# Panic on Wall Street

*Introduction to behavioral finance*

Anton Antonov
Ph.D., CQF, Quant at dxFeed
tonytonov@gmail.com, @tonytonov

2019-02-07

dxFeed

# Financial markets

- Many asset classes

# Financial markets

- Many asset classes
- Many exchanges

# Financial markets

- Many asset classes
- Many exchanges
- Significant impact

# Financial markets

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

# Financial markets

- Many asset classes
- Many exchanges
- Significant impact
- 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 $\Longrightarrow$ 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 $\Longrightarrow$ price quickly and correctly adjusts
  - *(Non-reaction to non-information)*
    No news about fundamentals $\Longrightarrow$ 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

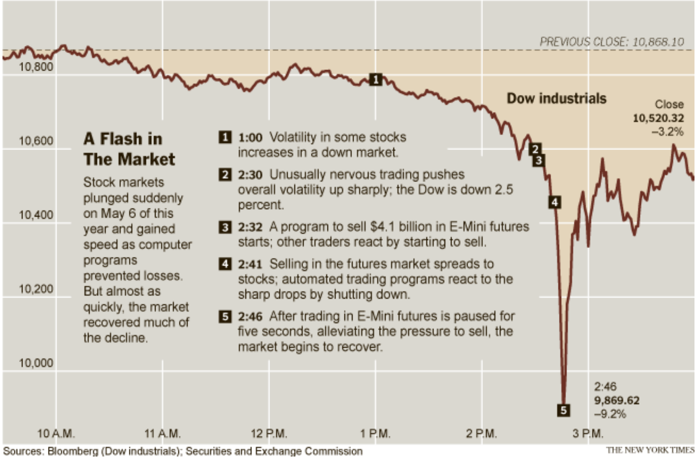



Figure 1: A trillion-dollar drop around 2:30 p.m. EST on May 6, 2010

# 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 financial gurus, portfolio managers

# Other effects

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


Figure 2: The Economist, November 1997

# Model building approach

- *Observe* existing phenomena

# Model building approach

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

# Model building approach

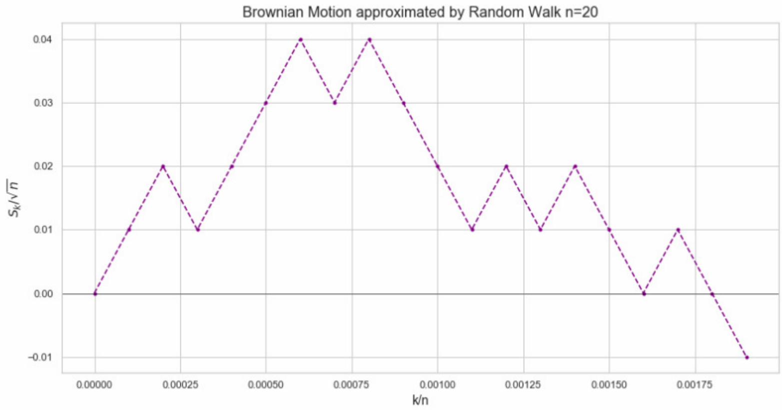
- *Observe* existing phenomena
- *Construct* a model using domain knowledge
- *Derive* quantifiable conclusions using mathematics

# Model building approach

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

# Random walk, brownian motion

- Random walk *converges* to brownian motion:



# Sornette, Andersen (2002)

- Option pricing (Black, Scholes, Merton (1973)):

$$\frac{dB}{B} = \mu dt + \sigma dW_t$$

# Sornette, Andersen (2002)

- Option pricing (Black, Scholes, Merton (1973)):

$$\frac{dB}{B} = \mu dt + \sigma dW_t$$

- Adding price jumps:

$$\frac{dB}{B} = \mu dt + \sigma dW_t - \kappa dj$$

# Sornette, Andersen (2002)

- *No-arbitrage* condition leads to

$$B(t) = \alpha^{\alpha} \frac{1}{\left(\mu_0(t_c - t) - \frac{\sigma_0}{B_0^m} W_t\right)^{\alpha}}$$

$$\alpha = \frac{1}{m-1}, t_c = \frac{y_0}{(m-1)\mu_0}$$

- Simplified version:

