
    title    = "Alpha Comparison Across Models",
    subtitle = "Annualized; adding factors changes-usually shrinks-measured alpha"
  ) |>
  tab_spanner(
    label   = "Annualized Alpha",
    columns = c(capm_alpha, ff3_alpha, carhart_alpha)
  ) |>
  data_color(
    columns = carhart_alpha,
    fn      = scales::col_numeric(
      palette = c("tomato", "white", "steelblue"),
      domain  = NULL
    )
  )
```

---

11 10. Visualizing Factor Loadings

```
loadings_long <- carhart_results |>
  select(name, b_mkt_estimate, b_smb_estimate,
         b_hml_estimate, b_umd_estimate) |>
  pivot_longer(
    cols      = -name,
    names_to  = "factor",
    values_to = "loading"
  ) |>
  mutate(factor = recode(factor,
    "b_mkt_estimate" = "MKT",
    "b_smb_estimate" = "SMB",
    "b_hml_estimate" = "HML",
```

Table 4: Alpha (annualized) across three models

Alpha Comparison Across Models

Annualized; adding factors changes—usually shrinks—measured alpha

	Annualized Alpha		
ETF	CAPM	FF3	Carhart
Nasdaq-100	0.045	0.029	0.033
Vanguard Growth	0.023	0.009	0.012
S&P 500	0.005	0.000	0.001
Financials Sector	-0.015	0.001	-0.002
Vanguard Value	-0.013	-0.008	-0.009
ARK Innovation	-0.042	-0.018	-0.014
Russell 2000	-0.058	-0.021	-0.024
Energy Sector	-0.067	-0.032	-0.024
Russell 1000 Value	-0.034	-0.026	-0.025
Real Estate (REIT)	-0.054	-0.048	-0.052

```
    "b_umd_estimate" = "UMD"
  ))

loadings_long |>
  ggplot(aes(x = loading, y = fct_reorder(name, loading),
             fill = factor)) +
  geom_col(show.legend = FALSE) +
  geom_vline(xintercept = 0, linewidth = 0.4) +
  facet_wrap(~factor, scales = "free_x", nrow = 1) +
  labs(
    title = "Carhart Factor Loadings",
    x = "Loading", y = NULL
  ) +
  theme_minimal() +
  theme(axis.text.y = element_text(size = 7))
```

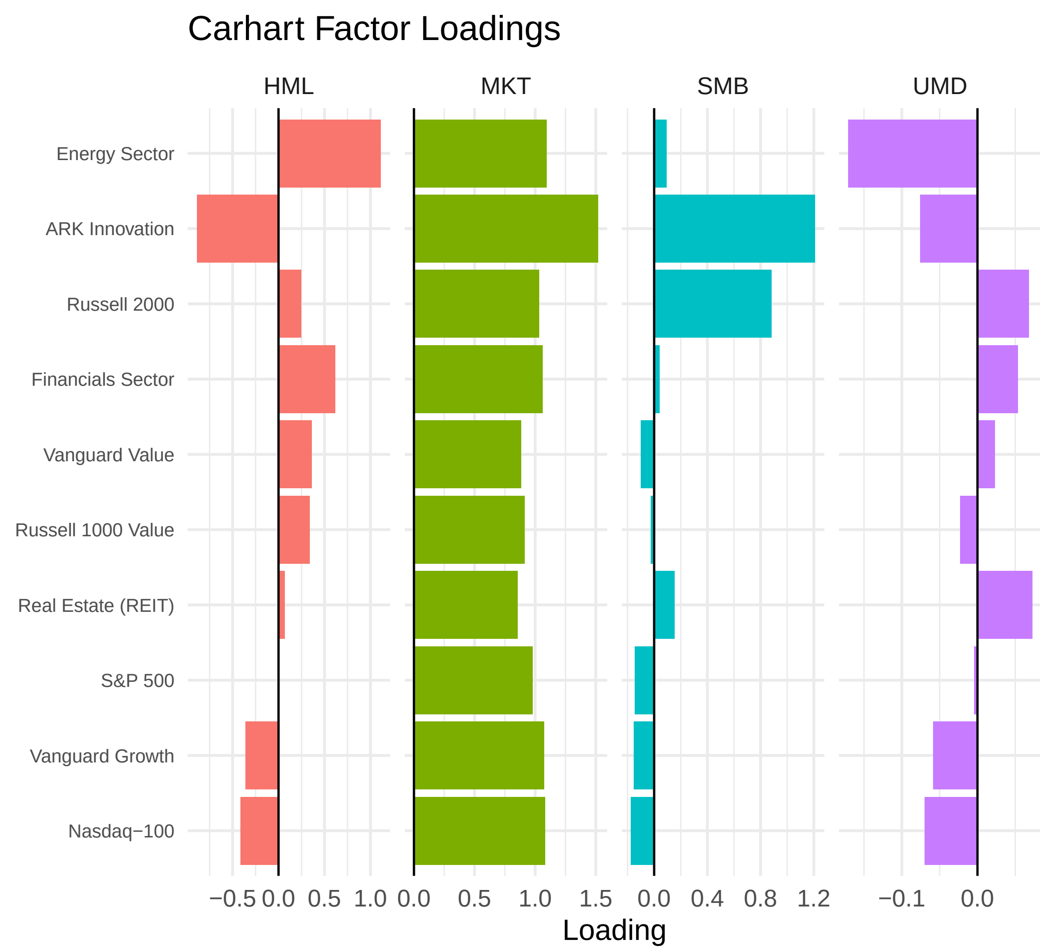

Figure 2: Carhart factor loadings across 10 US ETFs

---

12 11. R² as a Diagnostic

Wermers (2011) notes that managed long-only portfolios typically have $R^2 > 90\%$ under the Carhart model, precisely because these four factors capture most passive variation. Low R^2 signals idiosyncratic strategies not captured by the standard factors.

