WILLIAM MARSH RICE UNIVERSITY

Max-Median Rule

Spring 2023

Brandon Fantine, Edward Wong, Justin Jiang, Saswat Pati

Table of Contents

Method Descriptions

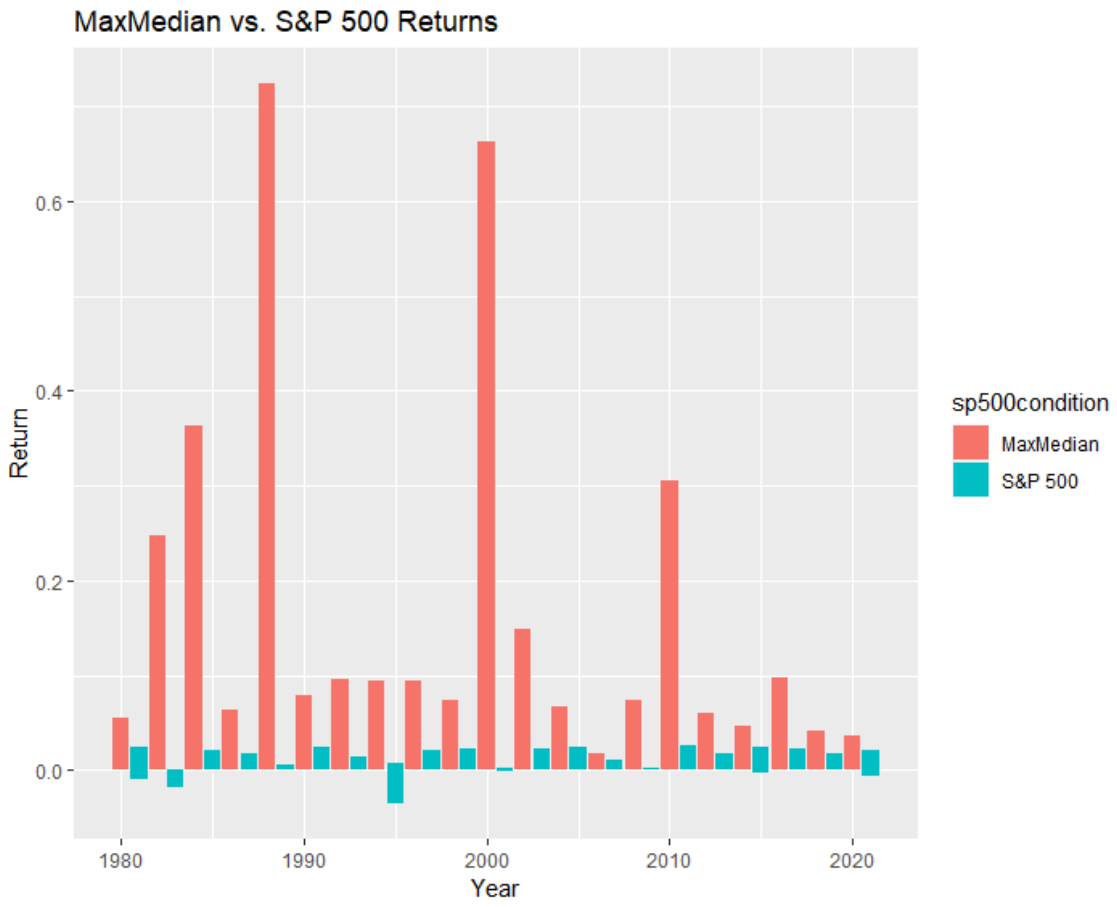
Max-Median

Introduction:

The use of quantitative techniques in finance has become increasingly popular over the years. In this paper, we present a quantitative investment strategy for the S&P 500 index based on the analysis of daily stock prices. Specifically, we look back at the historical prices of the 500 stocks in the index and compute the day-to-day ratios using the formula R(j,t) = S(j,t)/S(j,t-1). We then sort these ratios for the year's trading days and determine the median for each stock. By investing equally in the 20 stocks with the largest medians and holding them for a year before liquidating, we aim to achieve superior returns compared to traditional investment strategies.