$$B(t) = \frac{\alpha^{\alpha}}{(1 - W_t)^{\alpha}}$$

# Original result

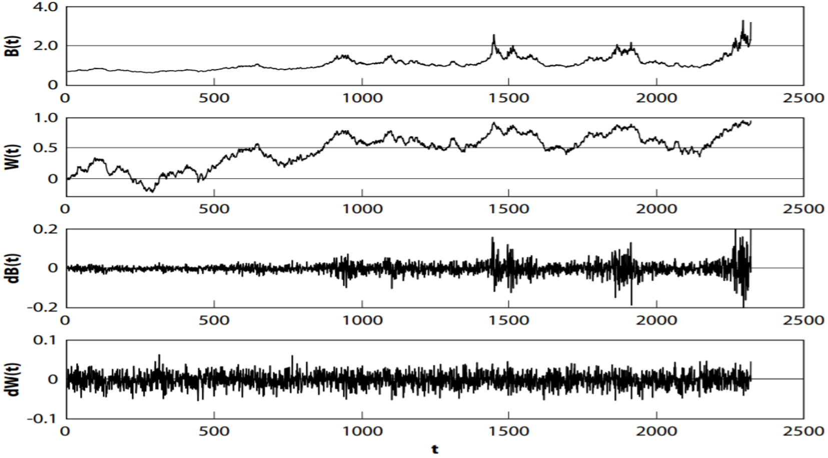



Figure 3: Bubble price and its' components. Source: [Sornette, Andersen (2002)]

# Reproducing in R

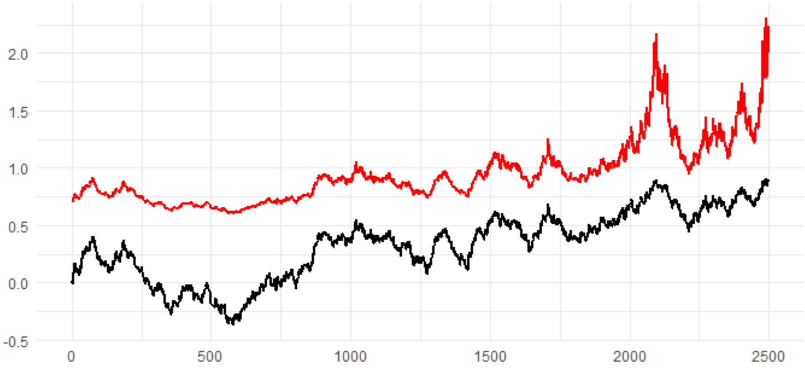



Figure 4: Brownian motion $W_t$ (black), price bubble $B(t)$ (red)

# Coin flipping

- Take a regular "unfair" coin: $P(Heads) = \theta$, $P(Tails) = 1 - \theta$

# Coin flipping

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

$$\begin{aligned} P(HHTHT) &= P(H)P(H)P(T)P(H)P(T) = \\ &= P(H)^3 P(T)^2 = \\ &= \theta^3 (1-\theta)^2 \end{aligned}$$

# Coin flipping

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

$$\begin{aligned} P(HHTHT) &= P(H)P(H)P(T)P(H)P(T) = \\ &= P(H)^3 P(T)^2 = \\ &= \theta^3 (1-\theta)^2 \end{aligned}$$

- In general, given a set $D$ of $\alpha_H$ heads and $\alpha_T$ tails

$$P(D|\theta) = \theta^{\alpha_H} (1-\theta)^{\alpha_T}$$

# MLE

- By observing data $D$, how we can estimate unknown parameter $\theta$?

# MLE

- By observing data $D$, how we can estimate unknown parameter $\theta$?
- Idea: pick $\theta$,so that the probability of observing $D$ is as high as possible

# MLE

- By observing data $D$, how we can estimate unknown parameter $\theta$?
- Idea: pick $\theta$,so that the probability of observing $D$ is as high as possible
- This is *maximum likelihood estimation (MLE):*

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

# MLE

- By observing data $D$, how we can estimate unknown parameter $\theta$?
- Idea: pick $\theta$,so that the probability of observing $D$ is as high as possible
- This is *maximum likelihood estimation (MLE):*

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

(whiteboard)

# Problem with MLE

- Calculation yields

$$\hat{\theta}_{MLE} = \frac{\alpha_H}{\alpha_H + \alpha_T}$$

# Problem with MLE

- Calculation yields

$$\hat{\theta}_{MLE} = \frac{\alpha_H}{\alpha_H + \alpha_T}$$

- Take $D = \{3\,Heads, 2\,Tails\}$, then $\hat{\theta} = 0.6$

# Problem with MLE

- Calculation yields

$$\hat{\theta}_{MLE} = \frac{\alpha_H}{\alpha_H + \alpha_T}$$

- Take $D = \{3\,Heads, 2\,Tails\}$, then $\hat{\theta} = 0.6$ — makes sense!

