Example: Equity premium

Stocks (or equities) tend to have more variable annual price changes (or is higher, as a way of compensating investors for the additional risk they bear. In most of this century, for example, stock returns were about 8% per year higher than bond returns. This was accepted as a reasonable return premium for equities until Mehra and Prescott (1985) asked how large a degree of risk-aversion is implied by this premium.

The answer is surprising-- under the standard assumptions of economic theory, investors must be extremely risk-averse to demand such a high premium. For example, a person with enough risk-aversion to explain the equity premium would be indifferent between a coin flip paying either $50,000 or $100,000, and a sure amount of $51,329.

Benartzi and Thaler(1997) offer prospect theory based explanation: Investors are not averse to the variability of returns, they are averse to loss (the chance that returns are negative). Since annual stock returns are negative much more frequently than annual bond returns are, loss-averse investors will demand a large equity premium to compensate them for the much higher chance of losing money in a year. (Note: Higher average return to stocks means that the cumulative return to stocks over a longer horizon is increasingly likely to be positive as the horizon lengthens.)

They compute the expected prospect values of stock and bond returns over various horizons, using estimates of investor utility functions from Kahneman and Tversky (1992), and including a loss-aversion coefficient of 2.25 (i.e., the disutility of a small loss is 2.25 times as large as the utility of an equal gain). Benartzi and Thaler show that over a one-year horizon, the prospect values of stock and bond returns are about the same if stocks return 8% more than bonds, which explains the equity premium.

Example: Disposition effect

Shefrin and Statman (1985) predicted that because people dislike incurring losses much more than they like incurring gains, and are willing to gamble in the domain of losses, investors will hold on to stocks that have lost value (relative to their purchase price) too long and will be eager to sell stocks that have risen in value. They called this the disposition effect.

The disposition effect is anomalous because the purchase price of a stock should not matter much for whether you decided to sell it. If you think the stock will rise, you should keep it; if you think it will fall, you should sell it. In addition, tax laws encourage people to sell losers rather than winners, since such sales generate losses which can be used to reduce the taxes owed on capital gains.

Disposition effects have been found in experiments by (Weber and Camerer, 1998).

On large exchanges, trading volume of stocks that have fallen in price is lower than for stocks that have risen.

The best field study was done by Odean. He obtained data from a brokerage firm about all the purchases and sales of a large sample of individual investors. He found that investors held losing stocks a median of 124 days, and held winners only 104 days.

Hold losers because they expect them to bounce back (or mean- revert)? Odean's sample, the unsold losers returned only 5% in the subsequent year, while winners that weres old later returned 11.6%. Tax incentives inverse behaviour.

Example: New York cab drivers working hours

Camerer, Babcock, Loewenstein and Thaler studied cab drivers in New York City about when they decide to quit driving each day. Most of the drivers lease their cabs, for a fixed fee, for up to 12 hours. Many said they set an income target for the day, and quit when they reach that target. While daily income targeting seems sensible, it implies that drivers will work long hours on bad days when the per-hour wage is low, and will quit earlier on good high-wage days. The standard theory of the supply of labor predicts the opposite: Drivers will work the hours which are most profitable, quitting early on bad day, and making up the shortfall by working longer on good days.

The daily targeting theory and the standard theory of labor supply therefore predict opposite signs of the correlation between hours and the daily wage. To measure the correlation, we collected three samples of data on how many hours drivers worked on different days. The correlation between hours and wages was strongly negative for inexperienced drivers and close to zero for experienced drivers.

This suggests that inexperienced drivers began using a daily income targeting heuristic, but those who did so either tended to quit, or learned by experience to shift toward driving around the same number of hours every day.

Daily income targeting assumes loss-aversion in an indirect way. To explain why the correlation between hours and wages for inexperienced drivers is so strongly negative, one needs to assume that drivers take a one-day horizon, and have a utility function for the day.

Example: Racetrack betting

In parimutuel betting on horse races, there is a pronounced bias toward betting on longshots , horses with a relatively small chance of winning. That is, if one groups longshots with the same percentage of money bet on them into a class, the fraction of time horses in that class win is far smaller than the percentage of money bet on them. Horses with 2% of the total money bet on them, for example, win only about 1% of the time (see Thaler and Ziemba, 1988; Hausch and Ziemba, 1995).

The fact that longshots are overbet implies favorites are underbet. Indeed, some horses are so heavily favored that up to 70% of the win money is wagered on them. For these heavy favorites, the return for a dollar bet is very low if the horse wins. (Since the track keeps about 15% of the money bet for expenses and profit, bettors who bet on such a heavy favorite share only 85% of the money with 70% of the people, a payoff of only about $2.40 for a $2 bet.) People dislike these bets so much that in fact, if you make those bets you can earn a small positive profit (even accounting for the track s 15% take).

There are many explanations for the favorite-longshot bias, each of which probably contributes to the phenomenon. Horses that have lost many races ina row tend to be longshots, so a gambler s fallacy belief that such horses are due for a win may contribute to overbetting on them. Prospect theoretic overweighting of low probabilities of winning will also lead to overbetting of longshots.

Examples: Insurance options

Samuelson and Zeckhauser (1988) coined the term status quo bias to refer to an exaggerated preference for the stat us quo, and showed such a bias in a series of experiments. They also reported several observations in field data which are consistent with status quo bias.

When Harvard University added new health-care plan options, older faculty members who were hired previously, when the new options were not available were, of course, allowed to switch to the new options. If one assumes that the new and old faculty members have essentially the same preferences for health care plans, then the distribution of plans elected by new and old faculty should be the same. However, Samuelson and Zeckhauser found that older faculty members tended to stick to their previous plans; compared to the newer faculty members, fewer of the old faculty elected new options.

