Lecture 1: Bayes' Rule
Information Cascades

Behavioral Game Theory and Experiments
Jörg Oechssler
Fall 2013

Prof. Jörg Oechssler, Ph.D.

Behavioral Game Theory (and Experiments)
Fall 2013, E877

- **Literature:** Fudenberg, D. and D. Levine (1998), *The Theory of Learning in Games*, Cambridge, Mass.: MIT-Press; Friedman, D. and S. Sunder (1994), *Experimental Methods*, Cambridge University Press. Colin F. Camerer, *Behavioral Game Theory*, Princeton University Press, 2003.

- **Prerequisite:** A course in Game Theory.

Syllabus

I. Rational Learning

1. **Bayes' Rule**. *Lit.*: *Judge et al. (1988) *Introduction to the Theory and Practice of Econometrics*, S. 14-18; *Friedman, D. (1998), "Monty Hall's Three Doors: Construction and Deconstruction of a Choice Anomaly", *American Economic Review,* 88, 933-946.

2. **Information cascades**. *Lit.:* Banerjee, A. (1992), "A Simple Model of Herd Behavior", *Quarterly Journal of Economics*, 107, 797-817; *Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992), "A Theory of Fads, Fashion, Custom, and Cultural Changes as Informational Cascades", *Journal of Political Economy*, 100, 992-1026; *Anderson, L. and Holt, C. (1997), "Information Cascades in the Laboratory", *American Economic Review,* 87, 847-862; *Huck, S. and Oechssler, J. (2000) "Informational Cascades in the Laboratory: Do They Occur for the Right Reasons?", *Journal of Economic Psychology,* 661-671; Drehmann, M., J. Oechssler, and A. Roider, "Herding and Contrarian Behavior in Financial Markets - An Internet Experiment", *American Economic Review* 95(5) December (2005). Weizsäcker, G. (2010): Do We Follow Others When We Should? A Simple Test of Rational Expectations, American Economic Review, 100(5), 2340-2360. Avery, C. And P. Zemsky (1998): Multidimensional Uncertainty and Herd Behavior in Financial Markets, American Economic Review, 88(4), 724-748. Salganik,

M. J., Dodds, P. S., and D.J. Watts (2006): Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market, Science, 311, 854-856. Plassmann, H., O'Doherty, J., Shiv, B. and A. Rangel (2008): Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness, PNAS, 105(3), 1050-1054.

II. Boundedly Rational Learning

1. **Belief–based Learning**: Cournot Best—Reply Dynamics, Fictitious Play. *Lit.:* *Fudenberg/Levine, Chap. 2. *Cheung Y. and D. Friedman (1997), "Individual Learning in Games: Some Laboratory Results" *Games and Economic Behavior*, 19, 46-76. *Huck, S., H.-T. Normann and J. Oechssler, (2002), Stability of the Cournot Process - Experimental Evidence, *International Journal of Game Theory*; Theocharis, R. (1960), "On the Stability of the Cournot Solution of the Oligopoly Problem", *Review of Economic Studies*, 73, 133-134.

2. **Reinforcement Learning**. *Lit.:* *Roth, A. and I. Erev (1995), "Learning in Extensive Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term", *Games and Economic Behavior* 8, 164-212; Börgers, T. and R. Sarin (1997),"Learning through Reinforcement and Replicator Dynamics", *Journal of Economic Theory* 77, 1-14.

3. **Experience Weighted Attractions**. *Lit.:* *Camerer, C. and T. Ho (1999), "Experience-weighted Attraction Learning in Normal Form Games", *Econometrica*, 67, 827-874.

4. **Imitation and Stochastic Stability**. *Vega–Redondo, F. (1997), "The Evolution of Walrasian Behavior", *Econometrica,* 65, 375-384*;* *Huck, S., H.-T. Normann and J. Oechssler, (1999), "Learning in Cournot Oligopoly - An Experiment", *Economic Journal*, 109, C80-C95; Apesteguia, J., S. Huck, and J. Oechssler „Imitation – Theory and Experimental Evidence" *Journal of Economic Theory,* 2007*.*

5. **Quntal response equilibrium.** Fey, M., McKelvey, R. D., and Palfrey, T. R. (1996): An Experimental Study of Constant-Sum Centipede Games, *International Journal of Game Theory*, 25, 269-287. Goeree, J. K. and Holt, C. A. (2001): Ten Little Treasures of Game Theory and Ten Intuitive Contradictions, *American Economic Review*, 91(5), 1402-1422. McKelvey, R. D., and Palfrey, T. R. (1992): An Experimental Study of the Centipede Game, *Econometrica*, 60(4), 803-836.

Example: 3 Prisoners (Casella and Berger)

Three prisoner's, A, B, and C, are on death row. The governor decides to pardon one of the three and chooses at random the prisoner to pardon. He informs the warden of his choice but requests that the name be kept secret for a few days. The next day, A tries to get the warden to tell him who had been pardoned. The warden refuses. A then asks which of B or C will be executed. The warden thinks for a while, then tells A that B is to be executed.

- Warden's reasoning: Each prisoner has a $\frac{1}{3}$ chance of being pardoned. Clearly, either B or C must be executed, so I have given A no information about whether A will be pardoned.
- A's reasoning: Given that B will be executed, then either A or C will be pardoned. My chance of being pardoned has risen to $\frac{1}{2}$.
- Who is right?

Conditional Probabilities

Definition

The set, S, of all possible outcomes of a particular experiment is called the *sample space* for the experiment.

Example: coin toss: $S = \{H, T\}$

Definition

An *event* is any collection of possible outcomes of an experiment, that is, any subset of S (including S itself).

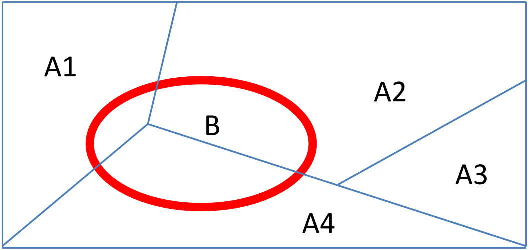
Definition

If A and B are events in S, and $P(B) > 0$, then the *conditional probability of A given B*, written $P(A|B)$, is
$P(A|B) = \frac{P(A \cap B)}{P(B)}$

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A) \qquad (*)$$

Definition

If $A_1, A_2, \ldots$ are pairwise disjoint and $\bigcup_{i=1}^{\infty} A_i = S$, then the collection $A_1, A_2, \ldots$ forms a *partition* of S.



$$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots$$

Since $(A_i \cap B)$ pairwise disjoint,

$$\begin{aligned} P(B) &= \sum_i P(A_i \cap B) \\ \text{by } (*) &= \sum P(B|A_i)P(A_i) \end{aligned}$$

Theorem

Bayes' Rule: Let $A_1, A_2, \ldots$ be a partition of the sample space, and let B be any set. Then, for each $i = 1, 2, \ldots,$

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}$$

- $P(A_i)$ is called the "prior"

Since $(A_i \cap B)$ pairwise disjoint,

$$\begin{aligned} P(B) &= \sum_i P(A_i \cap B) \\ \text{by } (*) &= \sum P(B|A_i)P(A_i) \end{aligned}$$

Theorem

Bayes' Rule: Let $A_1, A_2, \ldots$ be a partition of the sample space, and let B be any set. Then, for each $i = 1, 2, \ldots,$

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}$$

- $P(A_i)$ is called the "prior"
- $P(A_i|B)$ is called the "posterior"

- A, B, C the events that A, B, or C are pardoned, respectively

- A, B, C the events that A, B, or C are pardoned, respectively
- W is the event that the warden says B is executed