# Problem with MLE

- Calculation yields

$$\hat{\theta}_{MLE} = \frac{\alpha_H}{\alpha_H + \alpha_T}$$

- Take $D = \{3\,Heads, 2\,Tails\}$, then $\hat{\theta} = 0.6$ — makes sense!
- Take $D = \{5\,Heads\}$, then $\hat{\theta} = 1$

# Problem with MLE

- Calculation yields

$$\hat{\theta}_{MLE} = \frac{\alpha_H}{\alpha_H + \alpha_T}$$

- Take $D = \{3\,Heads, 2\,Tails\}$, then $\hat{\theta} = 0.6$ — makes sense!
- Take $D = \{5\,Heads\}$, then $\hat{\theta} = 1$ — a coin with two heads...?

# Bayes to the rescue!

- Formula (Reverend Thomas Bayes, 1702–1761):

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

# Bayes to the rescue!

- Formula (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!

- Formula (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 (MAP)* estimation:

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

(whiteboard)

# Bayesian estimates

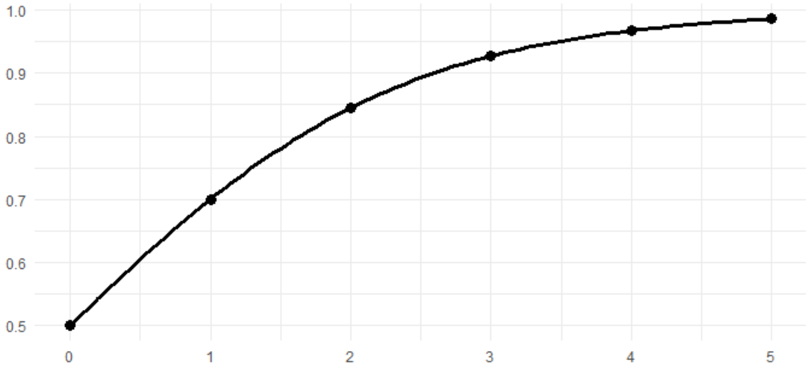



Figure 5: How a true Bayesian would answer

# Bayesian estimates

| 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

- Edwards (1968): excess conservatism, *underreaction*

# Bayesian estimates

| 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

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

# Background

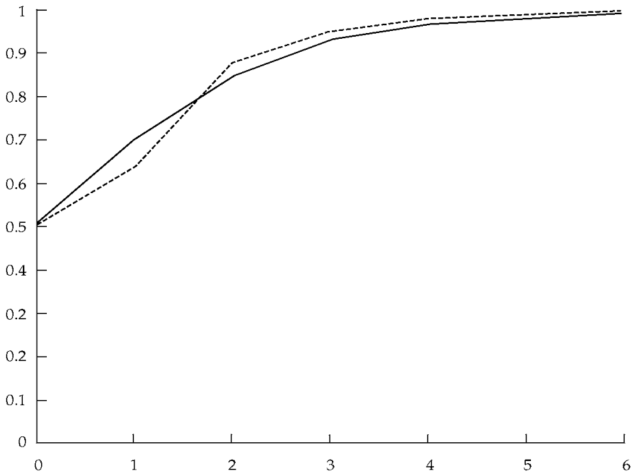
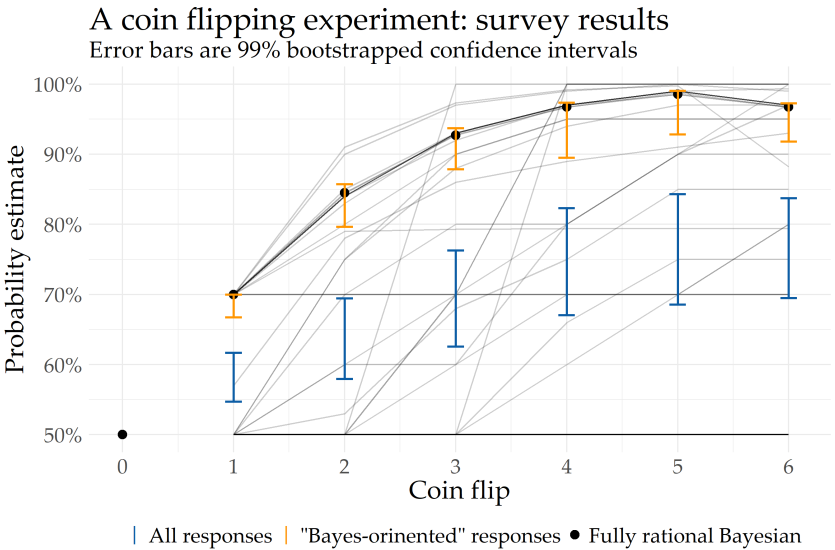