In cases where there is no status quo, people may have an exaggerated preference for whichever option is the default choice. Johnson, Hershey, Meszaros, and Kunreuther (1993) observed this phenomenon in decisions involving insurance purchases. At the time of their study, Pennsylvania and New Jersey legislators were considering various kinds of tort reform, allowing firms to offer cheaper automobile insurance which limited the rights of the insured person to sue for damages from accidents. Both states adopted very similar forms of limited insurance, but they chose different default options, creating a natural experiment. All insurance companies mailed forms to their customers, asking the customers whether they wanted the cheaper limited-rights insurance or the unlimited-rights insurance. One state made the limited-rights insurance the default-- the insured person would get that if they did not respond-- and the other made unlimited-rights the default. In fact, the percentage of people electing the limited-rights insurance was higher in the state where that was the default. An experiment replicated the effect.

Examples: Buying & selling prices

A closely related body of research on endowment effects established that buying and selling prices for a good are often quite different.

The paradigmatic experimental demonstration of this is the mugs experiments of Kahneman, Knetsch and Thaler (1990). In their experiments, some subjects are endowed (randomly) with coffee mugs and others are not. Those who are given the mugs demand a price about 2-3 times as large as the price that those without mugs are willing to pay, even though in economic theory these prices should be extremely close together. In fact, the mugs experiments were inspired by field observations of large gaps in hypothetical buying and selling prices in contingent valuations.

Contingent valuations are measurements of the economic value of goods which are not normally traded--like clean air, environmental damage, and so forth. These money valuations are used for doing benefit-cost analysis and establishing economic damages in lawsuits.

There is a huge literature establishing that selling prices are generally much larger than buying prices, although there is a heated debate among psychologists and economists about what the price gap means, and how to measure true valuations in the face of such a gap.

Effects: Endowment effect etc

Three phenomena:

status quo biases

default preference

endowment effects

All consistent with aversion to losses relative to a reference point. Making one option the status quo or default, or endowing a person with a good (even hypothetically), seems to establish a reference point people move away from only reluctantly, or if they are paid a large sum.

On normative theory

“Give me an axiom, and I’ll design the experiment that refutes it.”

Amos Tversky

Israeli-American mathematical psychologist
Heuristics & biases research program about probabilistic judgement and decision making under uncertainty, prospect theory
Daniel Kahneman briefly after receiving the Nobel Prize for prospect theory:
"*I feel it is a joint prize. We were twinned for more than a decade.*"
Grawemeyer Award in Psychology

On importance

“Nothing in life is quite as important as you think it is while you're thinking about it.”

Daniel Kahneman

Israeli-American psychologist
Psychology of judgment and decision-making, prospect theory, behavioral economics, hedonic psychology
Nobel Memorial Prize in Economic Sciences

On economists

"Economists think about what people ought to do. Psychologists watch what they actually do."

On mathematicians

"People who know math understand what other mortals understand, but other mortals do not understand them. This asymmetry gives them a presumption of superior ability."

Daniel Kahneman

Another descriptive approach: Ecological rationality

Gigerenzer's school of thought

In a nutshell:

Decision have to be made under uncertainty and under limited resources such as time and knowledge. Simple rules of thumb are successfully used by humans (and animals).

These heuristics are not inferior to so-called rational theory of decision-making, but produce good decisions in real-life situations.

Gigerenzer, Todd & the ABC Research Group, *Simple Heuristics That Make Us Smart, 1999,* OUP

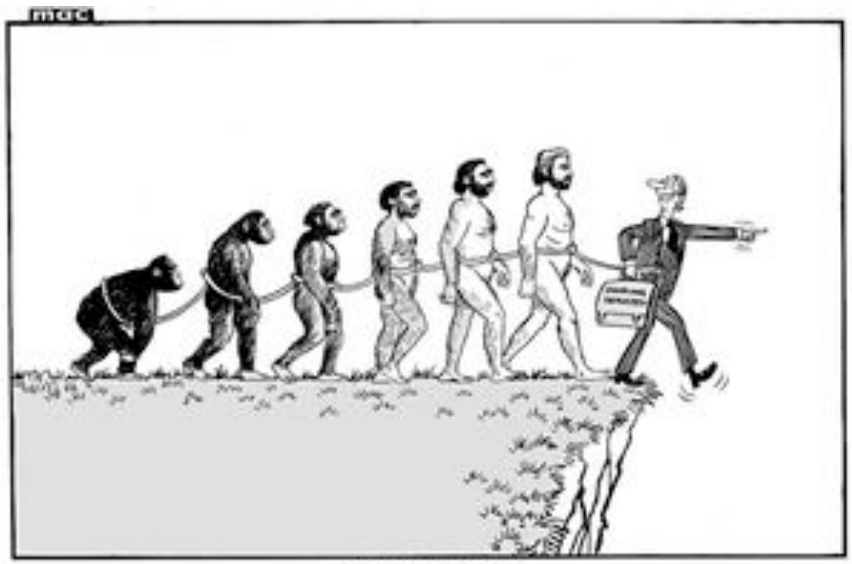
Gigerenzer and Brighton: *Homo Heuristicus: Why Biased Minds Make Better Inferences*, Topics in Cognitive Science 1 (2009) 107–143

More references and links at:

http://fastandfrugal.com/

https://www.mpib-berlin.mpg.de/en/research/adaptive-behavior-and-cognition/key-concepts

New main character: Homo heuristicus



Evolution stops here

Evolutionary perspective:
Humans adapt to environment

Animal's rely on heuristics to solve problems: e.g. in measuring, navigating, mating

heuristic (Greek) = "serving to find out or discover"

Heuristics and biases: from K & T's angle they are fallacies (violate logic, probability and other standards of rationality)

Heuristics: G emphasises their usefulness in problem solving

Heuristics - wrong and/or useful?

By the end of the 20th century, the use of heuristics became associated with shoddy mental software, generating three widespread misconceptions:

1. Heuristics are always second-best.
2. We use heuristics only because of our cognitive limitations.
3. More information, more computation, and more time would always be better.

Accuracy-effort trade-off: (general low of cognition)

Information and computation cost time and effort; therefore, minds rely on simple heuristics that are less accurate than strategies that use more information and computation.

