Herding Analysis in the Technology Industry

Eya MAHDHAOUI , Manon CASTANET, Aline RIVERA, Pedro José TRILLOS

2022-11-14

Introduction:

Herding in finance is often referred to as the phenomenon in which investors mimic the behavior of other investors; if others invest in a certain asset, they will invest too and vice versa. There are various reasons why this happens, which can be rational or rather irrational. Most of the times, herding occurs when people think certain investors have insight or privileged information on a certain security, and therefore act in the markets as they do looking to profit off of what they may know.

Herding is important to study because it has a real impact in the markets in which it takes place. Take for example the case of the 2008 mortgage crisis. When investors started to see the mortgage market bubble they withdrew their money from the banks to keep it safe, causing other investors to follow suit and also withdraw their money. Because of herding several banks went into bankruptcy. Another example that illustrates the impact of herding is the recent case of the GameStop stock. There was a movement in which investors tried to help the company to avoid going out of business and bought the stock for a low price, then, other investors mimicked this behavior and bought GameStop's stock, causing it to skyrocket in January of 2021. Whether it is being done rationally or irrationally, the effect of herding is not to be taken lightly.

In this project, we will use R to estimate if herding is taking place in the technology market by analyzing the closing stock prices of nine representative companies in this industry over the last 22 years. For this purpose we will be following Chang et al.'s (2000) approach. This means, we will be making a linear regression of the cross-sectional absolute deviation at time t (CSAD) in terms of the market return. The coefficients, their signs, and their significance will indicate whether or not herding can be identified in the observations we have gathered for the technology market from the year 2000 to 2022. We will also be estimating herding according to up/down market days, that is, when the returns of the market are positive and negative respectively. For this, a dummy variable of positive vs negative market returns will be created and the regression model will be extended.

Methodology and Findings:

First, we install in R packages that will be required to carry out this analysis.

The data that will be used can now be uploaded. Closing stock prices of Apple, Microsoft, Oracle, Intel, Motorola, Adobe, Texas Instruments, Sony, and Logitech from November, 2000 to November, 2022 have been retrieved from Yahoo Finance and arranged into a dataset. We have chosen these companies because they are some of the biggest names in the industry that may satisfy the approach's assumption that each of these companies have the same weight in the technology industry. The data is then transformed in R into a time-series object.

```
library (xts)

## Le chargement a nécessité le package : zoo

##
## Attachement du package : 'zoo'

## Les objets suivants sont masqués depuis 'package:base':
##
##     as.Date, as.Date.numeric

library (zoo)
library (PerformanceAnalytics)

##
## Attachement du package : 'PerformanceAnalytics'

## L'objet suivant est masqué depuis 'package:graphics':
##
##     legend

require (sandwich)

## Le chargement a nécessité le package : sandwich

require(lmtest)

## Le chargement a nécessité le package : lmtest

# Upload and arrange the data

techprices = read.csv("HerdingData.csv", header = TRUE) #importing the data
techprices$Date = as.POSIXct(techprices$Date,format="%d/%m/%Y", tz = "") #
converting the first column into date format
techprices.xts <- xts(techprices[,-1], order.by=techprices[,1]) # convert the
object into xts object (time series object)
techprices.zoo  <- as.zoo(techprices.xts) # convert the time series object
into zoo object
```

It is possible now to calculate each stocks return in reference to time. Return is the first logarithmic difference between a price at time 0 and a price at time -1. If the result is positive, it means the security has generated a profit in the established time period. If the result is negative, it means the security has generated a loss in the established time period. This is calculated automatically by R. It is seen that returns constantly vary in signs for our observations. This is because companies tend to switch from periods of growth and decay in terms of their performance as time advances; that being said, some companies are more able to sustain periods of growth (which reflect in positive returns) than others.

```
## Calculating the returns of each stock

return = Return.calculate(techprices.xts , method = "log") # automatically
calculate return
View(return)
```

A CSAD function must now be created. This function calculates the absolute deviations of each stock's returns from the market return at a given time, which will be helpful to estimate herding as explained further on. For this, the creation of the mean return of the companies (or market return) for each period of time must be obtained (Rm). We must then obtain the sum of the absolute values of the difference of the market return and each of the companies' individual returns for each period of time. Finally, we ask R to count the number of dates we have in our dataset to complete the needed information to create this function.

