

Week 3 Update: Passive Ownership Measurement and Preliminary Volatility Analysis

Ryan Levy, Srimaan Bekkari, Ryan Rammacher, West Plowman

2026-02-12

1. Data Sources and Passive Ownership Measure

This section describes the institutional ownership data used to construct a proxy for passive investment. We use 13F filings obtained through the SEC API, focusing on holdings reported by BlackRock, Vanguard, and State Street.

To measure passive ownership at the firm–year level, we compute the proportion of publicly available shares held by these institutions.

$$\text{PassiveOwnership}_{it} = \frac{\text{SharesHeldByPassive}_{it}}{\text{SharesOutstanding}_{it}}$$

```
# load in passive investment data
# gives shares held per reporting period per year aggregated between BlackRock, Vanguard, and State Street
p_own_raw <- read.csv("13f_api_outputs/BVS_13f_holdings_panel.csv")

p_own <- p_own_raw
# format date column
p_own$report_date <- as.Date(p_own$report_date)
p_own$year <- as.integer(format(p_own$report_date, "%Y"))
# divides shares held by shares outstanding for proxy passive ownership metric
p_own$percentage_owned <- p_own$shares_held_total / p_own$shares_outstanding
```

2. Illustrative Example: Apple Inc. (AAPL)

To illustrate the construction of the passive ownership metric, we present a demonstration using Apple Inc. (AAPL). We aggregate institutional holdings by year and compute annual average passive ownership.

```
# subset AAPL for demo
p_own_appl <- p_own[p_own$ticker == "AAPL",]

# consolidates averages per year
p_own_appl_yr <- aggregate(percentage_owned ~ year, data = p_own_appl, FUN = mean)

head(p_own_appl_yr)
```

```

##   year percentage_owned
## 1 2013      0.06950798
## 2 2014      0.16835350
## 3 2015      0.04710634
## 4 2016      0.06742886
## 5 2017      0.04316916
## 6 2018      0.04246568

```

3. Stock Price Data and Volatility Construction

We next introduce daily stock price data to construct a volatility measure that can be compared to passive ownership levels. Daily log returns are computed and aggregated into an annualized realized volatility measure.

```

# load in data
aapl_raw <- read.csv("ipo_to_2026_stock_volatility/aapl_us_d.csv")
aapl <- aapl_raw
# format date
aapl$data <- as.Date(aapl$data)
aapl$year <- as.integer(format(aapl$data, "%Y"))

# daily log return  $r_t = \ln(Close_t / Close_{t-1})$ 
aapl$log_ret <- c(NA, diff(log(aapl$Close)))

# year and annualized realized volatility
years <- sort(unique(aapl$year))
aapl_vol_annual <- data.frame(
  year = years,
  ann_vol = NA_real_,
  n_days = NA_integer_,
  stringsAsFactors = FALSE
)

# compute  $\sigma_y = sd(daily\ returns\ in\ year\ y) * \sqrt{252}$ 
for (i in seq_along(years)) {
  y <- years[i]

  r <- aapl$log_ret[aapl$year == y]
  r <- r[!is.na(r)]

  aapl_vol_annual$ann_vol[i] <- sd(r) * sqrt(252)
  aapl_vol_annual$n_days[i] <- length(r)
}

# Inspect
head(aapl_vol_annual, 10)

```

```

##   year ann_vol n_days
## 1 1984 0.4328972    79
## 2 1985 0.4491332   253
## 3 1986 0.3961446   253
## 4 1987 0.6663348   253

```

```

## 5 1988 0.3359687 253
## 6 1989 0.3408801 252
## 7 1990 0.4368052 253
## 8 1991 0.4700829 253
## 9 1992 0.3559901 254
## 10 1993 0.5131650 253

```

4. Merging Passive Ownership and Volatility Data

The passive ownership and volatility datasets are merged at the yearly level. This allows for direct comparison between changes in institutional ownership and realized stock volatility.

```

merged_df <- merge(
  p_own_aapl_yr,
  aapl_vol_annual,
  by = "year",
  all = FALSE
)

merged_df[,c("year", "percentage_owned", "ann_vol")]

##      year percentage_owned    ann_vol
## 1 2013        0.06950798 0.2881451
## 2 2014        0.16835350 0.2166937
## 3 2015        0.04710634 0.2668712
## 4 2016        0.06742886 0.2335747
## 5 2017        0.04316916 0.1754234
## 6 2018        0.04246568 0.2877036
## 7 2019        0.04403561 0.2629001
## 8 2020        0.16272072 0.4666700
## 9 2021        0.07683517 0.2508762
## 10 2022       0.09532210 0.3563516
## 11 2023       0.09812040 0.2027984
## 12 2024       0.03990372 0.2259901
## 13 2025       0.03202567 0.3221429

par(mar = c(5, 4, 4, 4) + 0.1)

plot(
  merged_df$year,
  merged_df$percentage_owned,
  type = "o",
  pch = 16,
  xlab = "Year",
  ylab = "Passive Ownership"
)

par(new = TRUE)

plot(
  merged_df$year,
  merged_df$ann_vol,