Question: Is more better?

Accuracy-effort trade-off:

Carnap: "Principle of total evidence"

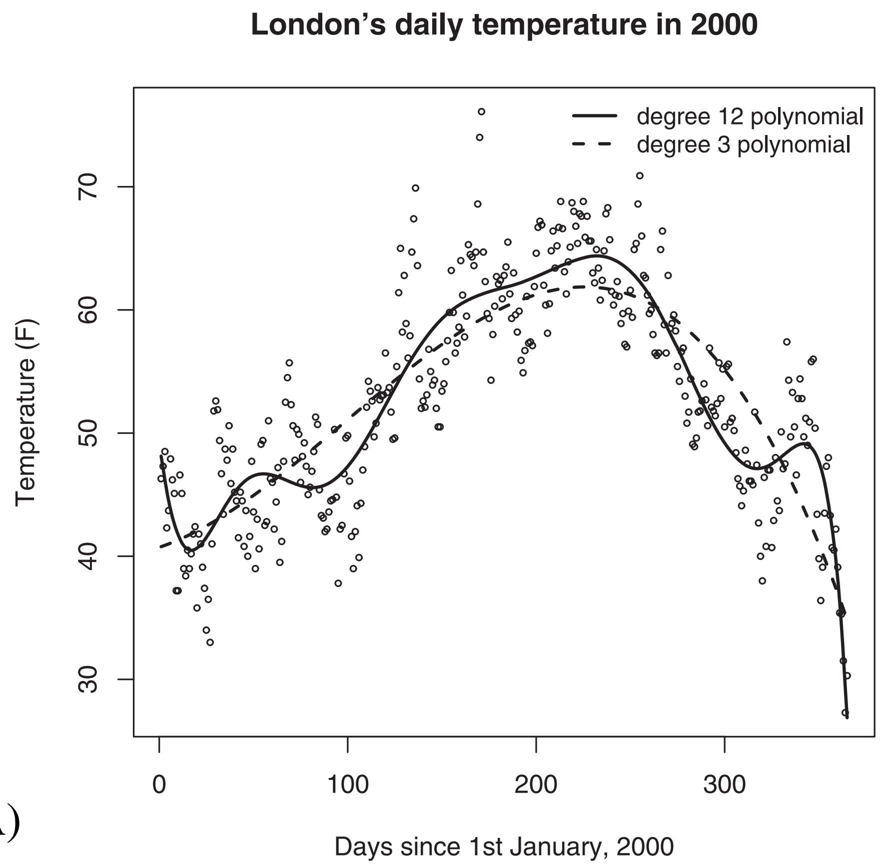
Good: "it is irrational to leave observations in the record but not use them"

First important discovery:
Heuristics can lead to more accurate inferences than strategies that use more information and computation

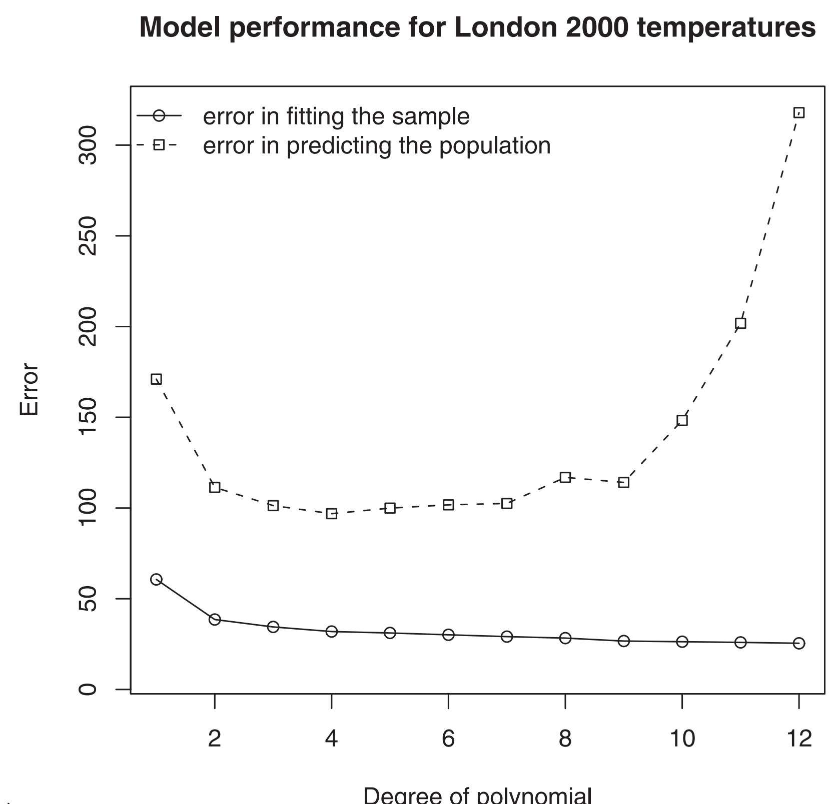
Less-is-more effects:
More information or computation can decrease accuracy; therefore, minds rely on simple heuristics in order to be more accurate than strategies that use more information and time. This relates to the existence of a point at which more information or computation becomes detrimental, independent of costs.

