Panic on Wall Street

Introduction to behavioral finance

Anton Antonov
`https://github.com/tonytonov/talks`

2019-03-05

dxFeed

Financial markets

- Many asset classes

Financial markets

- Many asset classes
- Many exchanges

Financial markets

- Many asset classes
- Many exchanges
- Significant impact on daily lives

Financial markets

- Many asset classes
- Many exchanges
- Significant impact on daily lives
- Huge experimental field, *but*

Financial markets

- Many asset classes
- Many exchanges
- Significant impact on daily lives
- Huge experimental field, *but*
- No control, no repeatability

EMH statement

Definition (Fama, 1970)

Financial market is *efficient*, if security prices always fully reflect available information.

EMH statement

Definition (Fama, 1970)

Financial market is *efficient*, if security prices always fully reflect available information.

Definition (Efficient Market Hypothesis)

Real-world financial markets *are* efficient.

Evidence supporing EMH

- Theoretical
  - Investors are rational

Evidence supporing EMH

- Theoretical
  - Investors are rational
  - Prices are random walks (i.e., unpredictable)

Evidence supporing EMH

- Theoretical
  - Investors are rational
  - Prices are random walks (i.e., unpredictable)
  - Irrational investors are eliminated by arbitrageurs

Evidence supporing EMH

- Theoretical
  - Investors are rational
  - Prices are random walks (i.e., unpredictable)
  - Irrational investors are eliminated by arbitrageurs
- Empirical
  - *(Reaction to information)*
    News arrive $\Rightarrow$ price quickly and correctly adjusts

Evidence supporing EMH

- Theoretical
  - Investors are rational
  - Prices are random walks (i.e., unpredictable)
  - Irrational investors are eliminated by arbitrageurs
- Empirical
  - *(Reaction to information)*
    News arrive $\Rightarrow$ price quickly and correctly adjusts
  - *(Non-reaction to non-information)*
    No news about fundamentals $\Rightarrow$ no significant price movements

Evidence against EMH

- *"Limited rationality"* (discussed further)

Evidence against EMH

- *"Limited rationality"* (discussed further)
- *"Physical reality":*
  - Arbitrage opportunities are limited

Evidence against EMH

- *"Limited rationality"* (discussed further)
- *"Physical reality":*
  - Arbitrage opportunities are limited
  - Excess volatility puzzle

Evidence against EMH

- *"Limited rationality"* (discussed further)
- *"Physical reality":*
  - Arbitrage opportunities are limited
  - Excess volatility puzzle
  - Momentum and other factors are consistent predictors

Evidence against EMH

- *"Limited rationality"* (discussed further)
- *"Physical reality":*
  - Arbitrage opportunities are limited
  - Excess volatility puzzle
  - Momentum and other factors are consistent predictors
  - Flash Crash (reaction to non-news)

Flash Crash

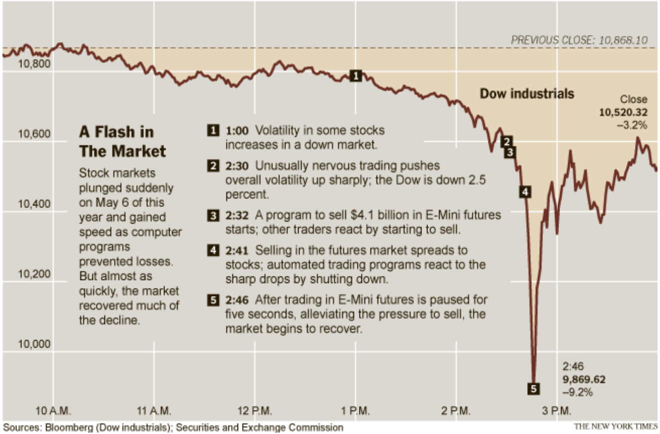


Figure 1: A trillion-dollar drop on May 6, 2010

Effect: noise trading

- Trading agents tend to
  - look for patterns in random data

Effect: noise trading

- Trading agents tend to
  - look for patterns in random data
  - ignore survivorship bias

Effect: noise trading

- Trading agents tend to
  - look for patterns in random data
  - ignore survivorship bias
  - systematically fail in absorbing new information

Effect: herding

- Traders, portfolio managers and algorithms
  - tend to mimic each other

Effect: herding

- Traders, portfolio managers and algorithms
  - tend to mimic each other
  - react on what others around them do

Effect: herding

- Traders, portfolio managers and algorithms
  - tend to mimic each other
  - react on what others around them do
  - often repeat each other's mistakes