- A, B, C the events that A, B, or C are pardoned, respectively
- W is the event that the warden says B is executed
- By Bayes' rule:

$$P(A|W) = \frac{P(W|A)P(A)}{P(W|A)P(A) + P(W|B)P(B) + P(W|C)P(C)}$$

- A, B, C the events that A, B, or C are pardoned, respectively
- W is the event that the warden says B is executed
- By Bayes' rule:

$$P(A|W) = \frac{P(W|A)P(A)}{P(W|A)P(A) + P(W|B)P(B) + P(W|C)P(C)}$$

- $P(W|A) = 0.5$ (assumption!), $P(W|B) = 0, P(W|C) = 1$

- A, B, C the events that A, B, or C are pardoned, respectively
- W is the event that the warden says B is executed
- By Bayes' rule:

$$P(A|W) = \frac{P(W|A)P(A)}{P(W|A)P(A) + P(W|B)P(B) + P(W|C)P(C)}$$

- $P(W|A) = 0.5$ (assumption!), $P(W|B) = 0, P(W|C) = 1$
-

$$P(A|W) = \frac{0.5\frac{1}{3}}{0.5\frac{1}{3} + 0 + 1\frac{1}{3}} = \frac{1}{3}$$

- A, B, C the events that A, B, or C are pardoned, respectively
- W is the event that the warden says B is executed
- By Bayes' rule:

$$P(A|W) = \frac{P(W|A)P(A)}{P(W|A)P(A) + P(W|B)P(B) + P(W|C)P(C)}$$

- $P(W|A) = 0.5$ (assumption!), $P(W|B) = 0, P(W|C) = 1$
- $$P(A|W) = \frac{0.5\frac{1}{3}}{0.5\frac{1}{3} + 0 + 1\frac{1}{3}} = \frac{1}{3}$$
- Therefore, the warden was right since $P(A) = P(A|W)$

Base-rate Neglect 3

- Here's an even more extreme example: Subjects had to guess what the employment of a person called "Linda" was. Among the options were "Linda is a bank teller" and "Linda is a bank teller and active in the feminist movement"

> *Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student she was deeply concerned with issues of discrimination and social justice and also participated in antinuclear demonstrations.*

- 89% of the probability group thought that Linda is more likely to be a feminist bank teller than a bank teller (conjunction fallacy)

Base-rate neglect

- Many people ignore or under-weigh the prior probability when updating their beliefs

Example: Aids test

Exp. by Dan Friedman, AER 1998

2. Information cascades and herding

Literature Herding and Information Cascades

- Ash, S. E. (1955): Opinions and Social Pressure, Scientific American, 193(5), 31-35
- Bond, R. and P. B. Smith (1996): Culture and Conformity: A Meta-Analysis of Studies Using Asch's (1952b, 1956) Line Judgment Task, Psychological Bulletin, 119(1), 111-137
- Bikhchandani, S., Hirschleifer, D. and I. Welch (1992): A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades, Journal of Political Economy, 100 (5), 992-1026.
- Anderson, L. R. and C. A. Holt (1997): Information Cascades in the Laboratory, American Economic Review, 87, 847-62.
- Weizsäcker, G. (2010): Do We Follow Others When We Should? A Simple Test of Rational Expectations, American Economic Review, 100(5), 2340-2360.

Literature Herding and Information Cascades

- Avery, C. And P. Zemsky (1998): Multidimensional Uncertainty and Herd Behavior in Financial Markets, American Economic Review, 88(4), 724-748
- Drehmann, M., Oechssler, J., and A. Roider (2005), Herding and Contrarian Behavior in Financial Markets: An Internet Experiment, American Economic Review, 95(5), 1403-1426
- Salganik, M. J., Dodds, P. S., and D.J. Watts (2006): Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market, Science, 311, 854-856
- Plassmann, H., O'Doherty, J., Shiv, B. and A. Rangel (2008): Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness, PNAS, 105(3), 1050-1054

Exp. on Cascades

Introduction

- “One of the most striking regularities of human society is localized conformity. Americans act American, Germans act German, and Indians act Indian” (Bikchandani et al., 1992, p. 992)
  - For example, the way people dress changes across time as well as across cultures
  - The tech bubble is also often seen as evidence for herding/conformity
  - The distribution of the consumption of goods like music or films or also cars is highly unequal: some producers sell a lot, most sell almost nothing
  - Opinion polls are thought to influence voters such that they tend to vote in favor of the leading candidate

Introduction

- Possible explanations

Introduction

- Possible explanations
  - Similar preferences
    - One product might have superior features and as a result, consumers with similar preferences independently choose it
    - but: there is not always enough information available to make an informed decision (e.g., cars, computers, stocks...)
    - but: these decisions would then always be correct (counterexample: Beta vs. VHS video systems)

Introduction

- Possible explanations
  - Similar preferences
    - One product might have superior features and as a result, consumers with similar preferences independently choose it
    - but: there is not always enough information available to make an informed decision (e.g., cars, computers, stocks...)
    - but: these decisions would then always be correct (counterexample: Beta vs. VHS video systems)
  - Payoff externalities
    - Example: Computer operating systems, possibly also music/movies/books

Introduction

- Possible explanations
  - Similar preferences
    - One product might have superior features and as a result, consumers with similar preferences independently choose it
    - but: there is not always enough information available to make an informed decision (e.g., cars, computers, stocks...)
    - but: these decisions would then always be correct (counterexample: Beta vs. VHS video systems)
  - Payoff externalities
    - Example: Computer operating systems, possibly also music/movies/books
  - Sanctions on deviants
    - It's probably not a good idea to wear LA-style clothes in Riad...

Introduction

- Possible explanations
  - Similar preferences
    - One product might have superior features and as a result, consumers with similar preferences independently choose it
    - but: there is not always enough information available to make an informed decision (e.g., cars, computers, stocks...)
    - but: these decisions would then always be correct (counterexample: Beta vs. VHS video systems)
  - Payoff externalities
    - Example: Computer operating systems, possibly also music/movies/books
  - Sanctions on deviants
    - It's probably not a good idea to wear LA-style clothes in Riad...
  - Information cascades
    - Each individual has limited information about the quality of the product but can observe other consumers' choices

Introduction

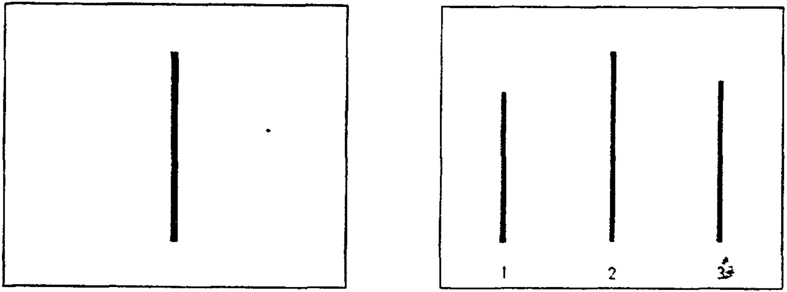
- Possible explanations
  - Similar preferences
    - One product might have superior features and as a result, consumers with similar preferences independently choose it
    - but: there is not always enough information available to make an informed decision (e.g., cars, computers, stocks...)
    - but: these decisions would then always be correct (counterexample: Beta vs. VHS video systems)
  - Payoff externalities
    - Example: Computer operating systems, possibly also music/movies/books
  - Sanctions on deviants
    - It's probably not a good idea to wear LA-style clothes in Riad...
  - Information cascades
    - Each individual has limited information about the quality of the product but can observe other consumers' choices
  - Preference for conformity

Ash Experiment: Purpose

- Ash addresses the question of "How, and to what extent, do social forces constrain people's opinions and attitudes?" (p .2)
- Experimental Design
  - 7-9 college students (all men)

Ash Experiment: Design

- The experimenter shows 2 white cards to the subjects
- Subjects are asked to indicate which one of the three lines on the card on the right has the same length as the line on the card on the left
- One line is of equal length, the others differ by up to 3/4 inch (=1.85cm)



Ash Experiment: Design

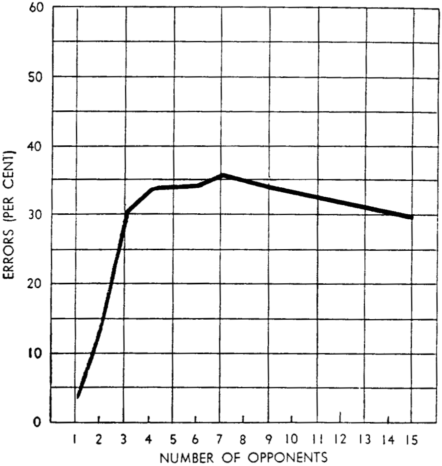
- Subjects publicly announce their judgement in a predetermined sequence
- Only the last subject in the sequence is actually a subject - all other participants are actors.
- During the first two trials, all actors agree on the correct line.
- Starting in the third trial, all actors agree on an incorrect line.
    - Occasionally, they choose the correct line to avoid suspicion (4 out of 18 trials)
    - Nevertheless, some subjects suspect collusion. These observations are dropped!