Example: Daily temperature series

Mean daily temperature for London in 2000 as a function of days and alternative polynomial fits



(A)

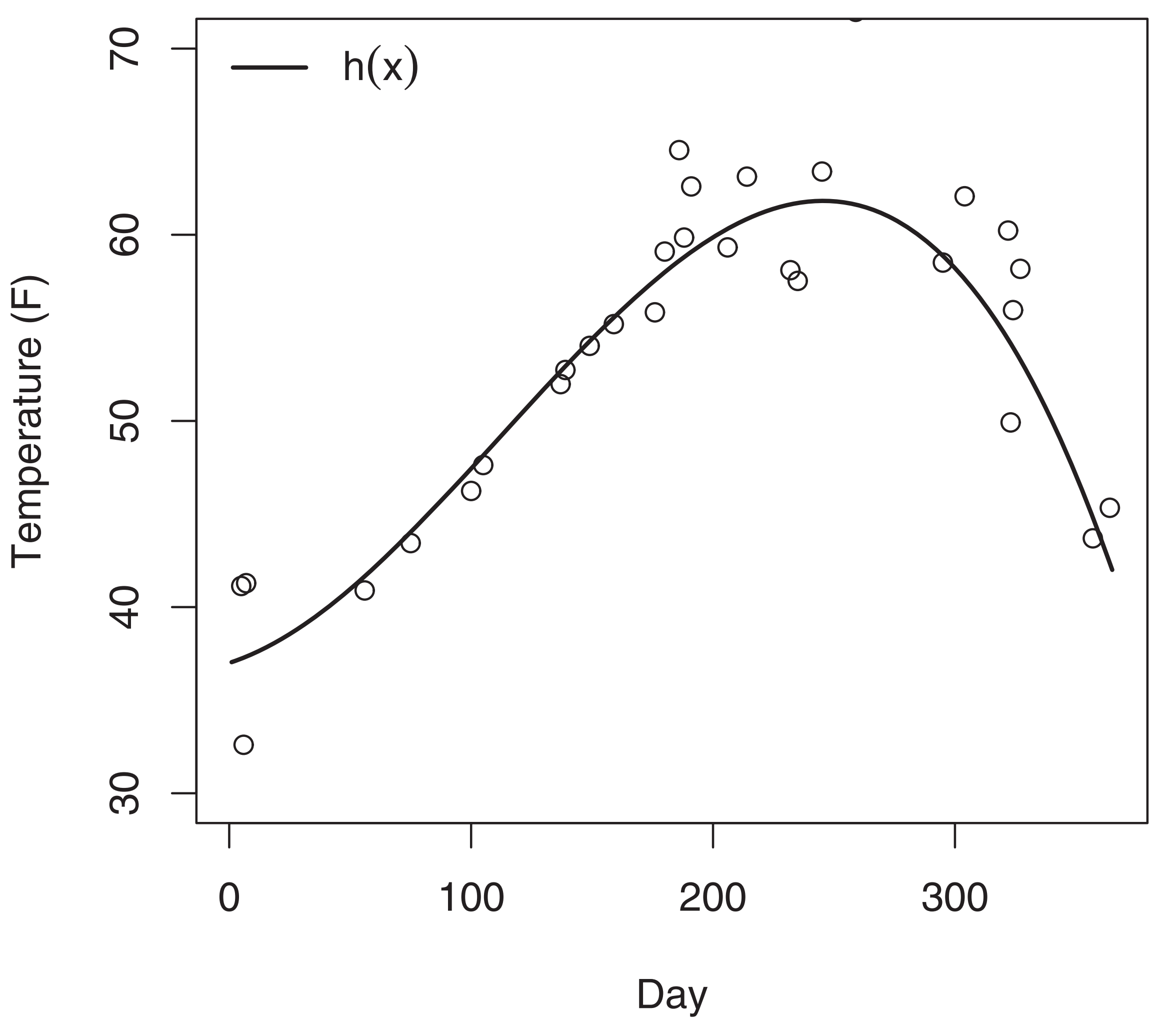


(B)

Achieving a good fit to observations does not necessarily mean we have found a good model, and choosing the model with the best fit is likely to result in poor predictions.

