Times series data updates

Zhongwei Yao

2023-06-08

Contents

List of Figures

List of Tables

1 Stock market returns

1.1 File path

```
market_pfmc_m_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□□/□□□□□□□/□□□□□□□□□/TRD_Cnmont1990-12 □ 2023
```

1.2 Monthly returns

```
mkt_pfmc_m <- readxl::read_excel(market_pfmc_m_path) %>%
  setlowercolnames() %>%
  #< 117: □□□A□□□□□□□□□□; 53: □□A□□□□□□□□□□
  filter(markettype == 117) %>%
  mutate(tradingmonth = ymd(str_c(trdmnt, "-01"))) %>%
  select(tradingmonth, cmretwdos) %>%
  mutate(cmretwdos = as.double(cmretwdos)) %>%
  na.omit()
```

1.2.1 Cumulative MRET: Backward

```
marketret_m_ds <- mkt_pfmc_m
for (i in seq(1, 59, 1)){
  m <- i + 1
  marketret_m_ds <- bind_cols(
    marketret_m_ds,
    tibble(
      "marketret_m_backward_{{m}}m" := slide_period(mkt_pfmc_m$cmretwdos, mkt_pfmc_m$tradingmonth, .period = "montl
    )
  )
}
```

1.2.2 Cumulative MRET: Forward

```
for (i in seq(1, 60, 1)){
  marketret_m_ds <- bind_cols(
    marketret_m_ds,
    tibble(
      "marketret_m_forward_{{i}}m" := slide_period(mkt_pfmc_m$cmretwdos, mkt_pfmc_m$tradingmonth, .period = "month'
    )
  )
}

marketret_m_csmar <- unnest(marketret_m_ds, everything()) %>%
  rename(marketret_m_backward_1m = cmretwdos) %>%
  mutate(tradingmonth = ceiling_date(tradingmonth, "m") -days(1))

#< Write to the local database
dbWriteTable(conn_macro,
    "marketret_m_csmar",
    value = marketret_m_csmar,
    overwrite = TRUE
  )
```

2 Market volatility

2.1 File path

```
market_pfmc_d_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□□/□□□□□□□/□□□□□□□□□/TRD_Cndalym1990-12-19 □ 2
```

2.2 Stock market variance

252 A

```
market_svar_m_csmar <- data.table::fread(market_pfmc_d_path) %>%
  tibble() %>%
  setlowercolnames() %>%
  filter(markettype == 117) %>%
  select(tradingdate = trddt, cdretwdos) %>%
  mutate(tradingdate = ymd(tradingdate), svar = cdretwdos * cdretwdos) %>%
  na.omit() %>%
  arrange(tradingdate) %>%
  mutate(svar = rollsumr(svar, 252, fill = NA)) %>%
  group_by(year(tradingdate), month(tradingdate)) %>%
  summarise(tradingdate = last(tradingdate), svar = last(svar)) %>%
  ungroup() %>%
  na.omit() %>%
  mutate(tradingmonth = ceiling_date(tradingdate, "m") - days(1)) %>%
  select(tradingmonth, svar)


#< Write to the local database
dbWriteTable(conn_macro,
    "market_svar_m_csmar",
    value = market_svar_m_csmar,
    overwrite = TRUE
  )
```

3 Market turnover

3.1 File path

```
stock_ret_month_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□□/□□□□□□□/□□□□□□/□□□□□□□/TRD_Mnth1990-12 □ 2
```

3.2 Monthly trading volume: All A shares

Bottom-up average with all a shares

```
market_turnover_m_csmar <- fread(stock_ret_month_path, colClasses = c(Stkcd = "character")) %>%
  setlowercolnames() %>%
  .[markettype %in% c(1, 4, 16, 32)] %>%
  .[, tradingmonth := ymd(paste0(trdmnt, "-01"))] %>%
  .[, .(stkcd, tradingmonth, mclsprc, mnshrtrd, mnvaltrd, msmvosd)] %>%
  .[, mnshrfloata := msmvosd * 1000 / mclsprc] %>%
  .[, `:=`(turnover_1 = mnshrtrd / mnshrfloata, turnover_2 = mnvaltrd / (msmvosd * 1000))] %>%
  na.omit() %>%
  .[, lapply(.SD, function(x) weighted.mean(x,w = msmvosd)), .SDcols = c("turnover_1", "turnover_2"), tradingmonth
  setorder(tradingmonth) %>%
  .[, (c("turnover_12m_mean_backward_1", "turnover_12m_mean_backward_2")) := lapply(.SD, RcppRoll::roll_meanr, n =
  .[, (c("turnover_12m_mean_forward_1", "turnover_12m_mean_forward_2")) := lapply(.SD, RcppRoll::roll_meanl, n = 1
  .[, tradingmonth := ceiling_date(tradingmonth, "m") -days(1)] %>%
  .[]

#< Write to the local database
dbWriteTable(conn_macro,
    "market_turnover_m_csmar",
    value = market_turnover_m_csmar,
    overwrite = TRUE
  )
```