Ash Experiment: Main Result

- Control treatment: no actors
  - Subjects make mistakes roughly 1% of the time
- Measure the influence of the group: error rate when all predecessors choose the wrong line
  - 123 subjects join the majority's wrong judgement 36.8% of the time
- Individual differences
  - 25% of subjects never agree with the erroneous judgement of the majority while others almost always do
  - Some subjects report that they yielded in order "not to spoil the results" (p. 4)

Ash Experiment: Size of the Majority

- How large does the majority have to be to affect the subject's decision?

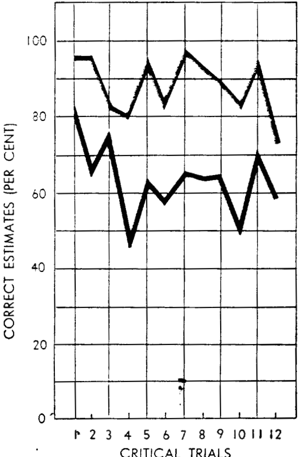


Ash Experiment: Ally

- Introducing a second dissenter massively reduces error rates
    - 8% errors when the other dissenter picks the correct line
    - 12% errors when the other dissenter picks the second best line
    - 9% errors when the other dissenter picks the worst possible line while the majority picks the second best line

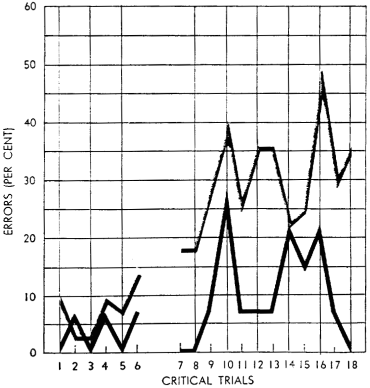
Ash Experiment: Ally

- Top line: one dissenter (subject or actor) picks the correct line
- Bottom line: no ally



Ash Experiment: Temporary Ally

- A dissenter leaves the room after 6 trials
  - Error rates only slightly increase
- A dissenter joins the majority after 6 trials
  - Error rates increase to the levels observed without a dissenter



Discussion

- “Life in society requires consensus as an indispensable condition” (p. 5)

Discussion

- "Life in society requires consensus as an indispensable condition" (p. 5)
  - Consensus about how decisions are made?
  - Consensus about values/norms (e.g., how much to contribute to a public good)?
  - Consensus about assessing objective facts?

Discussion

- “Life in society requires consensus as an indispensable condition” (p. 5)
  - Consensus about how decisions are made?
  - Consensus about values/norms (e.g., how much to contribute to a public good)?
  - Consensus about assessing objective facts?
- “That we have found the tendency to conformity in our society so strong that reasonably intelligent and well-meaning young people are willing to call white black is a matter of concern. It raises questions about our ways of education and about the values that guide our conduct.” (p. 5)

Discussion

- “Life in society requires consensus as an indispensable condition” (p. 5)
  - Consensus about how decisions are made?
  - Consensus about values/norms (e.g., how much to contribute to a public good)?
  - Consensus about assessing objective facts?
- “That we have found the tendency to conformity in our society so strong that reasonably intelligent and well-meaning young people are willing to call white black is a matter of concern. It raises questions about our ways of education and about the values that guide our conduct.” (p. 5)
  - But: it is more likely that one person is wrong than that 8 people are wrong. Therefore, deferring to the majority might be perfectly rational and could lead to better decisions

Discussion

- Ash's experiment has been successfully replicated many times but there have also been failed replication attempts where students did not follow the majority
- If the type of conformity measured in the Ash experiment is a matter of concern, it would be interesting to observe how it varies across culture and time
  - Example: Berry finds that high food-accumulating societies emphasize obedience (= high rates of conformity in the Ash experiment) while hunting and fishing peoples emphasize independence

Discussion

- Bond and Smith (1996) analyze 133 studies and find that conformity is affected by:
  - Size of the majority (+)
  - Consistency of the majority (+)
  - Proportion of female respondents (+)
  - Majority consists of out-group members (e.g., students vs. non-students) (-)
  - Whether the subject's decision is disclosed to the majority or not does not seem to have a significant effect
- They also find that conformity in the United States has declined since Ash conducted his experiment
- ...and that in highly collectivist societies (as measured by surveys), higher rates of conformity are observed

Information Cascades: Introduction

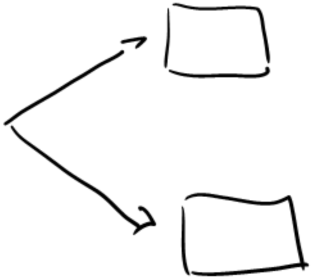
- Information cascades are how many economists like to explain observations such as conformity in the Ash experiment
- Idea:
  - Each person has some private information about which line is the correct line
  - If you can observe what others decide and you have no reason to believe that they are worse at the task than you, you might want to follow your predecessors even though you came to a different conclusion
  - This explanation is consistent with the observation that disclosing/not disclosing the subject's decision to the group has no effect on error rates (while fear of some kind of punishment would not be consistent)

Information Cascades: Definition

- "An informational cascade occurs when it is optimal for an individual, having observed the actions of those ahead of him, to follow the behavior of the preceding individual without regard to his own information" (Bikhchandani et al., 1992, p. 994).

Information Cascades: BHW-Model

- n agents decide in an exogenously determined sequence whether to *adopt* or *reject*
- Each individual observes the decisions of those ahead of him
- The cost of adopting is c, the value V of adopting is either 0 or 1 with equal probability, $c = \frac{1}{2} + \varepsilon$
- Each individual receives an independent private signal X_i, which is either H or L (high or low)
- H is observed with probability $p > 0.5$ if $V = 1$ and $1 - p$ if $V = 0$
- Assume that indifferent individuals adopt/reject with equal probability



Information Cascades: BHW Model

- Agents follow their predecessors as soon as one option has an advantage of 2
- At this point, an *information cascade* starts
- Private information is no longer aggregated: Agents follow the majority without considering their signals
- As a result, it is not unlikely that everybody adopts even though $V = 0$ (or nobody adopts even though $V = 1$)
- The lower the precision of private signals, the higher the probability of everybody choosing the wrong option
- Information cascades are fragile: It does not take much new information for everybody to change his mind because the information on which the decision to adopt/reject is based is weak

Discussion

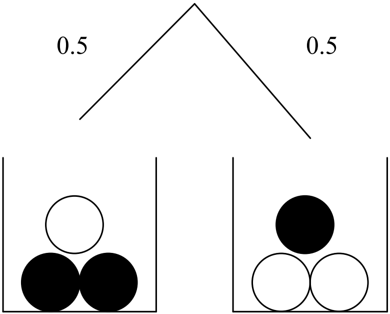
- Assumptions

Discussion

- Assumptions
  - All agents have the same preferences: Correlation would be sufficient
  - All agents receive a signal of equal quality: basic intuition also works when $p_i \neq p_j$
  - Distribution of signals/preferences is common knowledge
  - Exogeneous sequence of decisions

Experimental Evidence

- The first experimental study of information cascades is Anderson and Holt (1997)
- Experimental design
  - One of two urns is randomly chosen



other urn?