Example: Daily temperature series

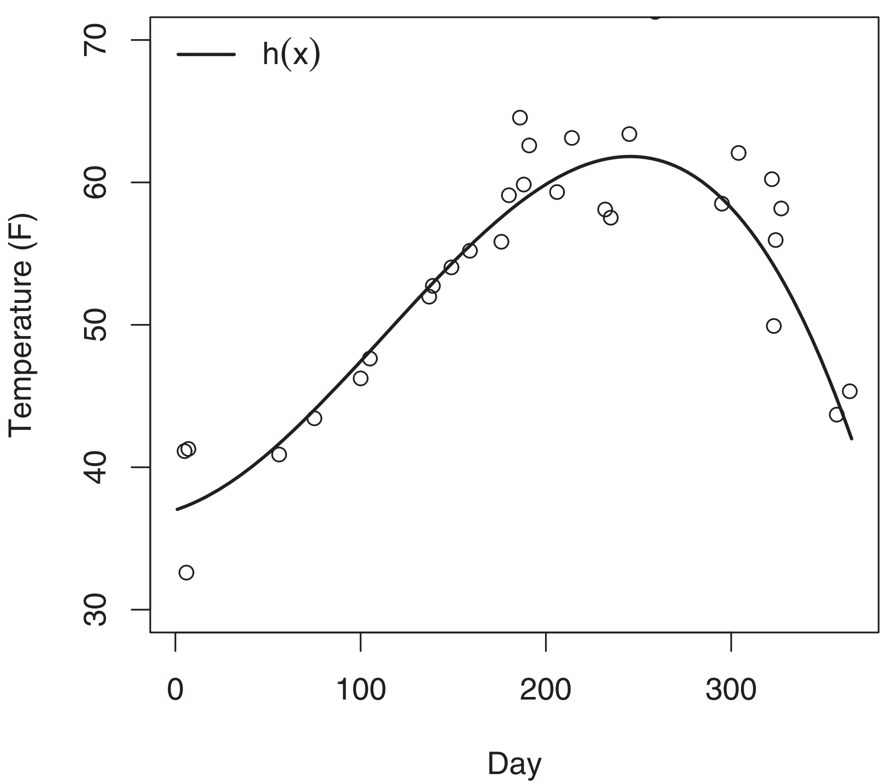
Study model fit with simulations



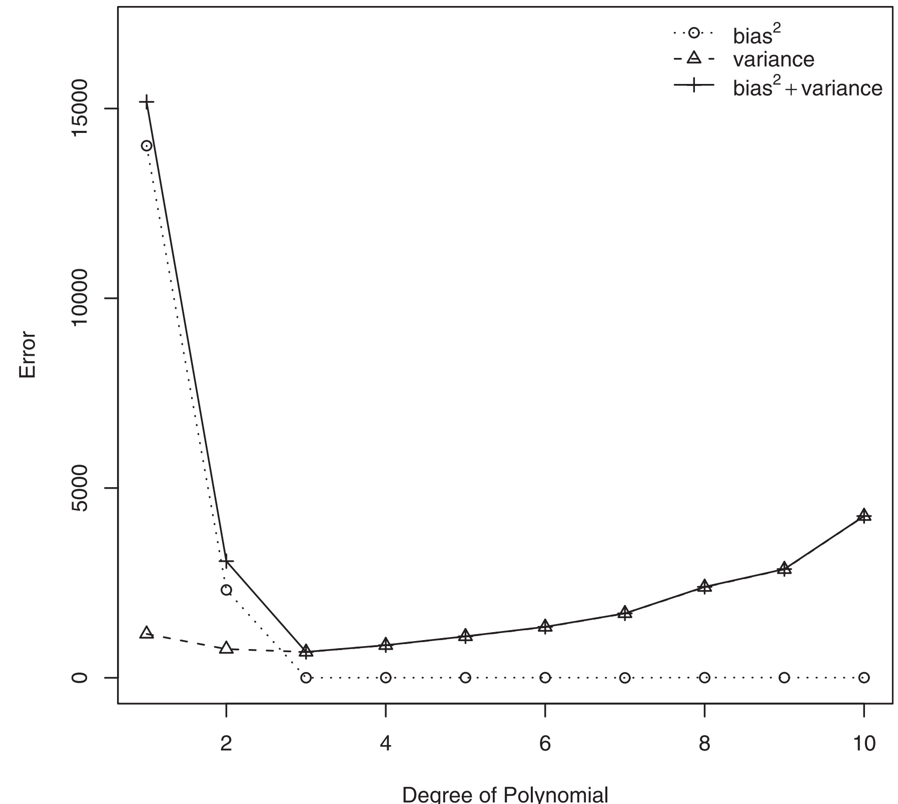
(A)

Underlying temperature pattern for some fictional location, along with a random sample of 30 observations with added noise.

Example: Daily temperature series

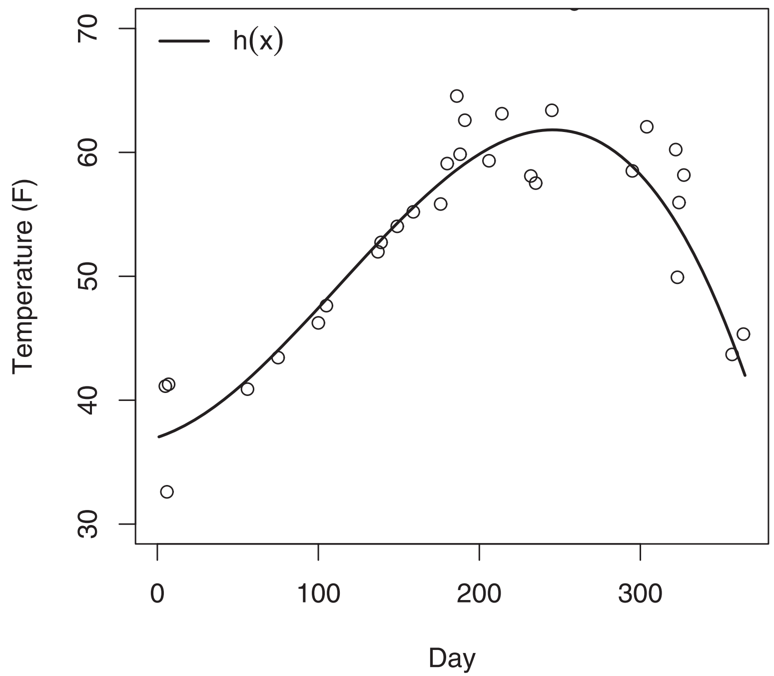


(A)

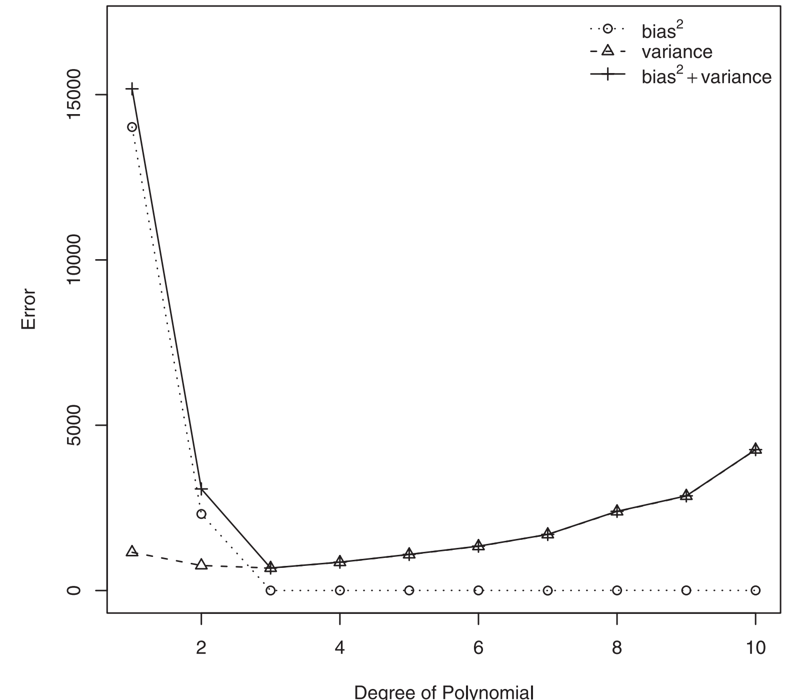


(B)

Plot (A) shows the underlying temperature pattern for some fictional location, along with a random sample of 30 observations with added noise. Plot (B) shows, as a function of degree of polynomial, the mean error in predicting the population after fitting polynomials to samples of 30 noisy observations. This error is decomposed into bias and variance, also plotted as function of degree of polynomial.