```
carhart_results |>
  select(name, r.squared) |>
  ggplot(aes(x = r.squared, y = fct_reorder(name, r.squared))) +
```

```r
  geom_col(fill = "steelblue") +
  geom_text(aes(label = percent(r.squared, accuracy = 0.1)),
            hjust = -0.1, size = 3) +
  scale_x_continuous(labels = percent_format(), limits = c(0, 1.05)) +
  labs(title = "R² from Carhart 4-Factor Model",
       x = "R-squared", y = NULL) +
  theme_minimal()
```

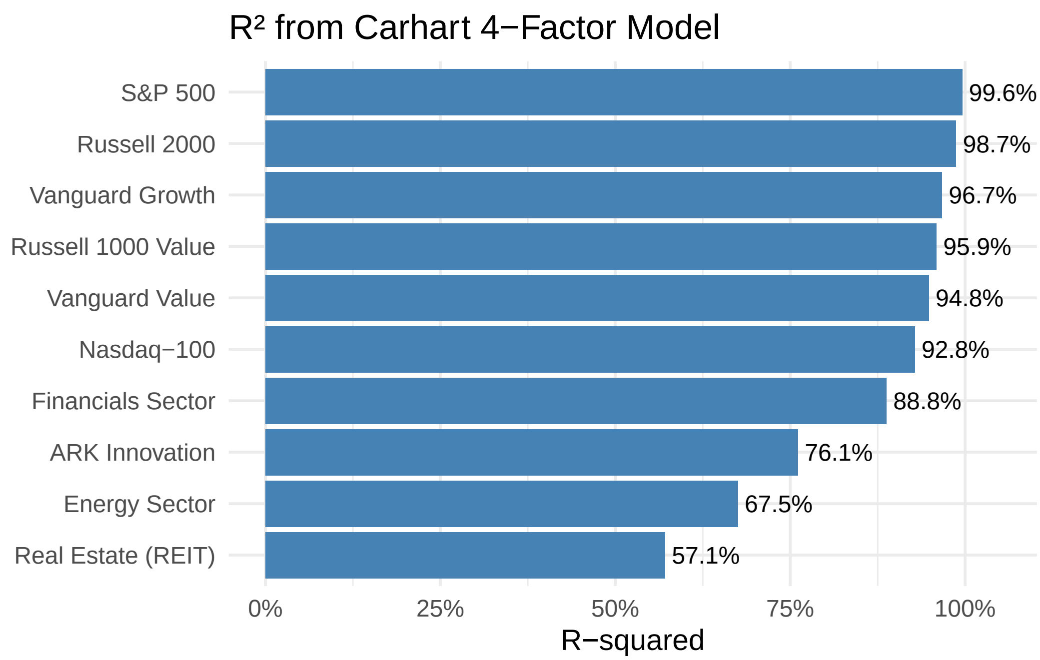

Figure 3: R^2 from Carhart model — how much return variation is explained by the four factors?

> **Note**
>
> **ARKK** typically has a lower R^2 than passive ETFs—consistent with its active, concentrated strategy in disruptive-technology stocks. Wermers (2011) discusses this in the context of hedge funds (Section 6.3): funds with lower R^2 tend to take more idiosyncratic bets, which may or may not translate into outperformance.

13 12. What Alpha Means for Passive ETFs

For **passive index ETFs** like SPY and IWM, alpha should be:

$$\alpha \approx -\text{expense ratio} \approx 0$$

Why? If markets are competitive and these factors are the right benchmark, a well-implemented passive fund should have zero alpha before costs. Any measured positive alpha from a passive fund is a red flag—it may indicate **benchmark misspecification**.

For **actively managed** funds like ARKK, we are genuinely asking: did the manager add value beyond passive factor exposure? The answer is in $\hat{\alpha}$ and its t-statistic.

> ❗ Important
>
> **A positive (or negative) alpha that is statistically insignificant should be interpreted cautiously.** With 10 years of monthly data (T 120), the power to detect even a 2% annual alpha at conventional significance levels is limited. Wermers (2011) (Section 3.3) discusses how the Kosowski et al. (2006) bootstrap method addresses this problem for large cross-sections of funds.

---

14 13. Summary Statistics