Experimental Evidence: Holt and Laury

- Experimental design
  - 6 subjects get a draw with replacement from the chosen urn ($p = 2/3$)
  - They then sequentially guess from which urn their draw came
  - They can observe what their predecessors choose
  - Agents who choose the correct urn receive $2, 0 otherwise
  - The same game is played 15 times
  - As soon as an imbalance of two decisions in one direction occurs, a cascade should form

Experimental Evidence: Holt and Laury

- What role do risk preferences play in this experiment?

Experimental Evidence: Holt and Laury

- What role do risk preferences play in this experiment?
  - None: the only uncertainty is about which option yields a *certain* payoff. Regardless of risk preference, all agents prefer more money to less money and therefore pick the option which is more likely to yield more money.

Experimental Evidence: Holt and Laury

- What role do risk preferences play in this experiment?
  - None: the only uncertainty is about which option yields a *certain* payoff. Regardless of risk preference, all agents prefer more money to less money and therefore pick the option which is more likely to yield more money.
- Is this a good test of Bayes' Rule?

Experimental Evidence: Holt and Laury

- What role do risk preferences play in this experiment?
  - None: the only uncertainty is about which option yields a *certain* payoff. Regardless of risk preference, all agents prefer more money to less money and therefore pick the option which is more likely to yield more money.
- Is this a good test of Bayes' Rule?
  - No, the prior that one option yields a positive payoff is 0.5
  - As a result, subjects should conform to the theoretical prediction even when suffering from base-rate neglect

Experimental Evidence: Holt and Laury

- Results overall: 87 / 122 cases correct
  - Let's only consider situations in which there is an imbalance of two or more decisions in one direction
  - In these situations, subjects should follow their predecessors (under common knowledge of rationality), regardless of their own signal
  - They fail to do so 26% of the time when their own signal is inconsistent with their predecessors' choices
  - Why?

Experimental Evidence: Weizsäcker

- Weizsäcker (2010) analyses data from 13 information cascade experiments
- We saw earlier that subjects rely too heavily on their own signal under the assumption of common knowledge of rationality
- Some subjects might not want to rely on the assumption that their predecessors are rational
- Weizsäcker computes *actual* payoffs and examines whether subjects behave optimally given *actual* behavior

Experimental Evidence: Weizsäcker

- To find out which actions are optimal given *actual* behavior, Weizäcker (2010) simply looks at situations in which the available information is identical (e.g., the first two agents both chose A)
- He then computes the fraction of such situations in which A yields a higher payoff than B
  - Need many observations to get a reliable estimate of expected payoffs given *actual* behavior
- We can now check whether following one's own signal in such a situation yields a higher or a lower payoff than not following one's own signal.

Experimental Evidence: Weizsäcker

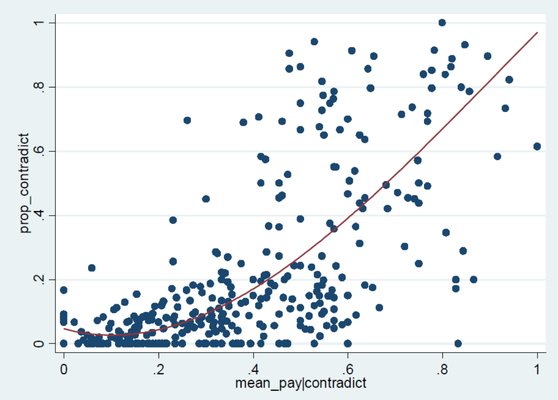
- In situations in which actual payoffs for going against one's own signal are higher than for following one's signal, subjects only choose the optimal course of action 44% of the time
- It may not be very costly to do so
  - Consider a situation in which the first two predecessors chose A. It could be that in these situations, A is correct 51% of the time and B 49% of the time
- Subjects only earn 53% of possible payoffs in situations in which it is *empirically* optimal to contradict one's own signal.
- They could have earned 64% of possible payoffs by never following their signal in these situations

Experimental Evidence: Weizsäcker

- For situations in which it is optimal given actual behavior to follow the private signal, subjects earn 73% of possible payoffs, while they could earn up to 75% by always behaving optimally
- In other words: subjects extract a much higher share of possible profits when it is best to follow their own signal rather than having to go against their own signal.
- These results suggest that we are not looking at equilibrium behavior. Rather, subjects follow their own signal too often.

Experimental Evidence: Weizsäcker

- x-axis: frequency with which own signal is wrong
- y-axis: frequency with which subjects contradict own signal
- Only situations with more than 10 observations



Information Cascades in Markets

- Alevy, Haigh and List (2007) write: “Arguably, however, the most well-known herds or cascades occur in financial markets, where bubbles and crashes may be examples of such behavior” (p. 151)
- How can an information cascade occur in a market?

Herding and Contrarian Behavior in Financial Markets:
An Internet Experiment

Mathias Drehmann, Jörg Oechssler, Andreas Roider

American Economic Review, Dez. 2005

- high volatility in financial markets often attributed to herding behavior of investors: “investors are like lemmings”

- high volatility in financial markets often attributed to herding behavior of investors: “investors are like lemmings”
- for various reasons difficult to verify in empirical work
  - do investors herd?

- high volatility in financial markets often attributed to herding behavior of investors: “investors are like lemmings”
- for various reasons difficult to verify in empirical work
  - do investors herd?
  - or, do they just follow the same information?

- high volatility in financial markets often attributed to herding behavior of investors: “investors are like lemmings”
- for various reasons difficult to verify in empirical work
  - do investors herd?
  - or, do they just follow the same information?
  - incidental clustering of actions?
  - private information unobservable $\Rightarrow$ experiments

Herding theories

- payoff externalities
- reputational externalities
- informational externalities: **informational cascades**

  $\Rightarrow$ ignore private informaton

  Bikhchandani/Hirshleifer/Welch (1992) **BHW**, Banerjee (1992)

- **but:** not directly applicable to financial markets.

  Avery/Zemsky (1998): market price efficiently aggregates information

  $\Rightarrow$ prevents cascades

  – Experiment by Cipriani/Guarino (2002)

The BHW model

- 2 investments: A or B, only one pays 10 Lotto–Euros, fixed price

The BHW model

- 2 investments: A or B, only one pays 10 Lotto–Euros, fixed price
- sequential decisions (e.g. 20 traders)
- private signal $s = a$ or b, only previous *decisions* are observable

base treatment:	$\cdot\ P(A) = 0.55$	(prior)
	$\cdot\ P(a\|A) = P(b\|B) = 0.6$	(signal precision)

The BHW model

- 2 investments: A or B, only one pays 10 Lotto–Euros, fixed price
- sequential decisions (e.g. 20 traders)
- private signal $s = a$ or b, only previous *decisions* are observable

  base treatment: $\cdot\ P(A) = 0.55$ (prior)

  $\cdot\ P(a|A) = P(b|B) = 0.6$ (signal precision)

- **rational herding** $\Delta = \#a - \#b =$ net balance of imputed signals

  $\Delta \geq 1 \quad \Rightarrow$ A-cascade

  $\Delta \leq -2 \ \Rightarrow$ B-cascade

Cascades on financial markets? Avery/Zemsky (1998)

- market price reflects all publicly available information (H_t)

Cascades on financial markets? Avery/Zemsky (1998)

- market price reflects all publicly available information (H_t)
- price of A: $p_t = 10P(A|H_t)$ (price of $B = 10 - p_t$)

Cascades on financial markets? Avery/Zemsky (1998)

- market price reflects all publicly available information (H_t)
- price of A: $p_t = 10P(A|H_t)$ (price of $B = 10 - p_t$)
- investment in A profitable iff

$$10P(A|H_t, s) - p_t = 10P(A|H_t, s) - 10P(A|H_t) > 0$$

- $\Rightarrow$ always follow own signal $\Rightarrow$ no herding

Cascades on financial markets? Avery/Zemsky (1998)

- market price reflects all publicly available information (H_t)
- price of A: $p_t = 10P(A|H_t)$ (price of $B = 10 - p_t$)
- investment in A profitable iff

$$10P(A|H_t, s) - p_t = 10P(A|H_t, s) - 10P(A|H_t) > 0$$

- $\Rightarrow$ always follow own signal $\Rightarrow$ no herding
- $E(p_{t+1}|H_t) = p_t$, martingale property of prices
- only with transaction costs can no–trade be rational

Objectives

- test cascade theory at face value

Objectives

- test cascade theory at face value
- find empirical regularities: do traders...
  - ...follow their own signals?
  - ...engage in herding?
  - ...behave like contrarians?