Effect: herding

- Traders, portfolio managers and algorithms
  - tend to mimic each other
  - react on what others around them do
  - often repeat each other's mistakes
- During stress periods
  - correlations increase sharply

Effect: herding

- Traders, portfolio managers and algorithms
  - tend to mimic each other
  - react on what others around them do
  - often repeat each other's mistakes
- During stress periods
  - correlations increase sharply
  - event cascades are triggered

Effect: herding

- Traders, portfolio managers and algorithms
  - tend to mimic each other
  - react on what others around them do
  - often repeat each other's mistakes
- During stress periods
  - correlations increase sharply
  - event cascades are triggered
  - nonlinearity kicks in

Other effects

- Win-loss asymmetry, influence of framing

Other effects

- Win-loss asymmetry, influence of framing
- Distortion by "gurus", portfolio managers

Other effects

- Win-loss asymmetry, influence of framing
- Distortion by "gurus", portfolio managers

Figure 2: The Economist, November 1997

Model building approach

John von Neumann

The sciences do not try to explain, they hardly even try to interpret, they mainly make models. [...]

Model building approach

John von Neumann

The sciences do not try to explain, they hardly even try to interpret, they mainly make models. [...]

- *Observe* existing phenomena

Model building approach

John von Neumann

The sciences do not try to explain, they hardly even try to interpret, they mainly make models. [...]

- *Observe* existing phenomena
- *Construct* a model using domain knowledge

Model building approach

John von Neumann

The sciences do not try to explain, they hardly even try to interpret, they mainly make models. [...]

- *Observe* existing phenomena
- *Construct* a model using domain knowledge
- *Derive* quantifiable conclusions using mathematics

Model building approach

John von Neumann

The sciences do not try to explain, they hardly even try to interpret, they mainly make models. [...]

- *Observe* existing phenomena
- *Construct* a model using domain knowledge
- *Derive* quantifiable conclusions using mathematics
- *Test* conclusions on real (simulated) data

Coin flipping

- "Unfair" coin:
  $P(Heads) = \theta,\ P(Tails) = 1 - \theta$

Coin flipping

- "Unfair" coin:
  $P(Heads) = \theta,\ P(Tails) = 1 - \theta$
- Coin flips are *independent and identically distributed*, then e.g.

$$P(HHTHT) = P(H)^3 P(T)^2 = \\ = \theta^3 (1-\theta)^2$$

Coin flipping

- "Unfair" coin:
  $P(Heads) = \theta,\ P(Tails) = 1 - \theta$
- Coin flips are *independent and identically distributed*, then e.g.

$$P(HHTHT) = P(H)^3 P(T)^2 = \\ = \theta^3 (1-\theta)^2$$

- Given a set D of K_H heads and K_T tails

$$P(D|\theta) = \theta^{K_H} (1-\theta)^{K_T}$$

MLE

- By observing data D, how we can estimate unknown parameter θ?

MLE

- By observing data D, how we can estimate unknown parameter θ?
- Idea: pick θ, s.t. the probability of observing sample D is as high as possible

MLE

- By observing data D, how we can estimate unknown parameter θ?
- Idea: pick θ, s.t. the probability of observing sample D is as high as possible
- This is *maximum likelihood estimation (MLE):*

$$\hat{\theta}_{MLE} = \underset{\theta}{\operatorname{argmax}}\, P(D|\theta) = \underset{\theta}{\operatorname{argmax}} \ln P(D|\theta)$$

Problem with MLE

- Calculation yields (derive for bonus points!)

$$\hat{\theta}_{MLE} = \frac{K_H}{K_H + K_T}$$

Problem with MLE

- Calculation yields (derive for bonus points!)

$$\hat{\theta}_{MLE} = \frac{K_H}{K_H + K_T}$$

- Take $D = \{3\,Heads, 2\,Tails\}$, then $\hat{\theta} = 60\%$

Problem with MLE

- Calculation yields (derive for bonus points!)

$$\hat{\theta}_{MLE} = \frac{K_H}{K_H + K_T}$$

- Take $D = \{3\,Heads, 2\,Tails\}$, then $\hat{\theta} = 60\%$ — makes sense!