```r
summary_tbl <- data_merged |>
  group_by(symbol, name, category) |>
  summarise(
    ann_ret     = (prod(1 + ret))^(12 / n()) - 1,
    ann_vol     = sd(ret) * sqrt(12),
    sharpe      = (mean(ret - RF) / sd(ret)) * sqrt(12),
    .groups     = "drop"
  ) |>
  left_join(
    carhart_results |>
      select(name, alpha_ann, b_mkt_estimate, r.squared),
    by = "name"
  ) |>
  arrange(desc(sharpe))
```

Table 5: Fund performance summary: annualized statistics

Performance Summary
Sample: 2015-01-01 to 2024-12-31

ETF	Category	Ann. Return	Ann. Vol.	Sharpe	Carhart	Beta	R^2
Nasdaq-100	Tech/Growth	18.4%	18.5%	0.919	0.033	1.082	0.928
Vanguard Growth	Large growth	15.9%	17.8%	0.826	0.012	1.074	0.967
S&P 500	Broad market	13.0%	15.3%	0.771	0.001	0.981	0.996
Vanguard Value	Large value	10.0%	15.0%	0.600	-0.009	0.886	0.948
Financials Sector	Sector	11.3%	19.8%	0.556	-0.002	1.061	0.888
Russell 1000 Value	Large value	8.3%	15.8%	0.479	-0.025	0.915	0.959
ARK Innovation	Active growth	12.2%	36.2%	0.450	-0.014	1.518	0.761
Russell 2000	Small cap	7.8%	20.7%	0.386	-0.024	1.032	0.987
Real Estate (REIT)	REIT	4.8%	18.2%	0.258	-0.052	0.857	0.571
Energy Sector	Sector	4.8%	29.9%	0.248	-0.024	1.093	0.675

```r
summary_tbl |>
  select(name, category, ann_ret, ann_vol, sharpe, alpha_ann, b_mkt_estimate, r.squared) |>
  gt() |>
  fmt_percent(columns = c(ann_ret, ann_vol), decimals = 1) |>
  fmt_number(columns = c(sharpe, alpha_ann, b_mkt_estimate, r.squared), decimals = 3) |>
  cols_label(
    name              = "ETF",
    category          = "Category",
    ann_ret           = "Ann. Return",
    ann_vol           = "Ann. Vol.",
    sharpe            = "Sharpe",
    alpha_ann         = "Carhart  ",
    b_mkt_estimate    = "Beta",
    r.squared         = "R²"
  ) |>
  tab_header(
    title    = "Performance Summary",
    subtitle = paste("Sample:", start_date, "to", end_date)
  )
```

15 14. Key Takeaways

1. **The Carhart four-factor model is the academic standard** for evaluating equity fund performance (Wermers 2011). It controls for market, size, value, and momentum exposures—the four main dimensions of passive strategy that any uninformed investor can exploit.
2. **Alpha is model-dependent.** CAPM alpha, FF3 alpha, and Carhart alpha can differ substantially. A fund that looks good in CAPM may simply be loading on the size or value premium, which requires no skill.
3. **For passive ETFs, $\alpha \approx 0$.** The expense ratio creates a small negative alpha. Any significant positive alpha from a passive ETF signals factor model misspecification.
4. **Factor loadings reveal investment style.** Negative HML loading → growth/tech; Positive SMB → small cap; Positive UMD → momentum strategy.
5. **R^2 near or above 90% is expected** for diversified long-only portfolios. Low R^2 (like ARKK) indicates concentrated, idiosyncratic bets—which may or may not reflect skill (Wermers 2011, sec. 3.3).
6. **Statistical power is limited** with 10 years of monthly data. A 2% annual alpha may not be detectable at the 5% level with only 120 observations.

---

16 Discussion Questions

1. SPY tracks the S&P 500. Theoretically, what should its Carhart alpha be? If you find a positive alpha, what are three possible explanations?
2. IWM (Russell 2000) will have a large positive SMB loading. Does this mean IWM *outperforms*? Or that it merely *holds small stocks*? What is the difference?
3. ARKK is actively managed. Suppose its Carhart alpha is positive but statistically insignificant ($p = 0.15$). How should an investor interpret this?
4. Compare the CAPM alpha and Carhart alpha of QQQ. Why do they differ? What factor explains most of the difference?
5. Wermers (2011) writes that the Carhart model regression of a managed portfolio typically yields $R^2 > 90\%$. Which ETF in our sample has the lowest R^2? Why does this make intuitive sense given that fund's investment strategy?

---

17 References

Carhart, Mark M. 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance* 52 (1): 57–82.

Fama, Eugene F., and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33 (1): 3–56.

Jegadeesh, Narasimhan, and Sheridan Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance* 48 (1): 65–91.

Jensen, Michael C. 1968. "The Performance of Mutual Funds in the Period 1945–1964." *Journal of Finance* 23 (2): 389–416.

Wermers, Russ. 2011. "Performance Measurement of Mutual Funds, Hedge Funds, and Institutional Accounts." *Annual Review of Financial Economics* 3: 537–74.