Max-Median Rule Performance:

We track the top 20 stocks in S&P 500 with the Max-Median Rule - buy and hold for a year. Here's our results comparing to S&P 500 annual return:

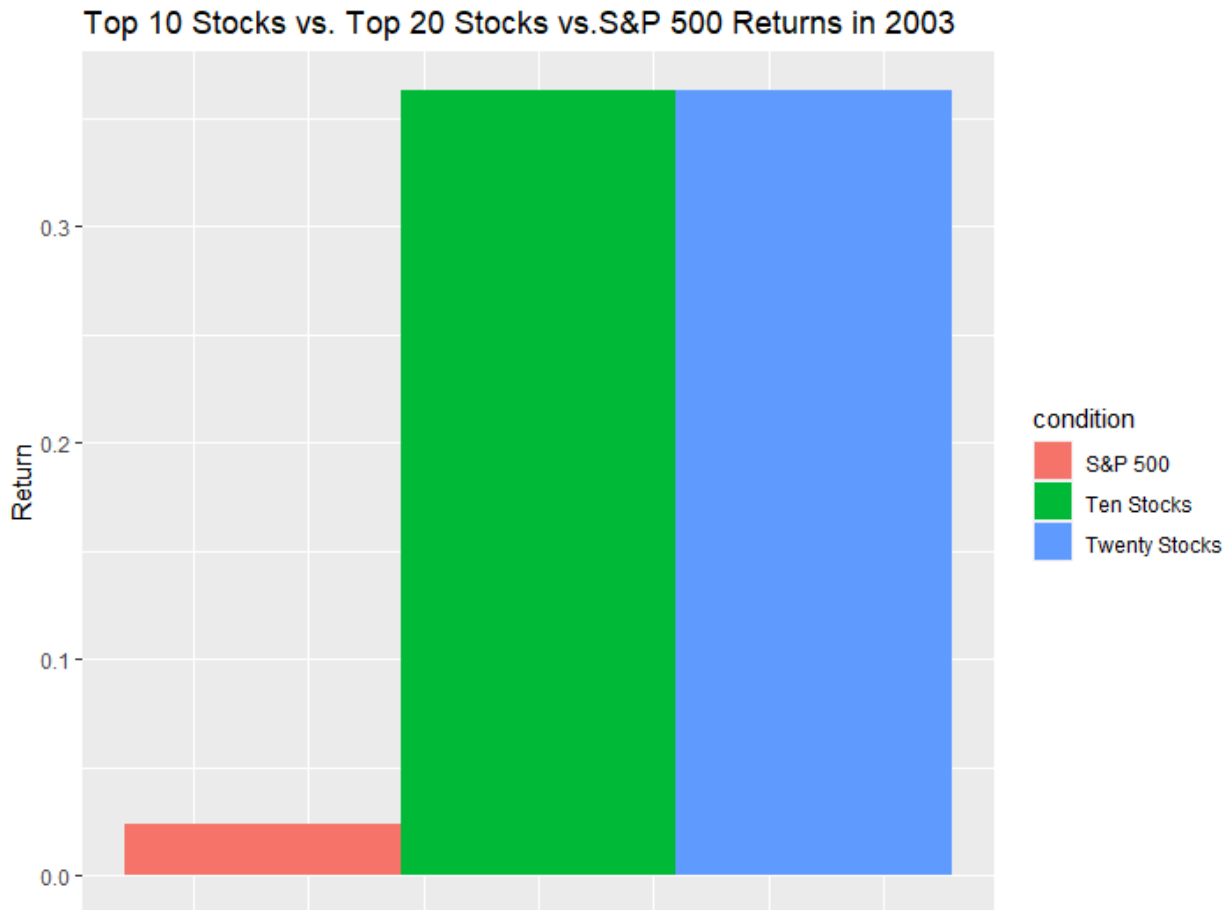


The demonstrated result is significantly higher than the S&P 500 - a result of market momentum. In the "good old days", the momentum wasn't correctly accounted for when people price the equities as we see in the 90s and early 2000s some returns skyrocketed to over 50%. However, now as people are developing a better understanding of the quantitative aspects of the market, the Max-Median rule doesn't work that great as they used to do.

The R code for this simulation is attached in the appendix.

Potential Improvements for Max-Median

One way of potential improvement we think that would work is actually decreasing the count of the stocks from 20 to 10 - we are overweighting the better performance stocks eventually to be the final cherry-picker of S&P 500. However, for some years the returns are similar or even lower like the year of 2003:



Works Cited

"S&P 500®." S&P Dow Jones Indices. Accessed Feburary 20, 2023. https://www.spglobal.com/spdji/en/indices/equity/sp-500/#data.

"S&P U.S. Indices - S&P Global." Accessed Feburary 20, 2023. https://www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-us-indices.pdf.