Problem with MLE

- Calculation yields (derive for bonus points!)

$$\hat{\theta}_{MLE} = \frac{K_H}{K_H + K_T}$$

- Take $D = \{3\,Heads, 2\,Tails\}$, then $\hat{\theta} = 60\%$ — makes sense!
- Take $D = \{5\,Heads\}$, then $\hat{\theta} = 100\%$

Problem with MLE

- Calculation yields (derive for bonus points!)

$$\hat{\theta}_{MLE} = \frac{K_H}{K_H + K_T}$$

- Take $D = \{3\,Heads, 2\,Tails\}$, then $\hat{\theta} = 60\%$ — makes sense!
- Take $D = \{5\,Heads\}$, then $\hat{\theta} = 100\%$ — a coin with two heads...!?

Coin tossing experiment

- Go to **bit.ly/pgi-math** to follow on-line!

Coin tossing experiment

- Go to **bit.ly/pgi-math** to follow on-line!
- The first coin is *heads-oriented:* H in 70% of outcomes, T in 30%. The second one is *tails-oriented:* 70% T, 30% H. I pick one randomly, the remaining coin is removed. At each step I'm telling you how the coin landed. Your task is to *estimate the probability of my coin being heads-oriented* as percentage from 0% to 100%. Let's go!

Bayes to the rescue!

- Reverend Thomas Bayes, 1702 – 1761:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

Bayes to the rescue!

- Reverend Thomas Bayes, 1702–1761:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

- Here $|$ means "given", "conditioned by": a probabilistic way to express *known information*

Bayes to the rescue!

- Reverend Thomas Bayes, 1702–1761:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}$$

- Here $|$ means "given", "conditioned by": a probabilistic way to express *known information*
- This is *maximum a posteriori* estimation:

$$\hat{\theta}_{MAP} = \operatorname*{argmax}_{\theta} P(\theta|D) = \operatorname*{argmax}_{\theta} P(D|\theta)P(\theta)$$

Bayes to the rescue!

- Reverend Thomas Bayes, 1702 – 1761:

$$P(H^*|Data) = \frac{P(Data|H^*)P(H^*)}{P(Data)}$$

Bayes to the rescue!

- Reverend Thomas Bayes, 1702 – 1761:

$$P(H^*|Data) = \frac{P(Data|H^*)P(H^*)}{P(Data)}$$

- Our case: two mutually exclusive events: H^*(head-oriented coin was chosen) and T^*(tail-oriented was chosen). $Data$ is what we observe: $Data = \{K_H\, Heads, K_T\, Tails\}$

Bayes to the rescue!

- Reverend Thomas Bayes, 1702 – 1761:

$$P(H^*|Data) = \frac{P(Data|H^*)P(H^*)}{P(Data)}$$

Bayes to the rescue!

- Reverend Thomas Bayes, 1702–1761:

$$P(H^*|Data) = \frac{P(Data|H^*)P(H^*)}{P(Data)}$$

$$P(Data) = P(Data|H^*)P(H^*) + P(Data|T^*)P(T^*) =$$
$$= 0.7^{K_H} 0.3^{K_T} 0.5 + 0.7^{K_H} 0.3^{K_T} 0.5$$

Bayes to the rescue!

- Reverend Thomas Bayes, 1702–1761:

$$P(H^*|Data) = \frac{P(Data|H^*)P(H^*)}{P(Data)}$$

Bayes to the rescue!

- Reverend Thomas Bayes, 1702 – 1761:

$$P(H^*|Data) = \frac{P(Data|H^*)P(H^*)}{P(Data)}$$

- Count K_H and K_T from $Data$, and $(\cancel{0.5})$

$$P(H^*|Data) = \frac{0.7^{K_H}0.3^{K_T}}{0.7^{K_H}0.3^{K_T} + 0.7^{K_H}0.3^{K_T}}$$

Results in theory

Roll	Outcome	History	Estimate
0			50%
1	H	H	70%
2	H	HH	84.5%
3	H	HHH	92.7%
4	H	HHHH	96.7%
5	H	HHHHH	98.6%
6	T	HHHHHT	96.7%

Table 1: Coin-tossing experiment, bayesian answers

Results from studies

- Edwards (1968): excess conservatism, *underreaction*
- Kahneman, Tversky (1974): representativeness heuristic, *overreaction*