```

```

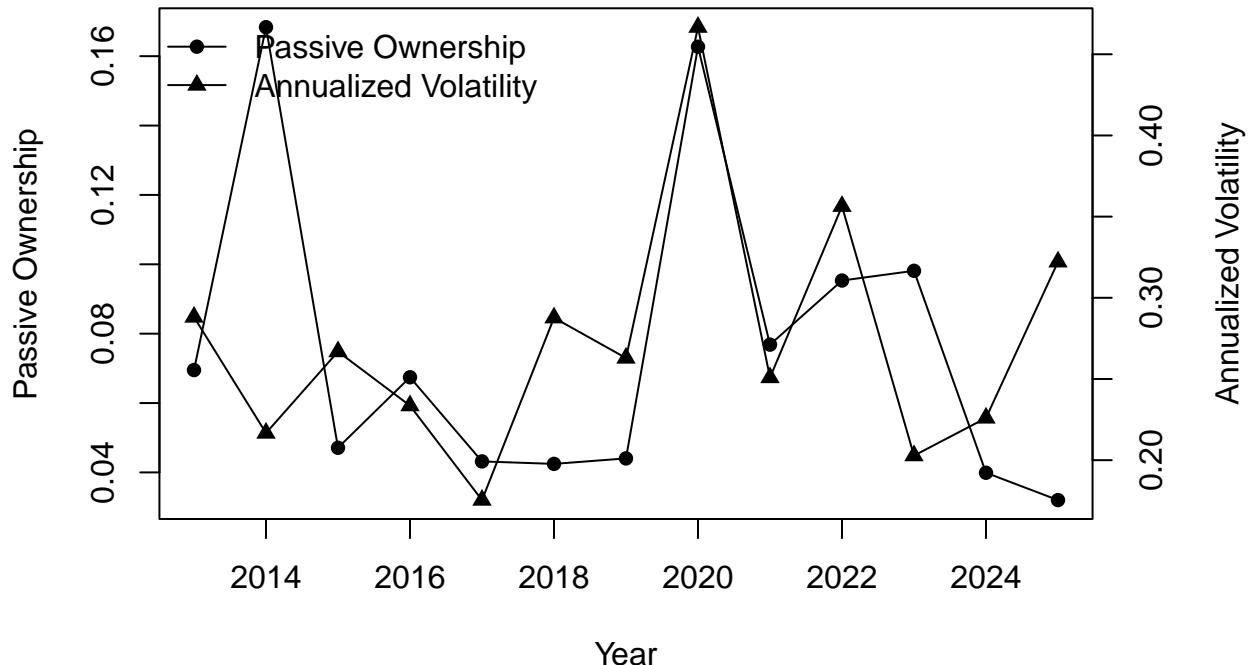
    type = "o",
    pch = 17,
    axes = FALSE,
    xlab = "",
    ylab = ""
)

axis(side = 4)
mtext("Annualized Volatility", side = 4, line = 3)
title("AAPL: Passive Ownership vs Annualized Volatility", line = 2)

legend(
  "topleft",
  legend = c("Passive Ownership", "Annualized Volatility"),
  pch = c(16, 17),
  lty = 1,
  bty = "n"
)

```

AAPL: Passive Ownership vs Annualized Volatility



Aside from the large spike in passive ownership in 2014, the two lines seem to follow a similar pattern.