Objectives

- test cascade theory at face value
- find empirical regularities: do traders...
  - ...follow their own signals?
  - ...engage in herding?
  - ...behave like contrarians?
- do traders believe in martingale property of prices?

Procedures

- **internet experiment** (`www.a-oder-b.de`)
  - phase I: consultants of an international consulting firm
  - phase II: posters at 40 universities, emails to doctoral students, ads in national newspapers: DIE ZEIT, Unicum, Audimax

Procedures

- **internet experiment** (`www.a-oder-b.de`)
  - phase I: consultants of an international consulting firm
  - phase II: posters at 40 universities, emails to doctoral students, ads in national newspapers: DIE ZEIT, Unicum, Audimax
- in total 6099 subjects (including 267 consultants)
  - each subject: 3 treatments in groups of 20 subjects (sometimes less)

Procedures

- **internet experiment** (`www.a-oder-b.de`)
  - phase I: consultants of an international consulting firm
  - phase II: posters at 40 universities, emails to doctoral students, ads in national newspapers: DIE ZEIT, Unicum, Audimax
- in total 6099 subjects (including 267 consultants)
  - each subject: 3 treatments in groups of 20 subjects (sometimes less)
- incentives: participants won lottery tickets
  - consultants: dinner vouchers, total value 1,200 Euros
  - phase II: 11 cash prizes à 1,000 Euros (exp. hourly pay $\approx$ 15 Euro (initially)

Properties of the subject pool

- very well educated: 31% are Ph.D. students, 13% hold a Ph.D.
- almost 50% trained in math, natural sciences, engineering
- 28% female subjects, average age: 28
- average duration $\approx$ 15 min., 687 subjects provided feedback

Treatments

T. group	Treatments	Description
P	P+D	Price, all decisions observable* (*=bet on final price)
	P-D	Price, no decisions observable
	Pe+D	Price, explicit formulation, all decisions observable*

Treatments

T. group	Treatments	Description
P	*P+D*	Price, all decisions observable* (*=bet on final price)
	P-D	Price, no decisions observable
	Pe+D	Price, explicit formulation, all decisions observable*
P-N	*P+D-N*	Price, all decisions observable, *N* not possible*
	P-D-N	Price, no decisions observable, *N* not possible

Treatments

T. group	Treatments	Description
P	P+D	Price, all decisions observable* (*=bet on final price)
	P-D	Price, no decisions observable
	Pe+D	Price, explicit formulation, all decisions observable*
P-N	P+D-N	Price, all decisions observable, N not possible*
	P-D-N	Price, no decisions observable, N not possible
P+T_i	P+D+T_i	+ transactions costs of size i (0.1 or 0.5)
	Pe+D+T_i	+ transactions costs of size i (0.1 or 0.5)

Treatments

T. group	Treatments	Description
P	P+D	Price, all decisions observable* (*=bet on final price)
	P-D	Price, no decisions observable
	Pe+D	Price, explicit formulation, all decisions observable*
P-N	P+D-N	Price, all decisions observable, N not possible*
	P-D-N	Price, no decisions observable, N not possible
$P+T_i$	$P+D+T_i$	+ transactions costs of size i (0.1 or 0.5)
	$Pe+D+T_i$	+ transactions costs of size i (0.1 or 0.5)
BHW		Bikhchandani/Hirshleifer/Welch
BHW+AS		BHW + all signals observable

- 7 different probability combinations (prior/precision)

Variables of interest

1. “Rational under common knowledge”: $ruck \in \{0, 1\}$

   Bayesian with respect to the imputed signal history assuming predecessors to be rational.

Variables of interest

1. "Rational under common knowledge": $ruck \in \{0, 1\}$

   Bayesian with respect to the imputed signal history assuming predecessors to be rational.

2. "Follow own signal": $own \in \{0, 1\}$

Theoretical predictions

1. P and $P-N$ treatments: $ruck = own = 1$

   $\rightarrow$ observed decisions (D) and no–trade option N should not make a difference

Theoretical predictions

1. P and $P - N$ treatments: $ruck = own = 1$

   $\rightarrow$ observed decisions (D) and no–trade option N should not make a difference

2. N should only be observed in $P + T$

Theoretical predictions

1. P and $P-N$ treatments: $ruck = own = 1$

   $\rightarrow$ observed decisions (D) and no–trade option N should not make a difference

2. N should only be observed in $P+T$

3. BHW and $BHW+AS$ treatments: $ruck = 1$ and $own < 1$

Theoretical predictions

1. P and $P-N$ treatments: $ruck = own = 1$

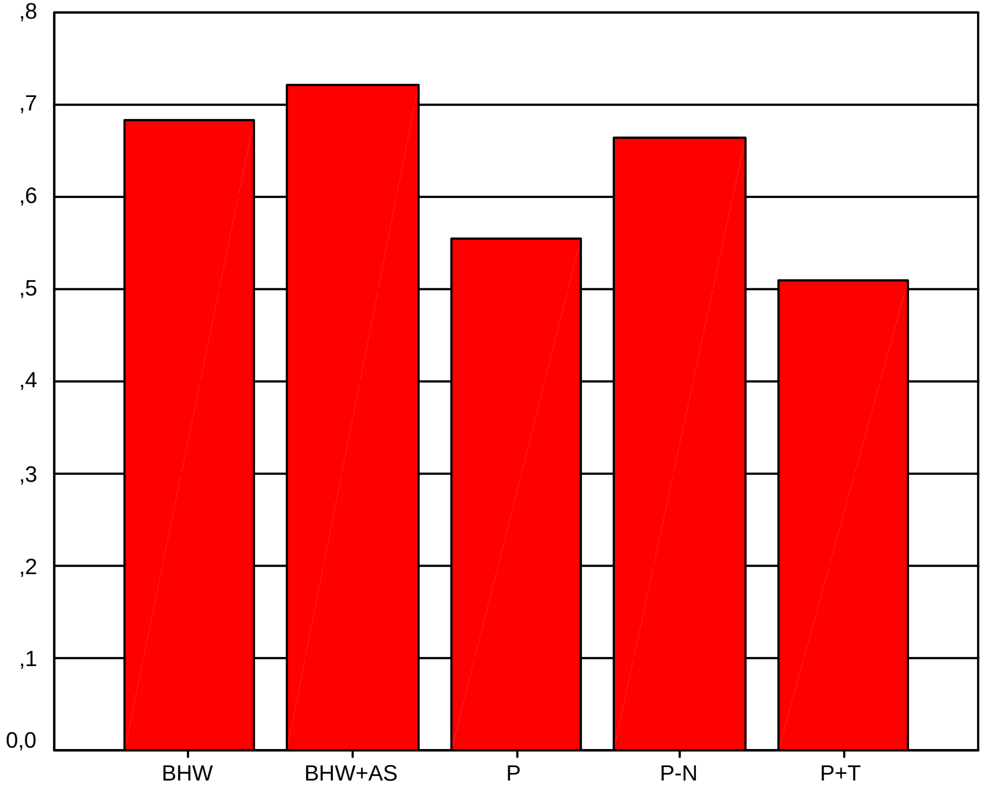
   $\rightarrow$ observed decisions (D) and no–trade option N should not make a difference