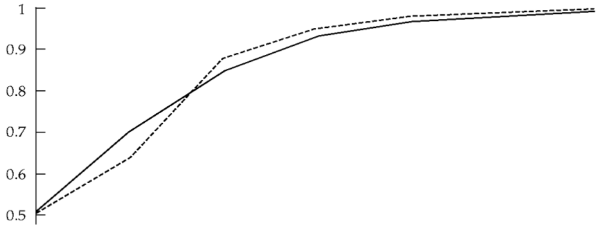
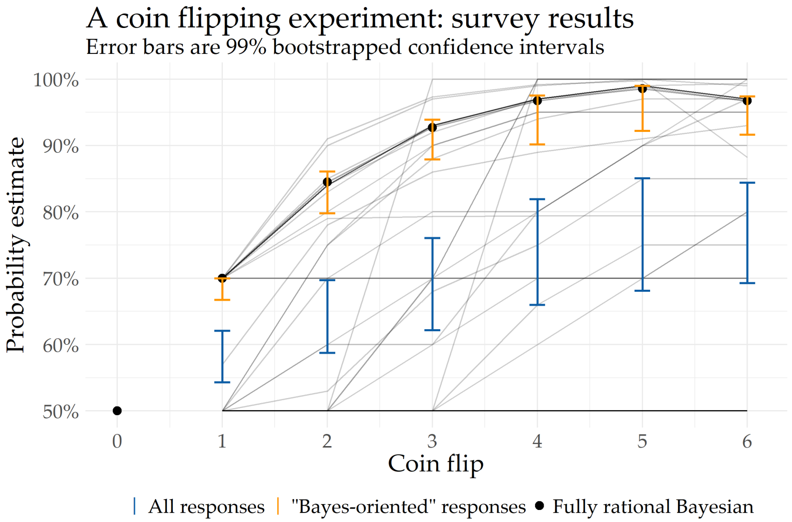


Figure 3: Bayesian (solid) and average response (dashed). Source: [Shleifer, 2000]

Our results

- Total respondents: 43
- Group 1: no reaction to information (all answers 50%)
- Group 2: non-Bayesian after one throw (all answers 70%)
- Group 3: Bayesian congregation
- Underreaction, no overreaction

Our results


A coin flipping experiment: survey results
Error bars are 99% bootstrapped confidence intervals
Probability estimate
100%
90%
80%
70%
60%
50%
0
1
2
3
4
5
6
Coin flip
All responses
"Bayes-oriented" responses
Fully rational Bayesian

Barberis et al. (1998)

- Featured in *Inefficient Markets* [Shleifer, 2000]

Barberis et al. (1998)

- Featured in *Inefficient Markets* [Shleifer, 2000]
- The stock price is driven by a random process $N_t = N_{t-1} + y_t$, where N_t is company's earnings, y_t is "shock"

Barberis et al. (1998)

- Featured in *Inefficient Markets* [Shleifer, 2000]
- The stock price is driven by a random process $N_t = N_{t-1} + y_t$, where N_t is company's earnings, y_t is "shock"
- Here y_t is either $+y$ or $-y$, i.e. positive or negative shock

Barberis et al. (1998)

- N_t is generated by a *regime-switching* model

Barberis et al. (1998)

- N_t is generated by a *regime-switching* model
- There are two states, M1 and M2, both are *Markov processes*

Barberis et al. (1998)

- N_t is generated by a *regime-switching* model
- There are two states, M1 and M2, both are *Markov processes*
- M1:

$$
\begin{array}{ccc}
 & y_{t+1} = y & y_{t+1} = -y \\
y_t = y & \pi_L & 1 - \pi_L \\
y_t = -y & 1 - \pi_L & \pi_L
\end{array}
$$

Barberis et al. (1998)

- N_t is generated by a *regime-switching* model
- There are two states, M1 and M2, both are *Markov processes*
- M1:

$$
\begin{array}{ccc}
 & y_{t+1} = y & y_{t+1} = -y \\
y_t = y & \pi_L & 1 - \pi_L \\
y_t = -y & 1 - \pi_L & \pi_L
\end{array}
$$

- M2 is similar, with π_H instead of π_L

Barberis et al. (1998)

- $\pi_L < 0.5 < \pi_H$, e.g. $\pi_L = 1/3$, $\pi_H = 3/4$

Barberis et al. (1998)

- $\pi_L < 0.5 < \pi_H$, e.g. $\pi_L = 1/3$, $\pi_H = 3/4$
- M1: "reversion", M2: "long memory"

Barberis et al. (1998)

- $\pi_L < 0.5 < \pi_H$, e.g. $\pi_L = 1/3$, $\pi_H = 3/4$
- M1: "reversion", M2: "long memory"
- Regime switching M1/M2 is also Markovian

$$\begin{array}{ccc} & s_{t+1} = 1 & s_{t+1} = 2 \\ s_t = 1 & 1 - \lambda_1 & \lambda_1 \\ s_t = 2 & \lambda_2 & 1 - \lambda_2 \end{array}$$