Figure 6: Bayesian (solid) and average response (dashed). Source: [Shleifer, 2000]

# Our results

- Total respondents: 43
- Group 1: non-Bayesian congregation (all answers $\leqslant 50\%$)
- Group 1a: non-Bayesian after first throw (all answers $= 70\%$)
- Group 2: MLE-infuenced (there is at least one 100% answer)
- Group 3: "I can feel it" and informed Bayesians

# 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-orinented" responses
Fully rational Bayesian

# Barberis et al. (1998)

- 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)

- 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 $\pi_H$ instead of $\pi_L$

# Barberis et al. (1998)

- $0 < \pi_L < 0.5$, $0.5 < \pi_H < 1$, e.g. $\pi_L = 1/3$, $\pi_H = 3/4$

## Barberis et al. (1998)

- $0 < \pi_L < 0.5$, $0.5 < \pi_H < 1$, e.g. $\pi_L = 1/3$, $\pi_H = 3/4$
- M1 is regime with "reversion", M2 — with "long memory"

# Barberis et al. (1998)

- $0 < \pi_L < 0.5$, $0.5 < \pi_H < 1$, e.g. $\pi_L = 1/3$, $\pi_H = 3/4$
- M1 is regime with "reversion", M2 — with "long memory"
- Regime switching between M1 and 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)

- $0 < \pi_L < 0.5$, $0.5 < \pi_H < 1$, e.g. $\pi_L = 1/3$, $\pi_H = 3/4$
- M1 is regime with "reversion", M2 — with "long memory"
- Regime switching between M1 and M2 is also Markovian

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

- $\lambda_1$ and $\lambda_2$ are small, $\lambda_1 + \lambda_2 < 1$, $\lambda_1 < \lambda_2$ (on average, M1 is more likely)

# 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 $r_+^n - r_-^n$ 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 $r_+^n - r_-^n$ for $n = 1, 2, 3, 4$ years
- 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 7: Difference between portfolios. Source: [Shleifer, 2000]

# Reproducing in R

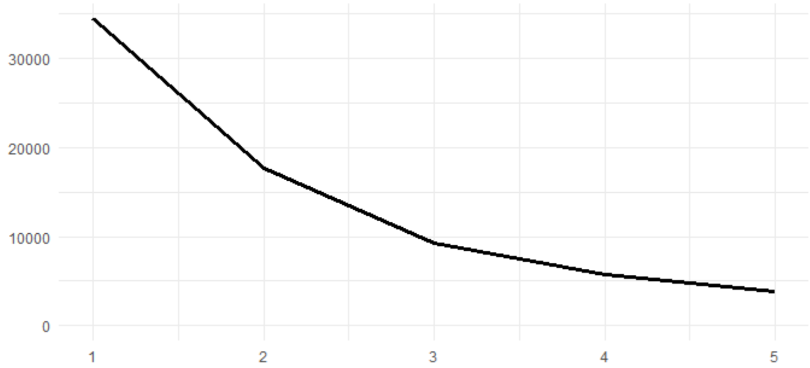



Figure 8: Difference between portfolios decays with $n$

# Takeaway

Key concepts and ideas:

- Efficient market hypothesis
- Behavioral effects in finance
- Price-driven crash model [Sornette, Andersen (2002)]
- Model of investor sentiment [Barberis et al. (1998)]

# Takeaway

Key concepts and ideas:

- Efficient market hypothesis
- Behavioral effects in finance
- Price-driven crash model [Sornette, Andersen (2002)]
- Model of investor sentiment [Barberis et al. (1998)]

## Key takeaway

Don't panic and become a Bayesian!

# Takeaway

Key concepts and ideas:

- Efficient market hypothesis
- Behavioral effects in finance
- Price-driven crash model [Sornette, Andersen (2002)]
- Model of investor sentiment [Barberis et al. (1998)]

## Key takeaway

Don't panic and become a Bayesian!

Thanks!

# 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.