

Times series data updates

Zhongwei Yao

2023-06-08

Contents

1	Stock market returns	2
1.1	File path	2
1.2	Monthly returns	2
2	Market volatility	3
2.1	File path	3
2.2	Stock market variance	3
3	Market turnover	4
3.1	File path	4
3.2	Monthly trading volume: All A shares	4
4	Inflation	4
4.1	File path	5
4.2	Monthly CPI month-to-month	5
5	Net equity expansion	5
5.1	File path	5
5.2	NTIS	5

6	Government bond yields	6
6.1	6
6.2	Short-term yield: 3m	6
6.3	Long-term yield: 10Yr	7
6.4	Termspread	7
7	Market valuation	7
7.1	D/P, E/P, B/M	7
8	Merge data	7

List of Figures

List of Tables

1 Stock market returns

1.1 File path

```
market_pfmc_m_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/国泰君安/国泰君安/国泰君安/TRD_Cnmont1990-12 至 2023
```

1.2 Monthly returns

```
mkt_pfmc_m <- readxl::read_excel(market_pfmc_m_path) %>%
  setlowercolnames() %>%
  #< 117: 国泰君安; 53: 国泰君安
  filter(marketype == 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 = "month", .fill = "na")
    )
  )
}

```

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", .fill = "na")
    )
  )
}

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

```

2 Market volatility

2.1 File path

```

market_pfmc_d_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/00000000/00000000/00000000/TRD_CndaIym1990-12-19 00:00:00"

```

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)

```

3 Market turnover

3.1 File path

```
stock_ret_month_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/00000000/00000000/000000/00000000/TRD_Mnth1990-12 0 1
```

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 = 12),
  .[, (c("turnover_12m_mean_forward_1", "turnover_12m_mean_forward_2")) := lapply(.SD, RcppRoll::roll_meanl, n = 12),
  .[, tradingmonth := ceiling_date(tradingmonth, "m") - days(1)] %>%
  .[]

```

4 Inflation

4.1 File path

```
cpi_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/中国/数据/宏观经济数据库/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)

```

5 Net equity expansion

5.1 File path

```
ntis_path <- "/Volumes/Samsung_T7/Research/Database/CSMAR/00000000/00000/00000/IP00000/ER_IP0.txt"
```

5.2 NTIS

12 / A

```
ntis_m_csma <- 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_pfm_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))]

```

6 Government bond yields

```
gb_path <- "/Volumes/Samsung_T7/Research/Database/GB/GB_20190101_20191231/"
```

6.1

```

gb_ts <- paste0(gb_path, dir(gb_path, ".xlsx")) %>%
map_dfr(
  ~readxl::read_xlsx(.x) %>%
  select(yield = `收益率(%)`, tenure = `期限`, tradingdate = `日期`) %>%
  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 %>%  
  left_join(sty_3m) %>%  
  mutate(termspread = lty - sty) %>%  
  select(tradingmonth, termspread)
```

7 Market valuation

```
mv_path <- "/Volumes/Samsung_T7/Research/Database/WIND/财报/财报数据/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_winda = 1 / dp_ratio) %>%  
  select(tradingmonth, ep_winda, bp_winda, dp_winda, pe_winda = pe, pb_winda = pb)
```

8 Merge data

```
econ_var_m <- marketret_m_csma %>%  
  left_join(market_svar_m_csma) %>%  
  left_join(market_turnover_m_csma) %>%  
  left_join(cpi_m_csma) %>%  
  left_join(ntis_m_csma) %>%  
  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()
```