(A)



(B)

Degree-3 polynomial achieves the lowest mean prediction error on this problem. Polynomials of degree-1 and degree-2 lead to significant estimation bias because they lack the ability to capture the underlying cubic function h(x) and will therefore always differ from the underlying function. Unbiased models are those of degree-3 or higher, but notice that the higher the degree of the polynomial, the greater the prediction error. The reason why this behavior is observed is that higher degree polynomials suffer from increased variance due to their greater flexibility. The more flexible the model, the more likely it is to capture not only the underlying pattern but unsystematic patterns such as noise. Recall that variance reflects the sensitivity of the induction algorithm to the specific contents of samples, which means that for different samples of the environment, potentially very different models are being induced. Finally, notice how a degree-2 polynomial achieves a lower mean prediction error than a degree-10 polynomial.

Question: Is more better?

Daily temperature example has demonstrated:
A biased model can lead to more accurate predictions than an unbiased model.

Illustrates a fundamental problem in statistical inference known as the ***bias–variance dilemma*** (Geman et al., 1992)

Diversity in the class of patterns that the model can accommodate is, however, likely to come at a price. The price is an increase in variance, as the model will have a greater flexibility, will enable it to accommodate not only systematic patterns but also accidental patterns such as noise. When accidental patterns are used to make predictions, these predictions are likely to be inaccurate.

Dilemma: Combating high bias requires using a rich class of models, while combating high variance requires placing restrictions on this class of models.

This is why "general purpose" models tend to be poor predictors of the future when data are sparse.

Implications: Cognitive processes

- cognitive system performs remarkably well when generalizing from few observations, so much so that human performance is often characterized as optimal (e.g., Griffiths & Tenenbaum, 2006; Oaksford & Chater, 1998).
- ability of the cognitive system to make accurate predictions despite sparse exposure to the environment strongly indicates that the variance component of error is successfully being kept within acceptable limits
- to control variance, one must abandon the ideal of general-purpose inductive inference and instead consider, to one degree or another, specialization

Bias–variance dilemma shows formally why a mind can be better off with an adaptive toolbox of biased, specialized heuristics.

A single, general-purpose tool with many adjustable parameters is likely to be unstable and to incur greater prediction error as a result of high variance.

Example: Hot hand vs gambler's fallacy

Gambler's fallacy: ○ ○ ○ ○ ○ ➔ ●

Hot hand fallacy: ○ ○ ○ ○ ○ ➔ ○

The gambler's fallacy and the hot hand fallacy are opposite intuitions: After a series of n similar events, the probability of the opposite event increases (gambler's fallacy), and after a series of similar events, the probability for same event increases (hot hand fallacy).

Both have been explained by the label "representativeness"

Sedlmeier, Hertwig, and Gigerenzer (1998)

Approach: Less is more

Ignoring cues, weights, and dependencies between cues.

- ove

R

Approach: Less is more

Ignoring cues, weights, and dependencies between cues.

- ove

R

Approach: Less is more

Ignoring cues, weights, and dependencies between cues.

- ove

R

Example: Should Darwin become a husband?

MARRY

Children—(if it please God)—constant companion, (friend in old age) who will feel interested in one, object to be beloved and played with—better than a dog anyhow—Home, and someone to take care of house—Charms of music and female chit-chat. These things good for one's health. Forced to visit and receive relations *but terrible loss of time.*
My God, it is intolerable to think of spending one's whole life, like a neuter bee, working, working and nothing after all.—No, no won't do.—
Imagine living all one's day solitarily in smoky dirty London House.—
Only picture to yourself a nice soft wife on a sofa with good fire, and books and music perhaps—compare this vision with the dingy reality of Grt Marlboro' St.

Not MARRY

No children, (no second life) no one to care for one in old age. . . . Freedom to go where one liked—Choice of Society *and little of it.* Conversation of clever men at clubs.—Not forced to visit relatives, and to bend in every trifle—to have the expense and anxiety of children—perhaps quarrelling.
Loss of time—cannot read in the evenings—fatness and idleness—anxiety and responsibility—less money for books etc—if many children forced to gain one's bread.—(But then it is very bad for one's health to work too much)
Perhaps my wife won't like London; then the sentence is banishment and degradation with indolent idle fool—

(Darwin, 1887/1969, pp. 232–233)

Darwin concluded that he should marry, writing "Marry—Marry—Marry Q. E. D." decisively beneath the first column. On the reverse side of the page he considered the consequences of his decision for his personal freedom, ending with the insight: "There is many a happy slave." The following year, Darwin married his cousin, Emma Wedgwood, with whom he eventually had 10 children. How did Darwin decide to marry, based on the possible consequences he envisioned—children, loss of time, a constant companion? He did not tell us. But we can use his "Question" as a thought experiment to illustrate various visions of rationality.

Imagine that Darwin had attempted to resolve his Question by maximizing his subjective expected utility. To compute his personal expected utility for marrying, he would have had to determine *all* the possible consequences that marriage could bring (e.g., children, constant companion, and an endless stream of further possibilities not included in his short list), attach quantitative probabilities to each of these consequences, estimate the subjective utility of each consequence, multiply each utility by its associated probability, and finally add all these numbers up. The same procedure would have to have been repeated for the alternative "not marry." Finally, he would have had to choose the alternative with the higher total expected utility. To acquire reliable information about the consequences and their probabilities and utilities, Darwin might have had to invest years of research—time he could have spent studying barnacles or writing *Origin of Species*.