Appendix
source code:

```
library(readxl)
library(tidyr)
library(stargazer)
library(grid)
library(gridExtra)
library(pdfetch)
library(xts)
library(texreg)
library(forecast)
library(vars)
library(pander)

dailyprice <- read.csv("C:/data/S&P 500 daily prices.csv")

date <- dailyprice$datedate

ratios = list()
tickers = list()

#RUNTIME: O(5429700) - 6:11
#go through every stock
for(i in 1:length(dailyprice$Ticker)){
  #check the year
  if(unlist(strsplit(date[i], "-"))[1] == "1985"){
    #calculate ratios
    if(dailyprice$datedate[i] < dailyprice$datedate[i+1]){
      ratios <- append(ratios, (dailyprice$DlyPrc[i+1]/dailyprice$DlyPrc[i]))
      tickers <- append(tickers, dailyprice$Ticker[i+1])
    }
    else{
      next
    }
  }
}

medians <- list()

#RUNTIME O(500-550) - 0:17
for(tick in unique(cbind(tickers))){
  min_index <- min(which(cbind(tickers) %in% tick))
  max_index <- max(which(cbind(tickers) %in% tick))
  medians <- append(medians, median(unlist(ratios[min_index:max_index])))
}

ticker_median <- cbind(unlist(tickers), unlist(medians))
ordered_tm <- ticker_median[order(ticker_median[,2],decreasing=TRUE),]
c(ordered_tm[,1][1:20], ordered_tm[,2][1:20])
```

```r
library("ggplot2")

# create a dataset
specie <- c(rep("sorgho" , 3) , rep("poacee" , 3) , rep("banana" , 3) ,
            rep("triticum" , 3) )
condition <- rep(c("normal" , "stress" , "Nitrogen") , 4)
dim(condition)
value <- abs(rnorm(12 , 0 , 15))
data <- data.frame(specie,condition,value)

# Grouped
ggplot(data, aes(fill=condition, y=value, x=specie)) +
  geom_bar(position="dodge", stat="identity")

year <- c(1980, 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990,
          1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001,
          2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012,
          2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021)

weighted_ret <- c(1.27218390804598, 7.46759825327511, 7.65925925925926,
                  1.53982300884956, 8.15151515151515, 2.68342857142857,
                  2.68769953051643, 5.4, 4.86153846153846, 7.32710280373832,
                  66.3, 14.9643010752688, 5.75636363636364, 1.78853801169591,
                  5.14533333333333, 9.17062801932367, 4.23214285714286,
                  2.32258064516129, 9.77076923076923, 4.21420118343195,
                  3.62385964912281, 5.44895591647332, 24.7594936708861,
                  36.2633333333333, 6.35163127912999, 72.3939393939394,
                  7.89067702552719, 9.52393272962484, 9.49875,
                  9.45701357466063, 2.4989235737352, 7.55851063829787,
                  5.80560782837787, 6.69847972972973, 1.00442922689976,
                  7.43163538873995, 30.4592233009709, 5.96755162241888,
                  4.67209507042254, 6.52971725331795, 1.01224092357783,
                  1.01749271137026)

sp500yearlydata <- c(0.025062583333333333, -0.003512833333333333,
                     0.018150333333333327, 0.017268166666666668, 0.00613375,
                     0.02396491666666667, 0.015193249999999998,
                     0.008065333333333336, 0.013532999999999998,
                     0.023568916666666665, -0.0013995,
                     0.02347033333333333, 0.006425,
                     0.008028916666666665, 0.0015516666666666673,
                     0.02710916666666667, 0.018000999999999996,
                     0.025402583333333336, 0.023479, 0.016917583333333337,
                     -0.006268750000000001, -0.008986916666666666,
                     -0.018624166666666664, 0.02172625, 0.008914,
                     0.004475166666666666, 0.012333000000000002, 0.004998,
                     -0.035293916666666675, 0.021627083333333335,
                     0.013263083333333333, 0.0024491666666666655,
                     0.01293975, 0.0240215, 0.010887583333333332,
                     0.001922666666666666, 0.009724833333333334, 0.0167585,
                     -0.002894416666666667, 0.023714500000000003, 0.01739075,
                     0.021560666666666667)

finalmat <- c()
sp500condition <- c()
for(i in 1:42){
  finalmat <- c(finalmat, weighted_ret[i]/100, sp500yearlydata[i])
  sp500condition <- c(sp500condition, "MaxMedian", "S&P 500")
}

sp500data <- data.frame(year, finalmat, sp500condition)

ggplot(sp500data, aes(fill=sp500condition, x=year, y=finalmat)) +
  geom_bar(position="dodge", stat="identity") +
  labs(x="Year", y="Return") + ggtitle("MaxMedian vs. S&P 500 Returns")
```