Behavior and Healthier Lifestyle: Methods

Running experiments and eliciting value

Paolo Crosetto

Behavioral Bridge Summer School – Valencia, July 2026

Experiments in economics: a primer

How to evaluate a behavioral policy

You want to make a policy intervention: *make people respect traffic rules more*. You have several options, and you don't know *ex-ante* which one is the best. How do you proceed?

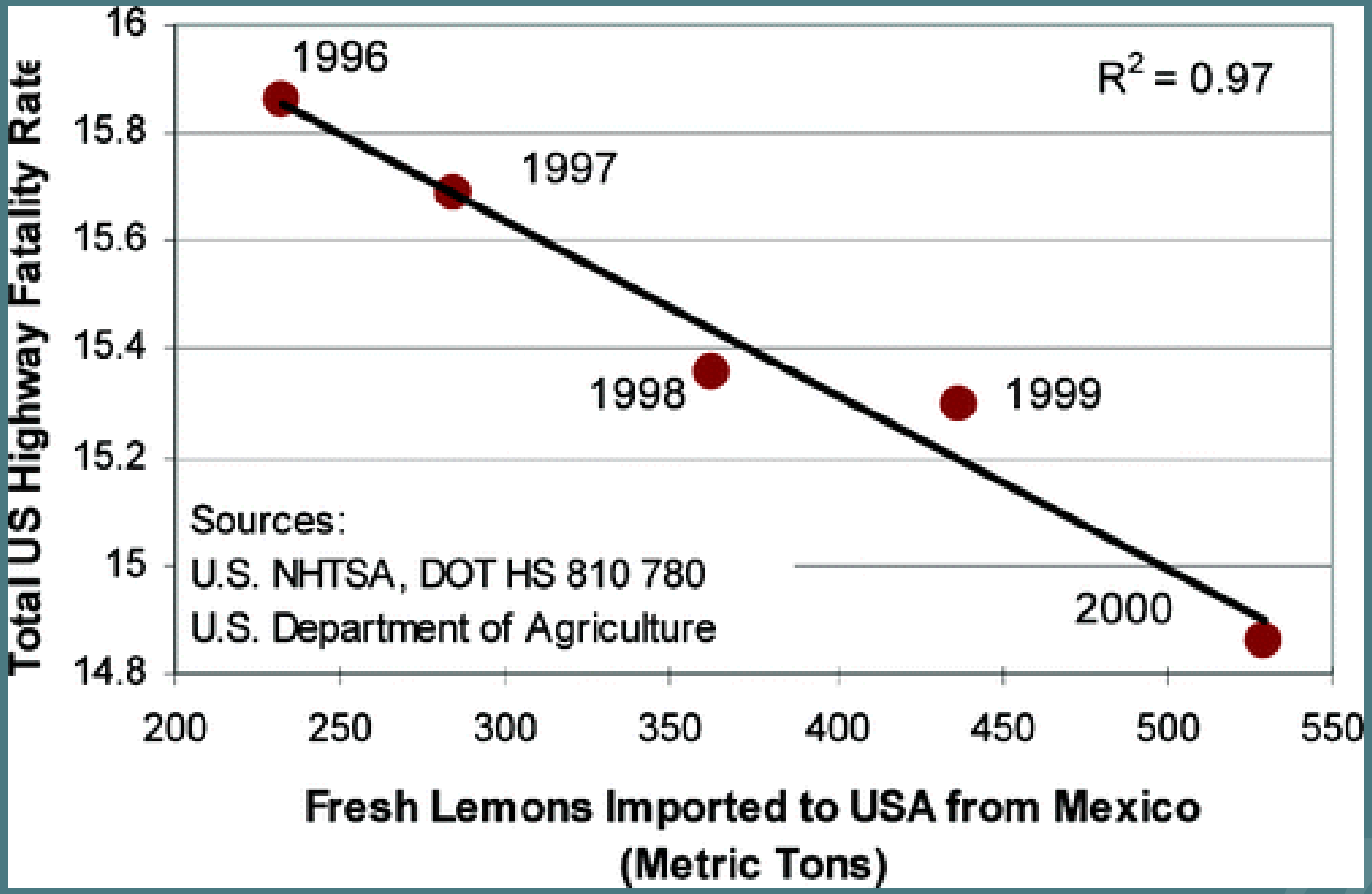
- Use theor(ies). Derive expectations from theories & parameters.
- Use data. Observe behavior, record data, use statistical techniques to fit the data.

Can you really say that intervention X caused effect Y?