2. N should only be observed in $P+T$

3. BHW and $BHW+AS$ treatments: $ruck = 1$ and $own < 1$

4. Different prior probabilities and signal precisions should not alter average $ruck$

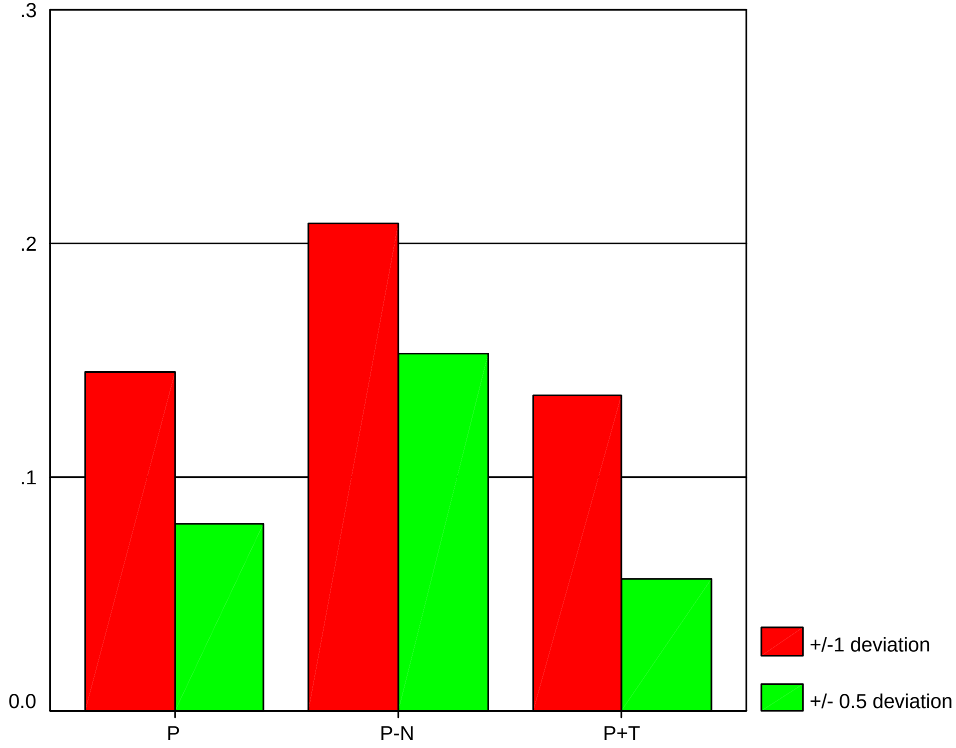
Ruck across treatments (55-60)



- No trade: 20% in P, 26% in $P{+}T$
- contrary to prediction: $own^{BHW} > own^{Price}$

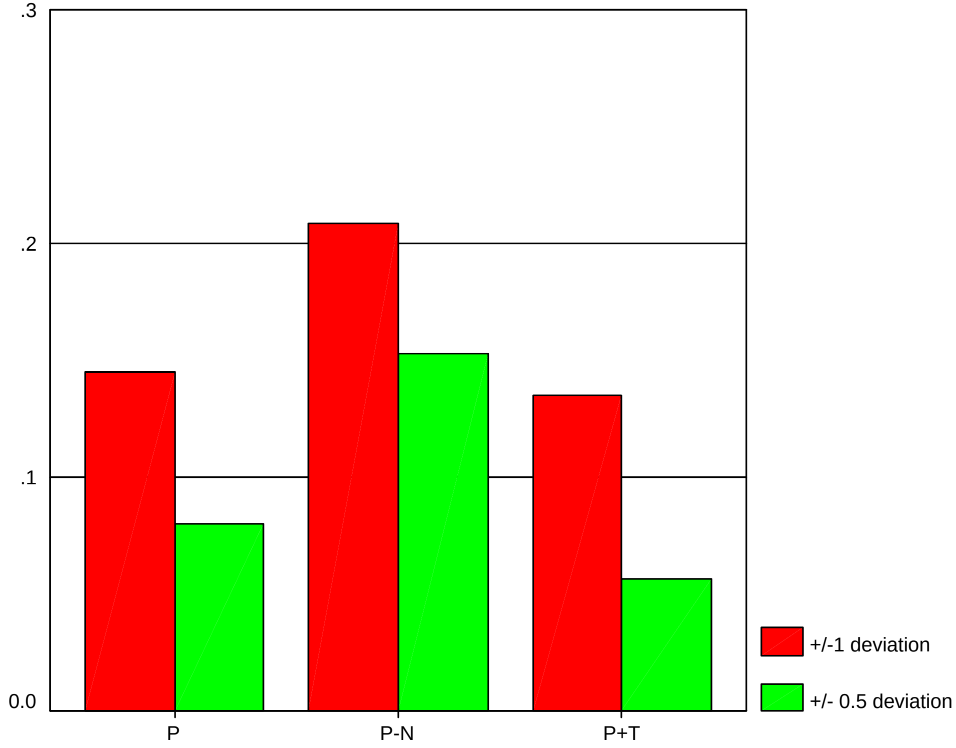
Convergence of the final price

Fraction of actual final prices within $\pm 1 (\pm 0.5)$ of final full-info price (all prob.)



Convergence of the final price

Fraction of actual final prices within $\pm 1 (\pm 0.5)$ of final full-info price (all prob.)



- Diff. betw. P-N and $P+T$ (P) significant at 1% (10%) level (only for ± 0.5)

Contrarian behavior

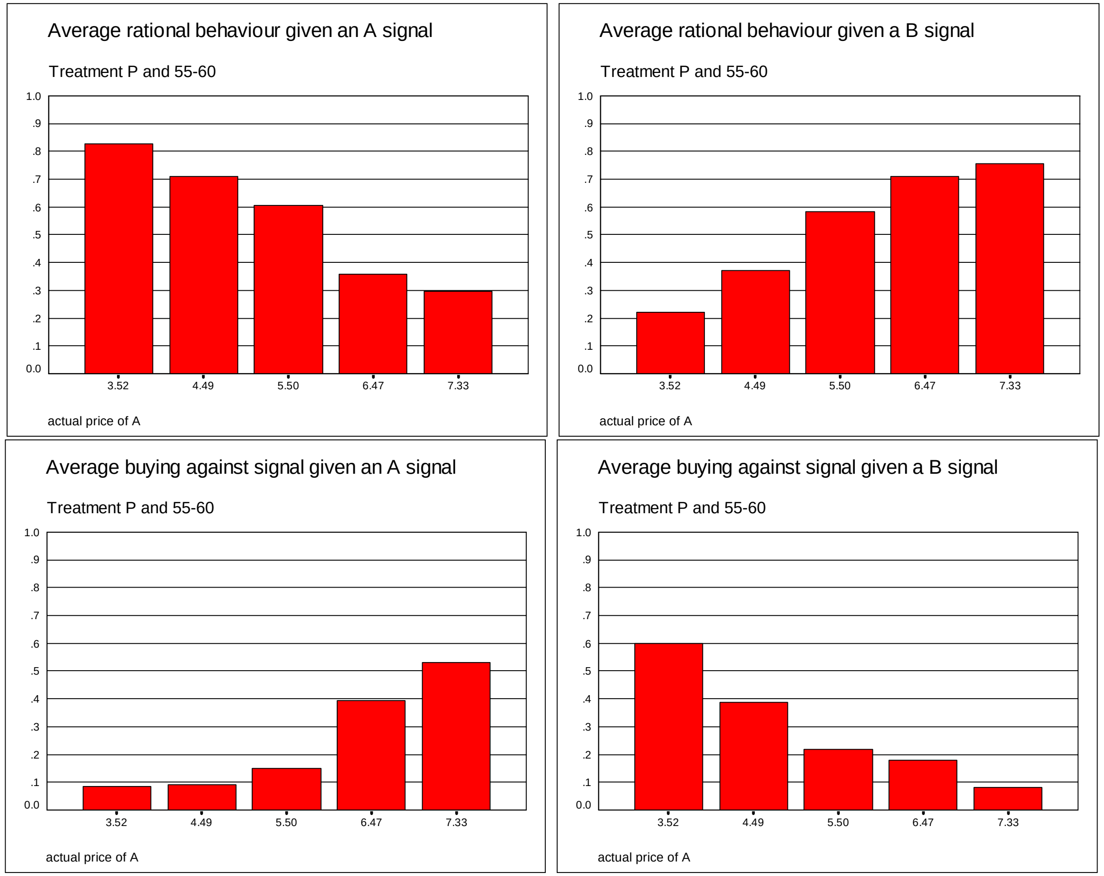
- contrarians = traders who trade against their signal and the market