5. High-Frequency Visualization (Non-Annualized Data)

To examine the relationship in finer detail, we plot quarterly passive ownership alongside daily stock price data beginning in 2010.

```

aapl_raw_2 <- aapl_raw
p_own_raw_2 <- p_own_raw

aapl_raw_2$Data <- as.Date(aapl_raw_2$Data)
p_own_raw_2$report_date <- as.Date(p_own_raw_2$report_date)

p_own_raw_2$percentage_owned <- p_own_raw_2$shares_held_total / p_own_raw_2$shares_outstanding

p_own_raw_2 <- p_own_raw_2[p_own_raw_2$ticker == "AAPL",]

head(aapl_raw_2)

##           Data      Open   Highest   Lowest    Close   Volume
## 1 1984-09-07 0.0992653 0.1004840 0.0980668 0.0992653 99149603
## 2 1984-09-10 0.0992653 0.0995698 0.0968786 0.0986760 76956267
## 3 1984-09-11 0.0995698 0.1022710 0.0995698 0.1004840 181467446
## 4 1984-09-12 0.1004840 0.1010710 0.0974581 0.0974581 158527291
## 5 1984-09-13 0.1028800 0.1031730 0.1028800 0.1028800 246900394
## 6 1984-09-14 0.1031730 0.1067600 0.1031730 0.1043630 293981051

head(p_own_raw_2)

##   group ticker mapped_cusip report_date shares_held_total
## 1   BVS   AAPL     37833100 2013-06-30          63964616
## 2   BVS   AAPL     37833100 2013-09-30          63684370
## 3   BVS   AAPL     37833100 2013-12-31          63233534
## 4   BVS   AAPL     37833100 2014-03-31          60095703
## 5   BVS   AAPL     37833100 2014-06-30          404130894
## 6   BVS   AAPL     37833100 2014-09-30          402715040
##   value_usd_thousands_total shares_outstanding percentage_owned
## 1                  25346521          938649000        0.06814540
## 2                  30361529          908497000        0.07009860
## 3                  35478810          899738000        0.07027994
## 4                  32255776          891989000        0.06737270
## 5                  37555883          861381000        0.46916625
## 6                  40573542          5987867000       0.06725517

cut_date <- as.Date("2010-01-01")

aapl_2010 <- subset(aapl_raw_2, Data >= cut_date)
p_2010 <- subset(p_own_raw_2, ticker == "AAPL" & report_date >= cut_date)

plot(aapl_2010$Data, aapl_2010$Close, type="l",
      xlab="Date", ylab="Close Price",
      main="AAPL: Close vs Passive Ownership (2010+)")
par(new=TRUE)

plot(p_2010$report_date, p_2010$percentage_owned,
      type="l", lty=3, axes=FALSE, xlab="", ylab="")
axis(side=4)

```

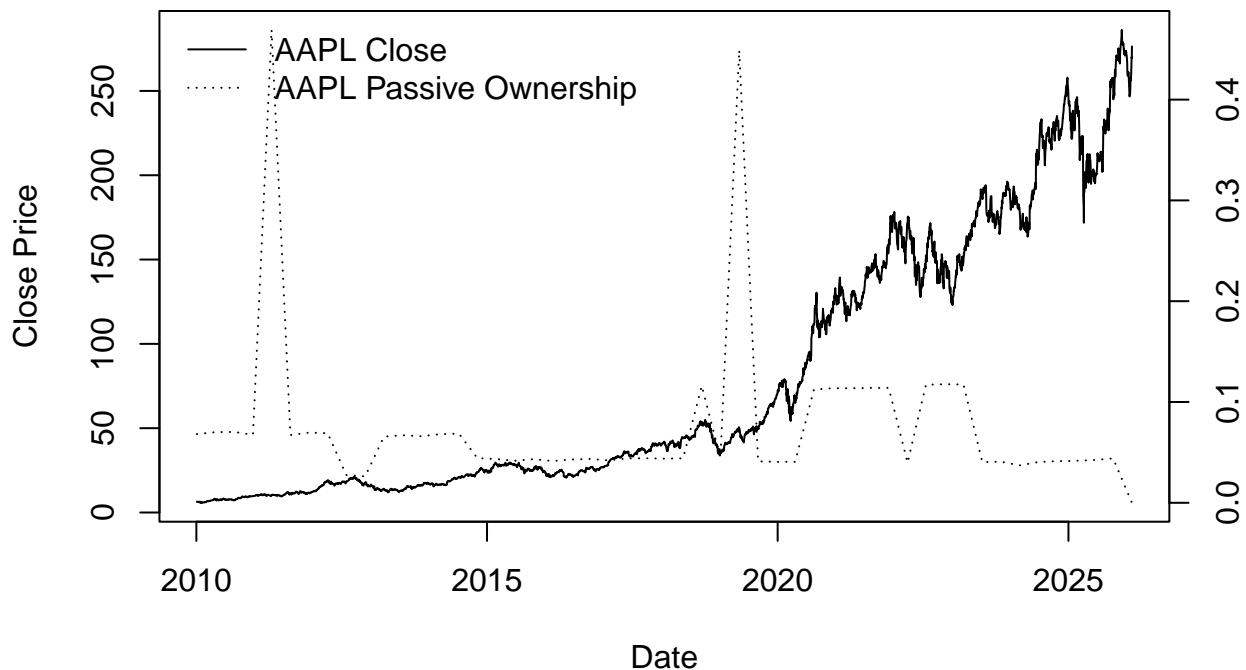
```

mtext("Passive Ownership", side=4, line=3)

legend("topleft",
       legend = c("AAPL Close", "AAPL Passive Ownership"),
       lty = c(1, 3),
       bty = "n")