Correlation $\neq$ causation

$A \sim B$ could imply:

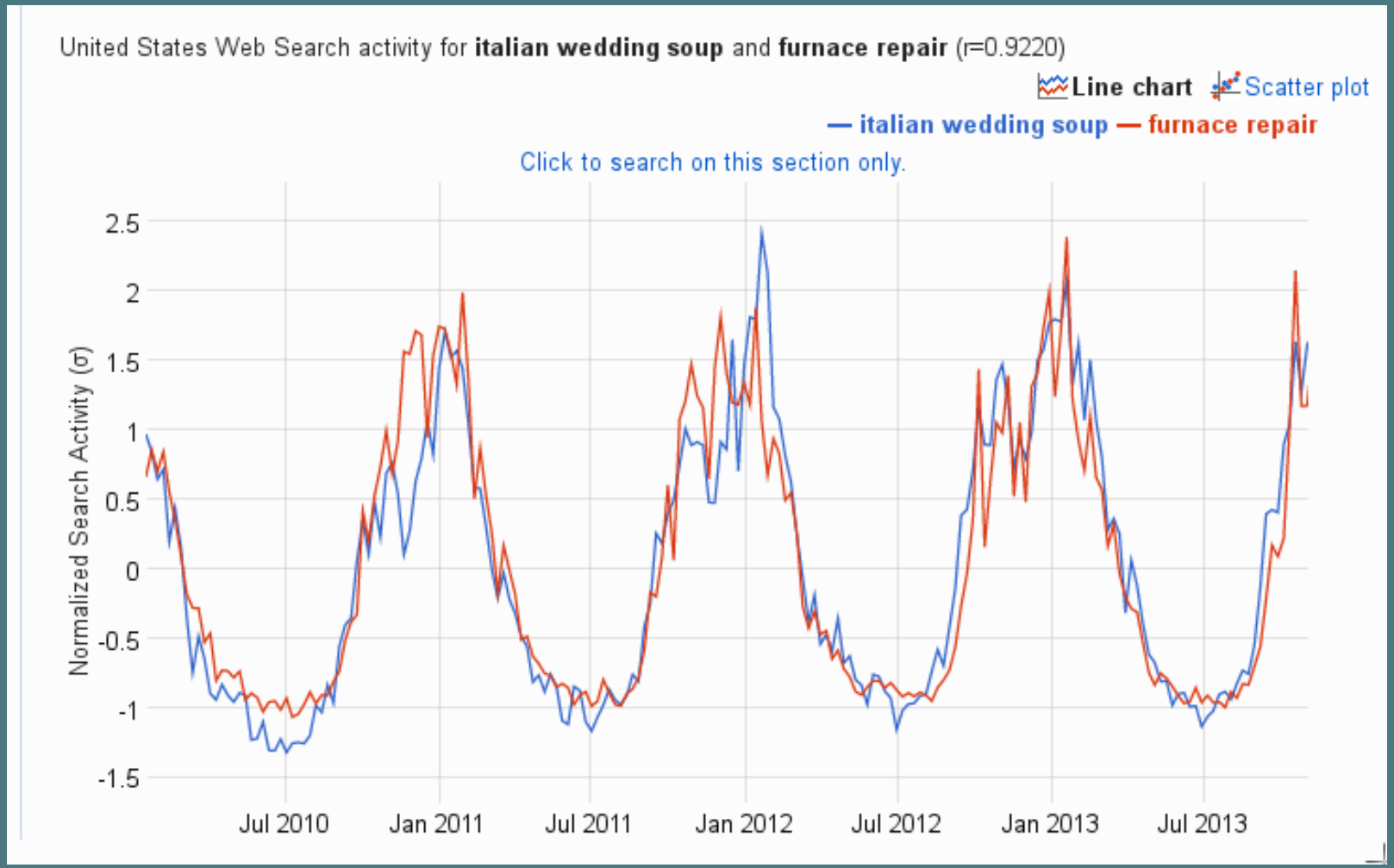
- no causality: a spurious correlation
- **reverse** causality: $B \Rightarrow A$
- a common **unobserved** cause: $C \Rightarrow A$ and $C \Rightarrow B$ but $A \not\Leftrightarrow B$
- causality! $A \Rightarrow B$