  i.e. buy other asset at high prices, contrary to signal

Contrarian behavior

- contrarians = traders who trade against their signal and the market

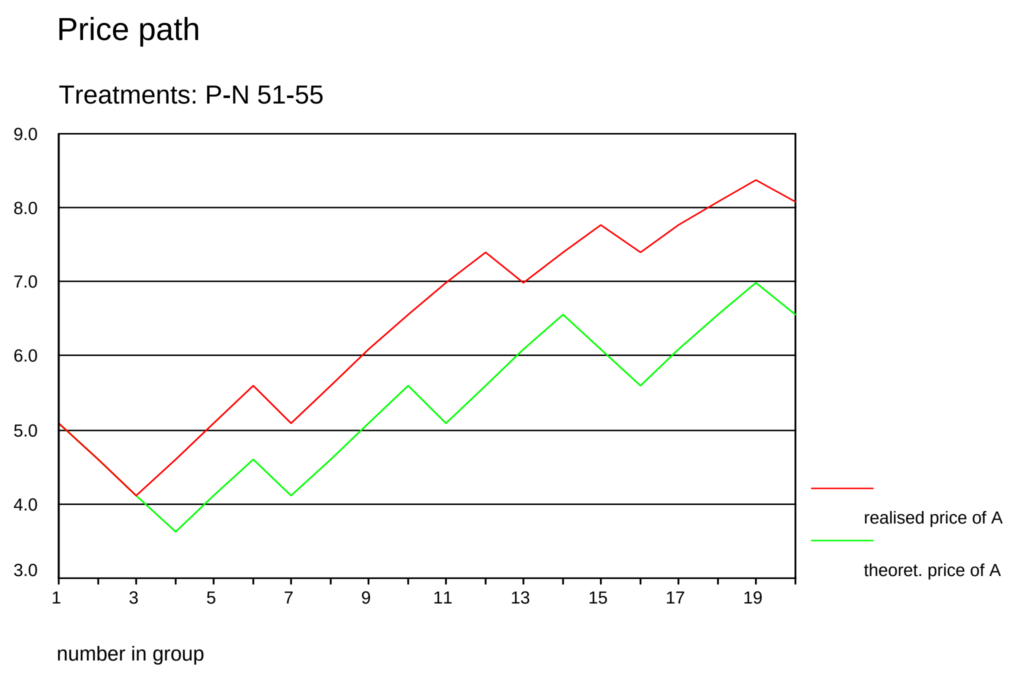
  i.e. buy other asset at high prices, contrary to signal



Is it a good idea to be a contrarian?

- paradox: contrarian behavior profitable if prices overshoot, but if many contrarians, prices don't overshoot.

  $\rightarrow$ but: contrarian behavior itself might lead to overshooting:



Is it a good idea to be a contrarian?

- mix of noise traders and contrarians in the market

Is it a good idea to be a contrarian?

- mix of noise traders and contrarians in the market
- if actual price high $\Rightarrow$ likely to be caused by noise traders

  $\Rightarrow$ theoretical price $<$ actual price (*regression to the mean*)

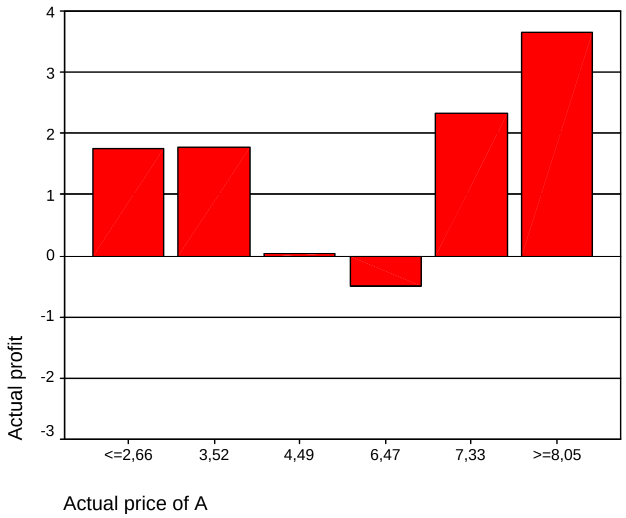
  $\Rightarrow$ contrarians fare better

Profits and *Ruck*

- On average it payed to be a Bayesian:
  - $ruck$ and $profits$ are positively correlated
- But ...

Profits of contrarians across prices

(P-N, 55-60)



Main results

1. Substantial fraction of subjects trade against own signal

Main results

1. Substantial fraction of subjects trade against own signal
2. Divergence of the prices from theoretical prices

Main results

1. Substantial fraction of subjects trade against own signal

2. Divergence of the prices from theoretical prices

3. Undershooting of prices

Main results

1. Substantial fraction of subjects trade against own signal

2. Divergence of the prices from theoretical prices

3. Undershooting of prices

4. No evidence for herding / imitation

Main results

1. Substantial fraction of subjects trade against own signal
2. Divergence of the prices from theoretical prices
3. Undershooting of prices
4. No evidence for herding / imitation
5. Strong evidence for contrarian behavior

Main results

1. Substantial fraction of subjects trade against own signal

2. Divergence of the prices from theoretical prices

3. Undershooting of prices

4. No evidence for herding / imitation

5. Strong evidence for contrarian behavior

6. Contrarian behavior can be rationalized through error models

Main results

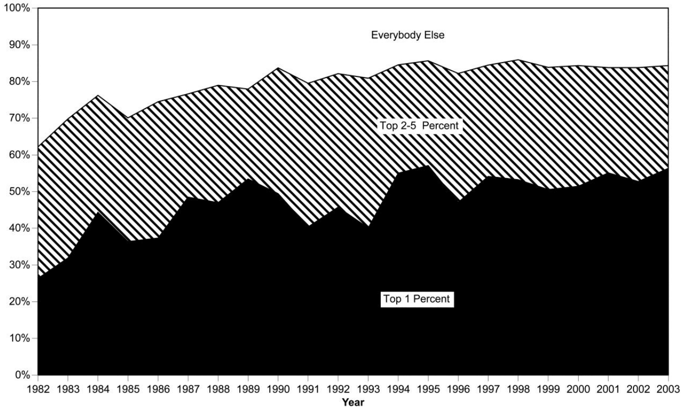
1. Substantial fraction of subjects trade against own signal
2. Divergence of the prices from theoretical prices
3. Undershooting of prices
4. No evidence for herding / imitation
5. Strong evidence for contrarian behavior
6. Contrarian behavior can be rationalized through error models
7. Control group of consultants not significantly different

Application: Cultural Markets

- Empirical Observations
  - The most successful songs, books and movies are much more successful than average
  - Experts are not very good at predicting which songs/movies/books will be successful

Example: Distribution of Ticket Revenue

- Share of total ticket revenue accruing to top performers, Krueger (2005)



Possible Explanations

Possible Explanations

- Mapping from quality to success is convex: small differences in quality lead to large differences in success
  - 10 mediocre concerts are not necessarily as good as 1 really good one. Partly because consumption is time consuming
  - Economies of scale in production and distribution allow few producers to service the entire market
  - Problem: If experts can measure quality, they should be able to predict success
- Consumers' decisions are influenced by the behavior of others

Experimental Design (Salganik et al., 2006)