E: P

Go

- ove

R

E: P

Go

- ove

R

E: P

Go

- ove

R

Applications: Financial regulation

FFT to determine risk of bank failure

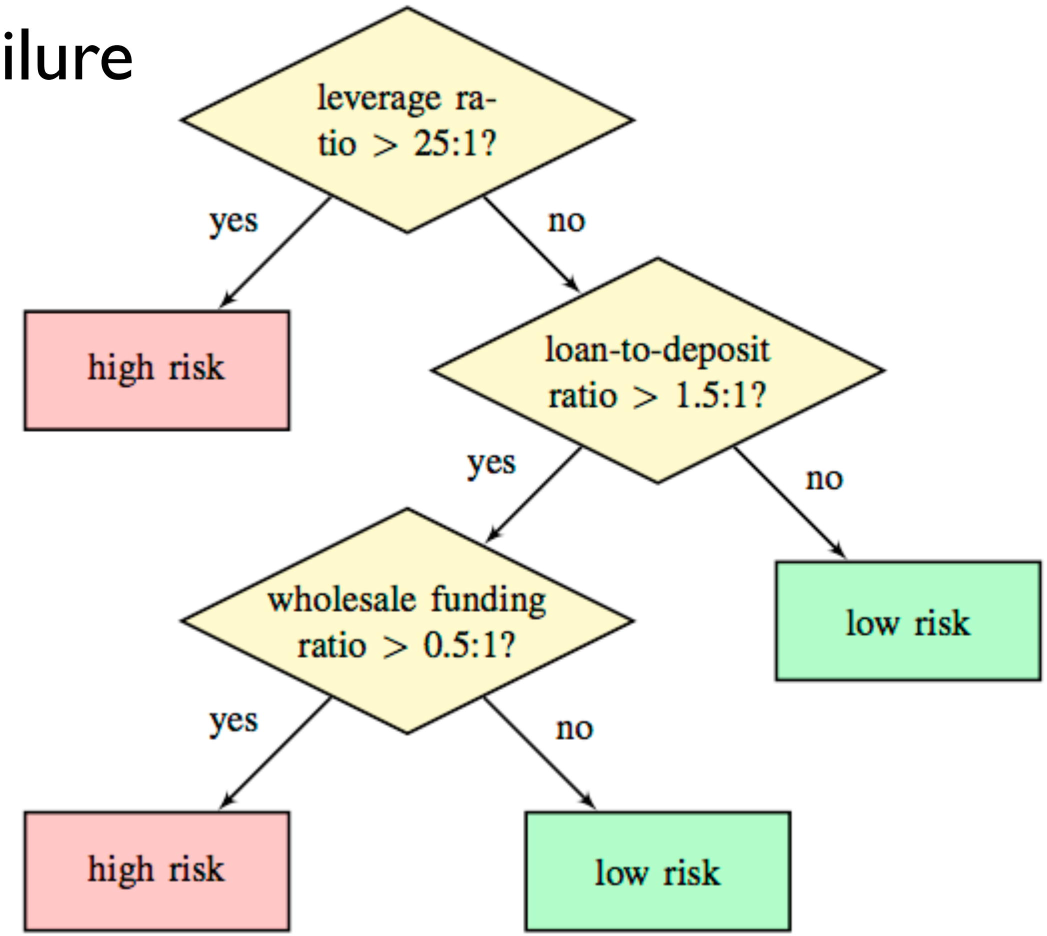


Figure 1. Example of a fast and frugal tree (FFT) to determine the risk of bank failure. (Threshold values are merely illustrative.)

Neth, H., Meder, B., Kothiyal, A. & Gigerenzer, G. (2014). *Homo heuristicus* in the financial world: From risk management to managing uncertainty. *Journal of Risk Management in Financial Institutions*, 7(2), 134–144.

On information

"It simply wasn't true that a world with almost perfect information was very similar to one in which there was perfect information."

J. E. Stiglitz (2010). Freefall: America, free markets, and the sinking of the world economy, p. 243

Joseph Eugene Stiglitz, ForMemRS, FBA
American economist, professor at Columbia University
Nobel Memorial Prize in Economic Sciences, John Bates Clark Medal