CSAD (m,t) = Σ (|R(i,t)-R(m,t)|) / n

Where:

CSAD(m,t) = market cross-sectional absolute deviation at time t

R(i,t)= return of company i in time t

R(m,t) = market return in time t

n = number of dates

```
# a function to create CSAD and Rm
csadfunction = function(return)
{
  n=ncol(return)
  Rm = rowMeans(return)
  temp_dif =abs(return-Rm)
  temp_sum = rowSums(temp_dif)
  CSAD = temp_sum / ncol(return)
  CSAD = cbind (CSAD, Rm)
  return (CSAD)
}
```

We now have obtained the market returns and cross-sectional absolute deviations for each date in our dataset.

```
f = csadfunction(return) # calling the function "exchange.herd" that
calculates CSAD and Rm
head (f) # show the first 6 rows

##              CSAD           Rm
## [1,]           NA           NA
```

```
## [2,] 0.02930458  0.070823023
## [3,] 0.02152117  0.009712961
## [4,] 0.02342477 -0.022328280
## [5,] 0.02616847 -0.003183105
## [6,] 0.03674829 -0.035853605
```

A data frame with these two new variables is created to simplify the calculations that will be performed next. The Chang et al. (2000) approach suggests, however, that we should not only take into consideration Rm for the construction of the model, but also this variable squared in order to properly be able to analyze herding as it will account for non-linearity that may be present between the deviations and the market return. Therefore, we also create this variable (Rm2). Since returns are calculated using the first logarithmic difference, one missing values row will be present in our dataset. We can simply remove this empty row; we now have one observation less than what we started with but since we have a relatively large number of observations (5533) this should not be a problem that interferes with our results. We now have the three required variables in our data frame to construct the model. For simplicity in writing the model in R, we name CSAD (our dependent variable) "y", |Rm| "x1", and Rm2 "x2" (our independent variables).

```
CSAD.df = fortify.zoo(f) # converting f into a dataframe (to simplify further calculations)
CSAD.df$Rm2 = CSAD.df$Rm^2 # calculating Rm^2
CSAD.df = CSAD.df[-c(1),] # removing the first row with NAs
head (CSAD.df) # show the first 6 rows

##   Index       CSAD           Rm          Rm2
## 2     2 0.02930458  0.070823023 5.015901e-03
## 3     3 0.02152117  0.009712961 9.434162e-05
## 4     4 0.02342477 -0.022328280 4.985521e-04
## 5     5 0.02616847 -0.003183105 1.013216e-05
## 6     6 0.03674829 -0.035853605 1.285481e-03
## 7     7 0.02534601 -0.005932503 3.519459e-05

tail (CSAD.df) # show the last 6 rows

##      Index        CSAD           Rm          Rm2
## 5529  5529 0.016584648 -0.017207068 2.960832e-04
## 5530  5530 0.018130115  0.029834619 8.901045e-04
## 5531  5531 0.013553750  0.019204314 3.688057e-04
## 5532  5532 0.007380406  0.006947571 4.826874e-05
## 5533  5533 0.007742778 -0.017791851 3.165500e-04
## 5534  5534 0.021739193  0.065612087 4.304946e-03

y = CSAD.df$CSAD  # reassign  columns as Y and Xs to look better in the regression model
x1 = abs (CSAD.df$Rm)
x2 = CSAD.df$Rm2
```

It is possible to proceed to the creation of the CSAD linear model for the technology market. R indicates it is built as follows.

CSAD (m,t)= 0,008053+0,236311(|R(m,t)|)+0,064886(R^2(m,t))+e(t)

As seen in the model summary, both coefficients are positive but the coefficient of |Rm| is highly significant while the coefficient of Rm2 is not at all. It is possible to interpret from this that herding is not present in the technology market referent to the company information (and time period) we have gathered. To understand this, we must focus on the coefficient of Rm2. As Rm2 is a non-linear variable, its coefficient significance indicates whether CSAD should be measured through a non-linear regression; if it is significant it means it should, if it is not significant it means it should not. Therefore, if Rm2's coefficient is significant it means CSAD would be smaller (if it is negative) when market return moves significantly up or down, which would mean that there is herding present because herding lowers cross-sectional deviation of returns compared to rational pricing. As it is not significant in this case, we can infer that the investors operate in the market more rationally without provoking herding as CSAD follows a rather linear function of the market returns. In other words, Chang et al. (2000) approach to herding analysis leads us to believe investors in the technology market are unaffected by what other similar investors do before them; they make their own investment decisions according to the information they possess.

```
#Linear model
csad_model <- lm(y~x1+x2)  # build linear regression model on full data
print(csad_model)

##
## Call:
## lm(formula = y ~ x1 + x2)
##
## Coefficients:
## (Intercept)           x1           x2
##    0.008053     0.236311     0.064886

summary(csad_model)

##
## Call:
## lm(formula = y ~ x1 + x2)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.021272 -0.003569 -0.001223  0.002153  0.050242
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0080530  0.0001294  62.227   <2e-16 ***
## x1          0.2363110  0.0129256  18.282   <2e-16 ***
## x2          0.0648860  0.2010573   0.323    0.747
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.005797 on 5530 degrees of freedom
```

```
## Multiple R-squared:  0.1978, Adjusted R-squared:  0.1975
## F-statistic: 681.9 on 2 and 5530 DF,  p-value: < 2.2e-16
```