4 Inflation

4.1 File path

```
cpi_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□□/□□□□/□□□□/□□□□□□□□□□□□□□□/CME_Mpi12003-02 □ 2023-02
```

4.2 Monthly CPI month-to-month

- Datasgn [ ] - PYM , PYP .
- Areasgn [ ] - 1= 2= 3=
- Epim0101 [ ] -
- Epim0102 [-] -
- Epim0103 [-] -
- Epim0104 [-] -
- Epim0105 [-] -

- Epim0106 [-] -
- Epim0107 [-] -
- Epim0108 [-] -
- Epim0109 [-] -

```
cpi_m_csmar <- readxl::read_xlsx(cpi_path) %>%
  slice(3:n()) %>%
  filter(Areasgn == 1, Datasgn == "PYM") %>%
  mutate(tradingmonth = ceiling_date(ymd(paste0(Staper, "-01")), "m") - days(1),
         year = year(tradingmonth),
         cpi = as.double(Epim0101)) %>%
  # filter(month(tradingmonth) == 12, !is.na(cpi), year >= 2000) %>%
  select(tradingmonth, cpi) %>%
  mutate(inflation = cpi - 100) %>%
  #< Because inflation information is released only in the following month, we wait for one month before using it
  mutate(inflation = dplyr::lag(inflation, 1)) %>%
  na.omit() %>%
  select(-cpi)

#< Write to the local database
dbWriteTable(conn_macro,
    "cpi_m_csmar",
    value = cpi_m_csmar,
    overwrite = TRUE
  )
```

5 Net equity expansion

5.1 File path

```
ntis_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□□/□□□□/□□□□/IPO□□□/ER_IPO.txt"
```

5.2 NTIS

12 / A

```
ntis_m_csmar <- fread(ntis_path) %>%
  .[CurrencyCode == "CNY", .(raisefund = RaiseFund, enddate = fifelse(is.na(EndDate), ListedDate, StartDate))] %>%
  .[, tradingmonth := floor_date(enddate, "m")] %>%
  .[, .(raisefund = sum(raisefund, na.rm = T)), tradingmonth] %>%
```

```
  .[data.table(tradingmonth = seq.Date(ymd("2000-1-1"), ymd("2023-1-1"), by = "month")), on = "tradingmonth"] %>%
  setorder(tradingmonth) %>%
  .[, ntis :=  roll_sumr(raisefund, n = 12, fill = NA, na.rm = T)] %>%
  .[, .(tradingmonth, ntis)] %>%
  left_join(
    readxl::read_excel(market_pfmc_m_path) %>%
      setlowercolnames() %>%
      filter(markettype == 117) %>%
      mutate(tradingmonth = ymd(str_c(trdmnt, "-01"))) %>%
      select(tradingmonth, cmmvosd) %>%
      na.omit()
  ) %>%
  as.data.table() %>%
  na.omit() %>%
  .[, .(tradingmonth = ceiling_date(tradingmonth, "m") - days(1), ntis = ntis / (cmmvosd * 1000))]
#< Write to the local database
dbWriteTable(conn_macro,
    "ntis_m_csmar",
    value = ntis_m_csmar,
    overwrite = TRUE
  )
```

6 Government bond yields

```
gb_path <- "/Volumes/Samsung_T7/Research/Database/□□/□□□□□□□□□□□□□□□/"
```

6.1

```
gb_ts <- paste0(gb_path, dir(gb_path, ".xlsx")) %>%
  map_dfr(
    ~readxl::read_xlsx(.x) %>%
      select(yield = `□□□(%)`, tenure = `□□□□□□`, tradingdate = `□□`) %>%
      mutate(yield = as.double(yield), tradingdate = ymd(tradingdate), tradingmonth = ceiling_date(tradingdate, "m
  ) %>%
  group_by(tradingmonth, tenure) %>%
  summarise_all(last) %>%
  ungroup() %>%
  select(-tradingdate)
```

6.2 Short-term yield: 3m

```
sty_3m_m_cb <- gb_ts %>%
  filter(tenure == "3m") %>%
  select(-tenure) %>%
  rename(sty = yield)
```

6.3 Long-term yield: 10Yr

```
lty_10yr_m_cb <- gb_ts %>%
  filter(tenure == "10y") %>%
  select(-tenure) %>%
  rename(lty = yield)
```

6.4 Termspread

```
termspread_m_cb <- lty_10yr_m_cb %>%
  left_join(sty_3m_m_cb) %>%
  mutate(termspread = lty - sty) %>%
  select(tradingmonth, termspread)

gby_m_cb <- sty_3m_m_cb %>%
  left_join(lty_10yr_m_cb) %>%
  left_join(termspread_m_cb)

#< Write to the local database
dbWriteTable(conn_macro,
    "gby_m_cb",
    value = gby_m_cb,
    overwrite = TRUE
  )
```