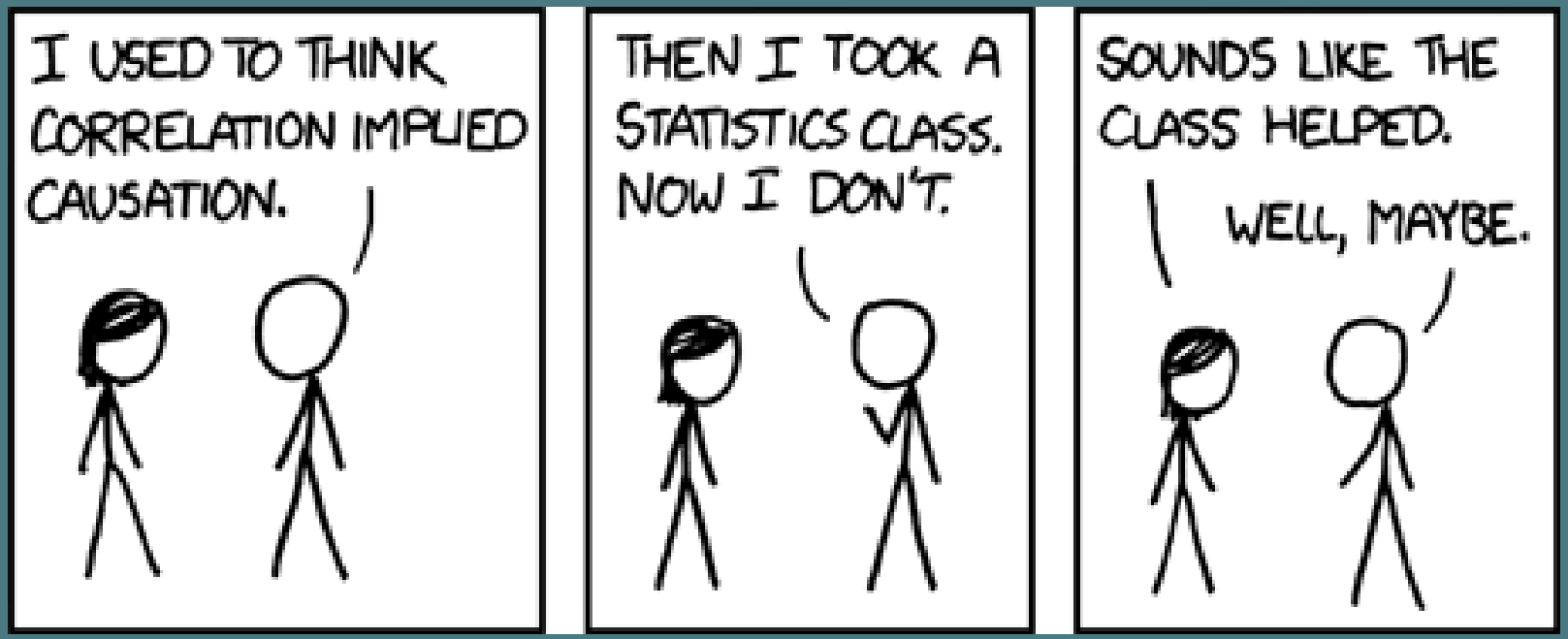
— A spurious correlation

In Italy and the United Kingdom, for example, where **lockdowns** *have been repeatedly imposed,* **death** *totals per million remain among the worst in the world. Meanwhile, in the United States, states with the most* **harsh lockdown rules***–such as New York, New Jersey, and Massachusetts are among the states with the* **worst total deaths***.*

— A bizarre claim of reverse causality



— A common unobserved cause



— Causation $\neq$ Correlation – XKCD

Experiments as the workhorse
for unearthing causal relations in science

Experiments, A to Z

- **Observe** and measure the penomenon of interest
- **Formulate** *testable hypotheses*
- **Control** all possible confounding factors
- *Ceteris paribus*: keep *all* things equal, but *one*
- Systematically **vary** the *one* variable of interest
- **Compare** results with hypothesis: rejected or not?
- **Compare** results to the real world: do they matter?

If the world does not accord with you, you are wrong.

Experiments: control + relevance

In experiments we need control + relevance

Control (or internal validity)

all possible confounds are controlled for, and we successfully isolate the phenomenon of interest.

Are you actually testing your hypothesis?

Relevance (or external validity)