- Treatments
  - Independent
    - Participants can see names of bands (18) and songs (48)
    - While listening to a song, they can rate it on a scale from one to five
    - After having rated a song, they have the opportunity to download it
  - Social influence
    - Same as independent except participants not only see the names of bands and songs but also the number of previous downloads
    - Experiment 1: Songs are randomly ordered in a 16 x 3 rectangular grid (also used in the indepenent treatment)
    - Experiment 2: Songs are presented in a single column format, ordered by previous downloads

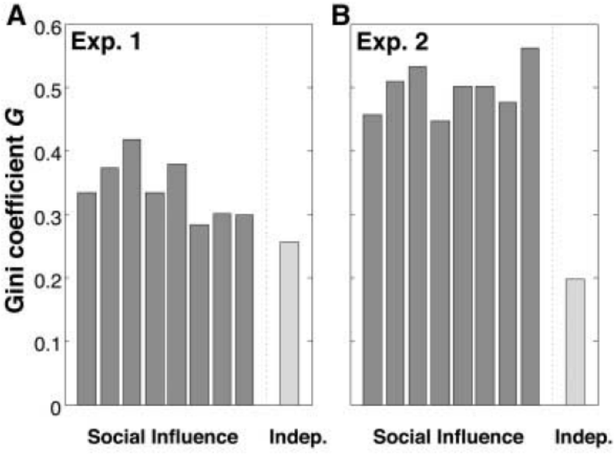
Experimental Design

- 8 separate worlds (plus one for the independent treatment): participants can only see the number of downloads within their world
- 14,341 participants, recruited on a teen-interest website, randomly assigned to a treatment and a world

Results: Inequality

- Measuring the distribution of success
  - Success of a song: $m_i = \frac{d_i}{\sum_{k=1}^{S} d_k}$
    - d_i: song i's download count
    - S: total number of songs (48)
  - Success inequality: $G = \frac{\sum_{i=1}^{S} \sum_{j=1}^{S} |m_i - m_j|}{2S \sum_{k=1}^{S} m_k}$ (average difference/sum of market shares)
- If social interaction leads to higher inequality of success, we expect G to be higher for treatment *social influence* than for treatment *independent*.
- The influence of other people's choices should be larger when it is more salient (i.e., larger difference in experiment 2 than in experiment 1)

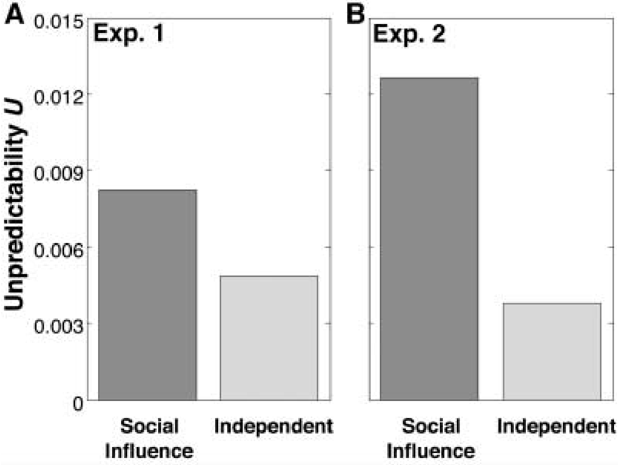
Results: Inequality



Results: Unpredictability

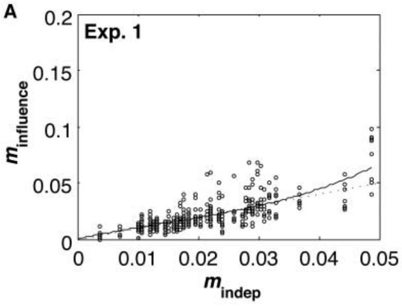
- Measuring unpredictability of success in the social influence treatment
  - $u_i = \sum_{j=1}^{W} \sum_{k=j+1}^{W} |m_{ij} - m_{ik}| / \binom{W}{2}$
  - m_{ij}: song i's market share in world j
  - $u_i =$ average absolute value of the difference in market shares of song i between two worlds
  - $U = \sum_{i=1}^{S} u_i / S$: average upredictability
- Measuring unpredictability of success in the independent treatment
  - Randomly split the single independent world into two subsets and compute the same measure U
  - Do this 1000 times and take the average of U
- If social influence increases unpredictability we would expect U to be higher in the social influence treatment than in the independent treatment. Moreover, this difference should be larger in experiment 2 than in experiment 1

Results: Unpredictability



Results: Unpredictability

- Relationship between quality (market share in treatment independent) and success (market share in treatment social influence)



Unpredictability

- Quality and success are positively correlated
- The mapping from quality to success is convex when social influence is strong
- Success is hard to predict based on quality

Summary

- Higher inequality of success in the presence of social influence
- Higher unpredictability of success in the presence of social influence
- Quality is only weakly correlated with success

Discussion

- How can we explain why social influence increases both inequality and unpredictability of success?

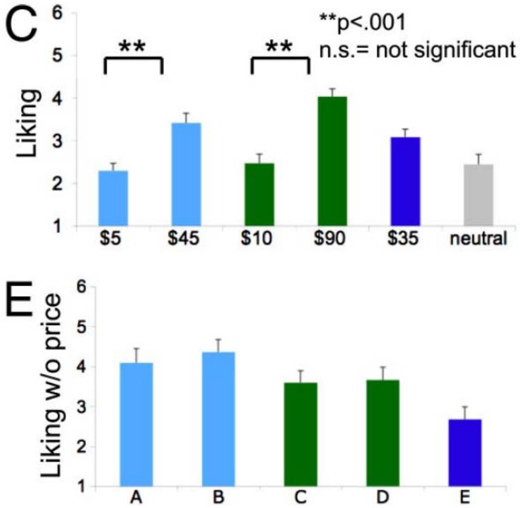
Discussion

- How can we explain why social influence increases both inequality and unpredictability of success?
- Information Cascades?
  - If subjects have similar preferences and no time to listen to all songs, they will only listen to those that have been downloaded by others
  - As a result, the first few subjects have a strong influence on the success of a song
  - Example: purchase your own book in order to make sure it gets on the bestseller list
- Preferences depend on other's actions?
  - A song is more valuable if I can talk to others about it
  - I might enjoy listening to a song more if I know that it is considered to be a masterpiece

Interdependent Preferences

- Does the utility a good generates depend on how others assess it?
- Plassmann et al. (2008) provide evidence supporting this claim
- Experimental design
  - Tell subjects that they are tasting 5 different wines that sell at different prices per bottle: \$5 (wine 1), \$10 (wine 2), \$35 (wine 3), \$45 (wine 1), \$90 (wine 2)
  - Ask subjects to rate these wines
  - Control treatment: Let the same subjects taste the same wines 8 weeks later but no price information

Results


C
Liking
1
2
3
4
5
6
**
**
**p<.001
n.s.= not significant
$5
$45
$10
$90
$35
neutral
E
Liking w/o price
1
2
3
4
5
6
A
B
C
D
E

Results

- Subjects rate the same wine significantly higher when they think it is more expensive
- Subjects are Caltech students - it might be a good idea for them to rely on information other than their own impression when rating a wine...
- Or: do they really enjoy wines that are supposedly more expensive more?

Results

- Plassmann et al. (2008) claim that they can measure the “experienced pleasantness” of tasting wine by measuring activity of a brain region associated with experiencing pleasure
- Is this “experienced pleasure” equivalent to utility?

Results

- Plassmann et al. (2008) claim that they can measure the "experienced pleasantness" of tasting wine by measuring activity of a brain region associated with experiencing pleasure
- Is this "experienced pleasure" equivalent to utility?
- The utility of a good might not be limited to the immediately experienced pleasure
    - pleasant expectation before consumption takes place
    - unpleasant effects the next morning
    - unpleasant effect, for example when driving after having tasted an excellent wine
    - brain regions interact: the same level of activation might have a different effect depending on differences in activity levels elsewhere

Discussion

- Reporting a different price has a significant effect on the level of activation of a brain region associated with experiencing pleasure
- Would you still experience more pleasure if you knew it was the same wine but packaged and labeled differently and sold at a higher price?