7 Market valuation

```
mv_path <- "/Volumes/Samsung_T7/Research/Database/WIND/□□/□□□□□□□/STK_INDEX_VALUATION_update202302.xlsx"
```

7.1 D/P, E/P, B/M

```
market_valuation_m_wind <- readxl::read_xlsx(mv_path, "WINDA") %>%
      mutate(tradingmonth = ymd(ceiling_date(ymd, "month") - days(1)), ep_winda = 1 / pe, bp_winda = 1 / pb, dp_wi
      select(tradingmonth, ep_winda, bp_winda, dp_winda, pe_winda = pe, pb_winda = pb)

#< Write to the local database
dbWriteTable(conn_macro,
    "market_valuation_m_wind",
    value = market_valuation_m_wind,
    overwrite = TRUE
  )
```

8 Risk factors

8.1 File path

```
ff3_m_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□/Fama-French□□/STK_MKT_THRFACMONTH1990-12 □ 2023-0
ff3_d_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□/Fama-French□□/STK_MKT_THRFACDAY1990-12-19 □ 2023
ff5_m_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□/Fama-French□□/STK_MKT_FIVEFACMONTH1994-01 □ 2023
ff5_d_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□/Fama-French□□/STK_MKT_FIVEFACDAY1994-01-03 □ 202
ch4_m_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□/□□□□/STK_MKT_CARHARTFOURFACTORS1990-12 □ 2023-03
cn4_m_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□/CH_4_fac_update_20211231.csv"
cn4_d_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/□□□□□□/CH_4_fac_daily_update_20211231.csv"
```

8.2 Factor returns

8.2.1 FF3

```
ff3_m_csmar <- data.table::fread(ff3_m_path) %>%
  tibble() %>%
  setlowercolnames() %>%
  filter(markettypeid == "P9714") %>%
  select(tradingmonth, rp_ff3 = riskpremium1, smb_ff3 = smb1, hml_ff3 = hml1) %>%
  mutate(tradingmonth = ymd(str_c(tradingmonth, "-01"))) %>%
  arrange(tradingmonth)
```

```
for (i in seq(1, 36, 1)){
  ff3_m_csmar <- bind_cols(
    ff3_m_csmar,
    tibble(
      "hml_ff3_m_forward_{{i}}m" := slide_period(ff3_m_csmar$hml_ff3, ff3_m_csmar$tradingmonth, .period = "month",
    )
  )
}

ff3_m_csmar <- unnest(ff3_m_csmar, everything()) %>%
  mutate(tradingmonth = ceiling_date(tradingmonth, "m") -days(1)) %>%
  filter(tradingmonth >= "2000-01-01")

#< Write to the local database
dbWriteTable(conn_macro,
    "ff3_m_csmar",
    value = ff3_m_csmar,
    overwrite = TRUE
  )
```

8.2.1.1 Cumulative FF3 returns: Forward

8.2.2 FFC4

```
ch4_m_csmar <- data.table::fread(ch4_m_path) %>%
  tibble() %>%
  setlowercolnames() %>%
  filter(markettypeid == "P9714") %>%
  select(tradingmonth,rp_ch4=riskpremium1,smb_ch4=smb1, hml_ch4 = hml1, umd_ch4 = umd1) %>%
  mutate(tradingmonth = ymd(str_c(tradingmonth, "-01")))
```

8.2.3 LSY4

https://finance.wharton.upenn.edu/~stambaug/

```
cn4_m_lsy <- data.table::fread(cn4_m_path, skip = 9) %>%
  as_tibble() %>%
  setlowercolnames() %>%
  mutate(tradingmonth = floor_date(ymd(mnthdt), unit = "m")) %>%
  select(-mnthdt) %>%
  # ■■■■■■■■■■■%
```

```
  mutate_at(vars(rf_mon:pmo), ~.x/100) %>%
  select(tradingmonth, mrf = rf_mon, rp_cn4 = mktrf, vmg_cn4 = vmg, smb_cn4 = smb, pmo_cn4 = pmo)
```

8.2.4 FF5

```
ff5_m_csmar <- data.table::fread(ff5_m_path) %>%
  tibble() %>%
  setlowercolnames() %>%
  filter(markettypeid == "P9714", portfolios == 1) %>%
  select(tradingmonth,rp_ff5=riskpremium1,smb_ff5=smb1,hml_ff5 = hml1,rmw_ff5 = rmw1,cma_ff5 = cma1) %>%
  mutate(tradingmonth = ymd(str_c(tradingmonth, "-01"))) %>%
  arrange(tradingmonth)
```

9 Merge data

```
econ_var_m <- marketret_m_csmar %>%
  left_join(market_svar_m_csmar) %>%
  left_join(market_turnover_m_csmar) %>%
  left_join(cpi_m_csmar) %>%
  left_join(ntis_m_csmar) %>%
  left_join(sty_3m_m_cb) %>%
  left_join(lty_10yr_m_cb) %>%
  left_join(termspread_m_cb) %>%
  left_join(market_valuation_m_wind) %>%
  as.data.table()
```