Barberis et al. (1998)

- $\pi_L < 0.5 < \pi_H$, e.g. $\pi_L = 1/3$, $\pi_H = 3/4$
- M1: "reversion", M2: "long memory"
- Regime switching M1/M2 is also Markovian

$$\begin{array}{ccc} & s_{t+1}=1 & s_{t+1}=2 \\ s_t=1 & 1-\lambda_1 & \lambda_1 \\ s_t=2 & \lambda_2 & 1-\lambda_2 \end{array}$$

- λ_1 and λ_2 are small, $\lambda_1 + \lambda_2 < 1$, $\lambda_1 < \lambda_2$

Barberis et al. (1998)

Proposition

Under some conditions on $\pi_L, \pi_H, \lambda_1, \lambda_2$, the price exhibits both under- and overreaction to N_t.

Barberis et al. (1998)

Proposition

Under some conditions on $\pi_L, \pi_H, \lambda_1, \lambda_2$, the price exhibits both under- and overreaction to N_t.

- Pick two portfolios: "winners" and "losers"

Barberis et al. (1998)

Proposition

Under some conditions on $\pi_L, \pi_H, \lambda_1, \lambda_2$, the price exhibits both under- and overreaction to N_t.

- Pick two portfolios: "winners" and "losers"
- Calculate difference in performance (returns) $r^n_+ - r^n_-$ for $n = 1, 2, 3, 4$ years

Barberis et al. (1998)

Proposition

Under some conditions on $\pi_L, \pi_H, \lambda_1, \lambda_2$, the price exhibits both under- and overreaction to N_t.

- Pick two portfolios: "winners" and "losers"
- Calculate difference in performance (returns) $r_+^n - r_-^n$ for $n = 1, 2, 3, 4$ years
- It *decreases monotonically*

Barberis et al. (1998)

- It *decreases monotonically:*

$r_+^1 - r_-^1$	0.0391
$r_+^2 - r_-^2$	0.0131
$r_+^3 - r_-^3$	−0.0072
$r_+^4 - r_-^4$	−0.0309

Figure 4: Difference between portfolios. Source: [Shleifer, 2000]

Reproducing in R

R code:
https://github.com/tonytonov/talks/blob/master/behfin_intro/R/shleifer.R

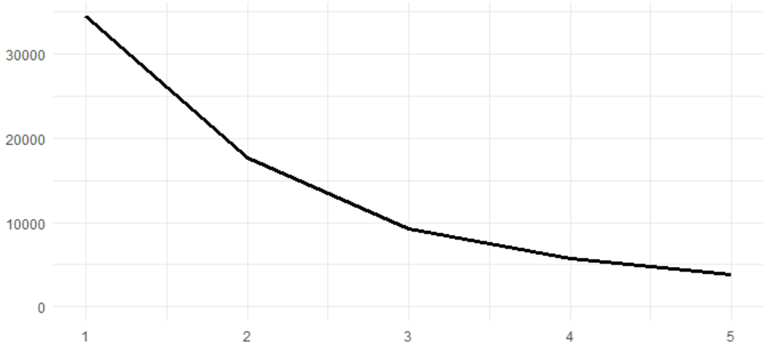


Figure 5: Difference between portfolios decays with n

Takeaway

- Efficient market hypothesis
- Behavioral effects in finance
- Model of investor sentiment [Barberis et al. (1998)]

Takeaway

- Efficient market hypothesis
- Behavioral effects in finance
- Model of investor sentiment [Barberis et al. (1998)]

Key takeaway

Don't panic and become a Bayesian!

Takeaway

- Efficient market hypothesis
- Behavioral effects in finance
- Model of investor sentiment [Barberis et al. (1998)]

Key takeaway

Don't panic and become a Bayesian!

Thanks! Questions?

References

- Sornette, D.; Andersen, J. V. (2002). *A Nonlinear Super-Exponential Rational Model of Speculative Financial Bubbles.* International Journal of Modern Physics C, Volume 13, Issue 02, pp. 171–187.
- Sornette, D. *Why stock markets crash.* Princeton University Press, 2003.
- Barberis, N.; Shleifer, A.; Vishny, N. (1998). *A model of investor sentiment.* Journal of Financial Economics 49, pp. 307–343.
- Shleifer, A. *Inefficient Markets.* Oxford University Press, 2000.
- Taleb, N.N. *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets.* Random House, 2001.