Similar estimates across the literature

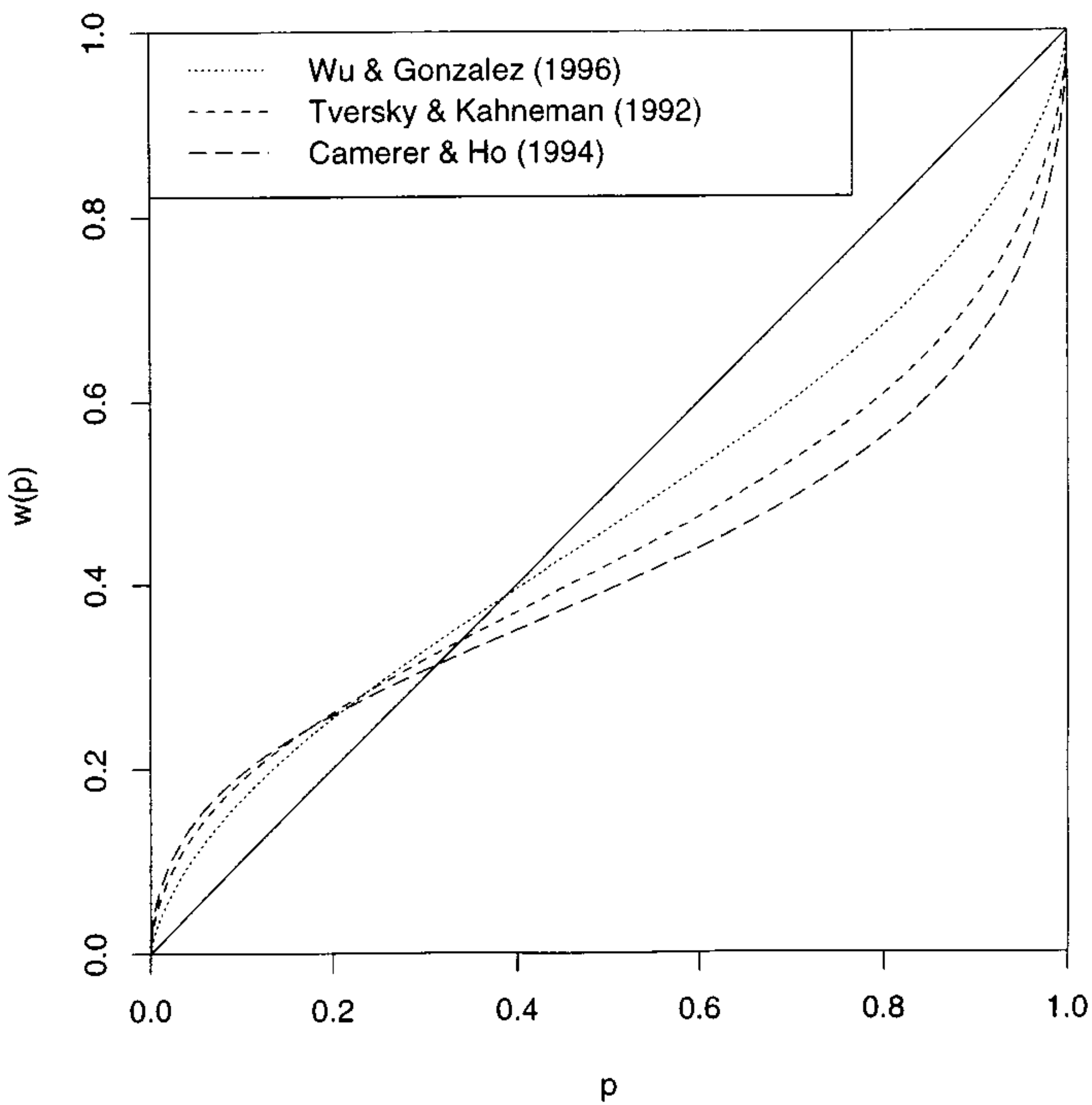


FIG. 2. One-parameter weighting functions estimated by Camerer and Ho (1994), Tversky and Kahneman (1992), and Wu and Gonzalez (1996) using $w(p) = (p^{\beta}/(p^{\beta} + (1 - p)^{\beta})^{1/\beta})$. The parameter estimates were .56, .61, and .71, respectively.

Concepts: PT utility function (value function)

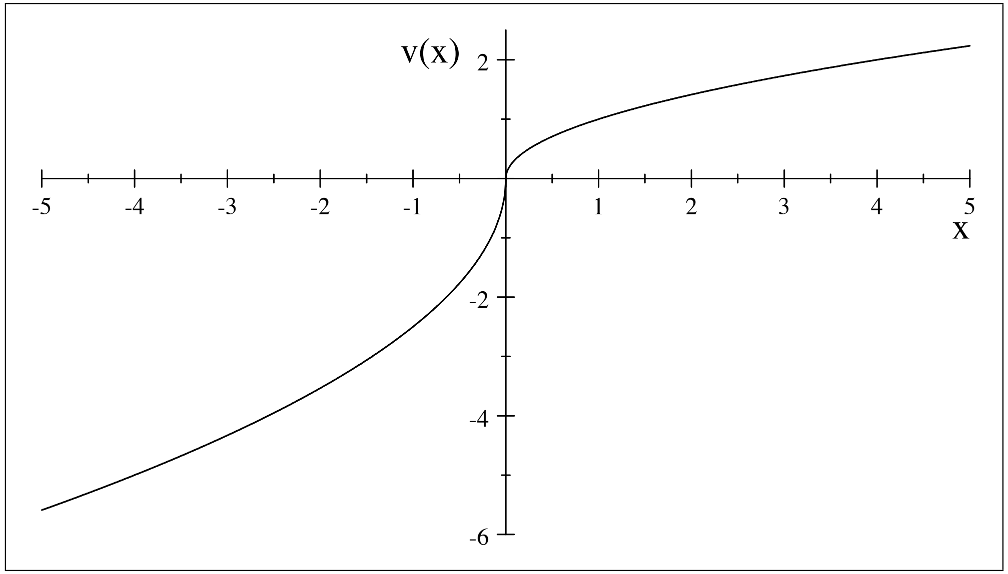
Definition *(Tversky and Kahneman, 1979). A utility function, $v(x)$, is a continuous, strictly increasing, mapping $v : \mathbb{R} \to \mathbb{R}$ that satisfies:*

1. $v(0) = 0$ (reference dependence).

2. $v(x)$ is concave for $x \geq 0$ (declining sensitivity for gains).

3. $v(x)$ is convex for $x \leq 0$ (declining sensitivity for losses).

4. $-v(-x) > v(x)$ for $x > 0$ (loss aversion).



$$v(x) = \begin{cases} x^\alpha & x \geq 0 \\ -\lambda(-x)^\beta & x < 0 \end{cases}$$

$\alpha > 0$: degree of risk aversion in gains

$\beta > 0$: degree of risk seeking in losses

$\lambda > 0$: degree of loss aversion

Normative Theories

Purpose of the normative theories is to express, how people *should* behave when they are confronting risky decisions. Thus the behavioral models based on EUT stresses the rationality of decisions. We are not interested so much on how people behave in real life or in empirical experiments. Notice that one of these theories is the SEUT. Furthermore, the EUT can be also defended on grounds that it works satisfactory in many cases.

Descriptive Theories

From the descriptive point of view, we are concerned with *how* people make decisions (rational and not rational) in real life. The starting point for these theories has been in empirical experiments, where it has been shown that people's behavior is inconsistent with the normative theories. These theories are, for example, prospect theory and regret theory. We will present some of them more precisely in the next sections.

Prescriptive Point of View

Prescriptive thinking in risky decisions means that our purpose is to *help* people to make good and better decisions. In short, the aim is to give some practical aid with choices to the people, who are less rational, but nevertheless aspire to rationality. This "category" includes, for example, operation research and management science.

http://epublications.uef.fi/pub/urn_isbn_978-952-458-985-7/urn_isbn_978-952-458-985-7.pdf