```

AAPL: Close vs Passive Ownership (2010+)



Two periods of sharp increases in passive ownership are visible and warrant further investigation. Perhaps also removing the linear trend from the data to only see the volatility would help.

```

# compute log returns (centered around 0)
aapl_2010$log_ret <- c(NA, diff(log(aapl_2010$Close)))

plot(aapl_2010$Data, aapl_2010$log_ret, type="l",
      xlab="Date", ylab="Daily Log Return",
      main="AAPL Returns vs Passive Ownership (2010+")

abline(h = 0, lty = 2) # zero reference line

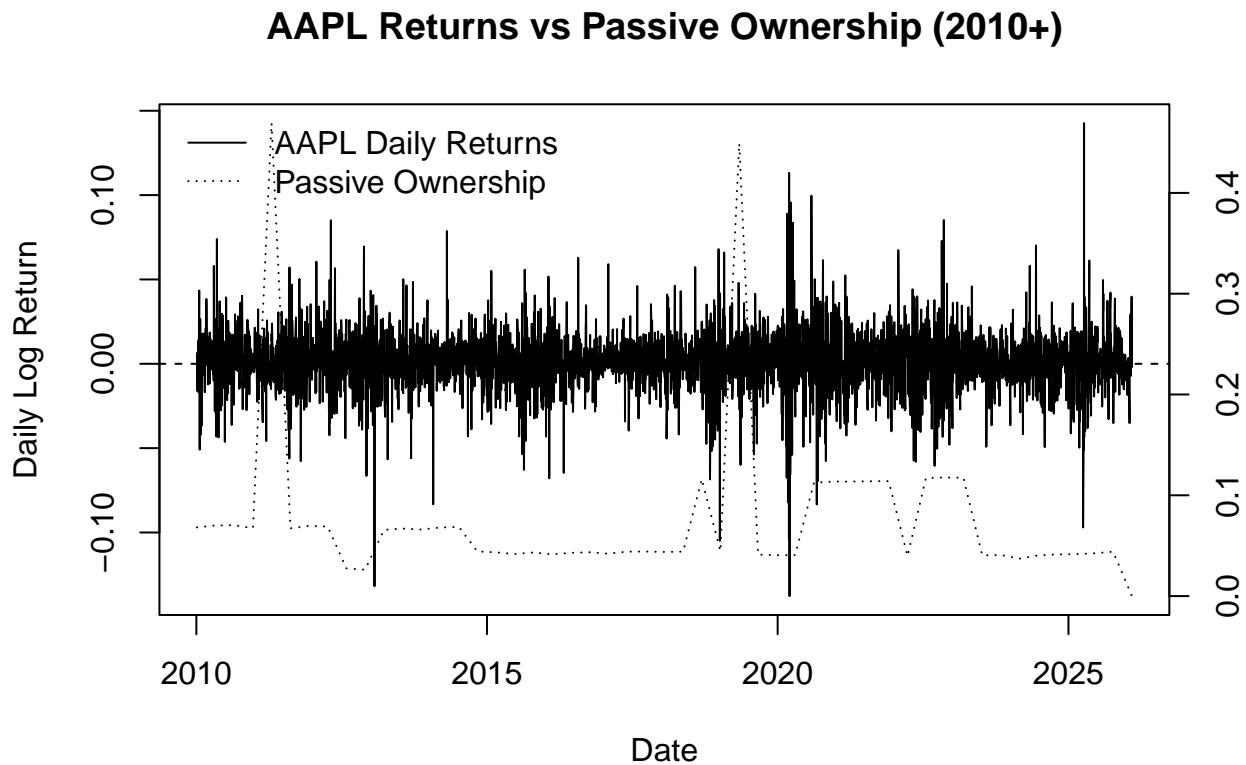
par(new=TRUE)

plot(p_2010$report_date, p_2010$percentage_owned,
      type="l", lty=3, axes=FALSE, xlab="", ylab "")

axis(side=4)
mtext("Passive Ownership", side=4, line=3)

```

```
legend("topleft",
       legend = c("AAPL Daily Returns", "Passive Ownership"),
       lty = c(1, 3),
       bty = "n")
```



It seems like large spikes of volatility follow the sharp increase in passive ownership, but further analysis is still to be done.