Now, herding in up/down technology market days shall be analyzed. This is done using an extended version of the previous CSAD model whose coefficients would demonstrate whether there is presence of herding either in days of positive market returns or in days of negative market returns. For this, we must first create a dichotomous variable that indicates whether the technology market returns at a certain date were positive (1) or negative (0).

```
# UP/DOWN Days
dup <- CSAD.df$Rm
dup[dup<0] <- 0
dup[dup>0] <- 1
```

The model for analyzing herding on up/down days requires the creation of four new variables; the dummy variable multiplied by the market return, one minus the dummy variable multiplied by the market return, and the same but with the market return squared. This time, we will focus on B3 and B4 as they are the coefficients of the squared variables.

```
x1_updown = dup*x1
x2_updown = ((1-dup)*x1)
x3_updown = dup*x2
x4_updown = ((1-dup)*x2)
```

The model is estimated as follows.

$CSAD(m,t) = 0,008071 + 0,221930(D(up)|R(m,t)|)+ 0,246751((1-D(up))|R(m,t)|)+0,129153(D(up)\ R^{(m,t))+0,059050((1-D(up))R}2(m,t))+e(t)$

We see again that all the coefficients are positive and that the coefficients related to the squared market returns are insignificant while the ones related to the market returns are highly significant. We have therefore found no evidence that indicates that there is herding happening in neither days of positive market returns nor in days of negative market returns. In other words, investors do not financially follow other investors neither when the company is generating profits for them nor when they are generating losses for them. This is again because the positive and insignificant coefficients of and indicate that CSAD does not move as much as Rm when Rm suffers great changes. If herding was present on up days the coefficient of would have been negative and significant, while if herding was present on down days we would have had a negative and significant coefficient for .

```
csad_model_extended <- lm(y~x1_updown+x2_updown+x3_updown+x4_updown)
print(csad_model_extended)

##
## Call:
## lm(formula = y ~ x1_updown + x2_updown + x3_updown + x4_updown)
##
## Coefficients:
```

```
## (Intercept)    x1_updown    x2_updown    x3_updown    x4_updown
##    0.008071     0.221930     0.246751     0.129153     0.059050
```

```
summary(csad_model_extended)
```

```
##
## Call:
## lm(formula = y ~ x1_updown + x2_updown + x3_updown + x4_updown)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.022331 -0.003573 -0.001223  0.002157  0.050000
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0080713  0.0001297  62.238   <2e-16 ***
## x1_updown   0.2219299  0.0151504  14.648   <2e-16 ***
## x2_updown   0.2467506  0.0157802  15.637   <2e-16 ***
## x3_updown   0.1291533  0.2421643   0.533    0.594
## x4_updown   0.0590500  0.2886166   0.205    0.838
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.005795 on 5528 degrees of freedom
## Multiple R-squared:  0.1986, Adjusted R-squared:  0.198
## F-statistic: 342.5 on 4 and 5528 DF,  p-value: < 2.2e-16
```

Conclusion:

Following Chang et al.'s (2000) approach to study herding within a market, it has been found that no herding behavior was present in the technology market from November of 2000 to November of 2022. It has also been found that this is true regardless of whether the market is experiencing an up (predominant positive returns) or down (predominant negative returns) day. Because the technology industry is a rather uncertain one (it is difficult to predict whether a technological innovation will bring profits or not), people might think investors would tend to follow the behavior of others who they believe could have some sort of insight. However, the findings of this project indicate that investors in this market act on their own knowledge and estimations without being biased or influenced by what other investors in the market are doing.