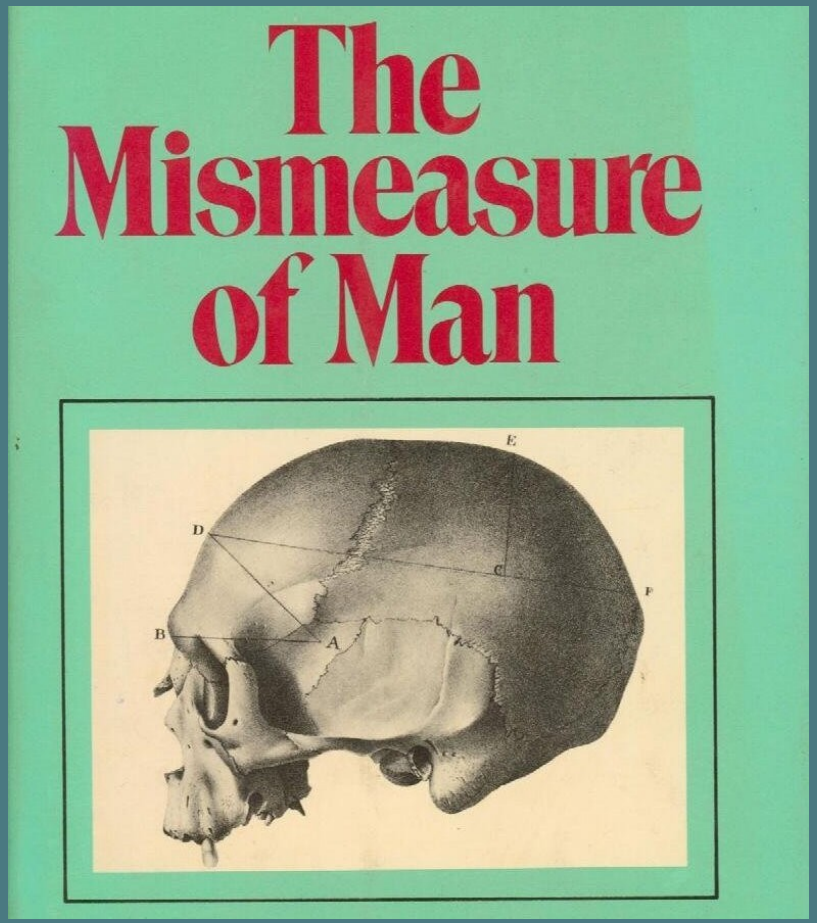
our experimental manipulation does not alter reality so much that our results do not hold *outside of the lab*.

Does your test actually matter?

Experiments' checklist: Measure

Observe and measure the penomenon of interest

- Observation must be *objective* [it's 9°C today *vs.* it's cold]
- Measure must follow some standards (loi Carrez)
- Both observation and measure must be **replicable**



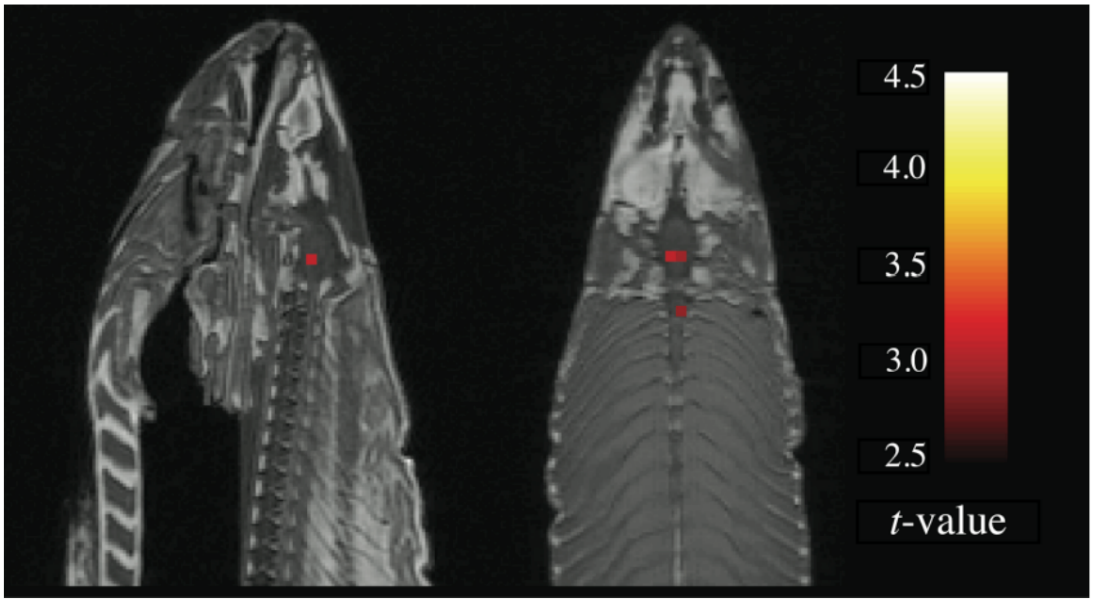
— Stephen Jay Gould and measure biases

Formulate testable hypotheses

- Testable hypotheses must be apt to be *proven wrong*: **falsifiable**.
- You should lay them down *before* running the experiment...
- ...because you will **always** find something...

“

GLM RESULTS

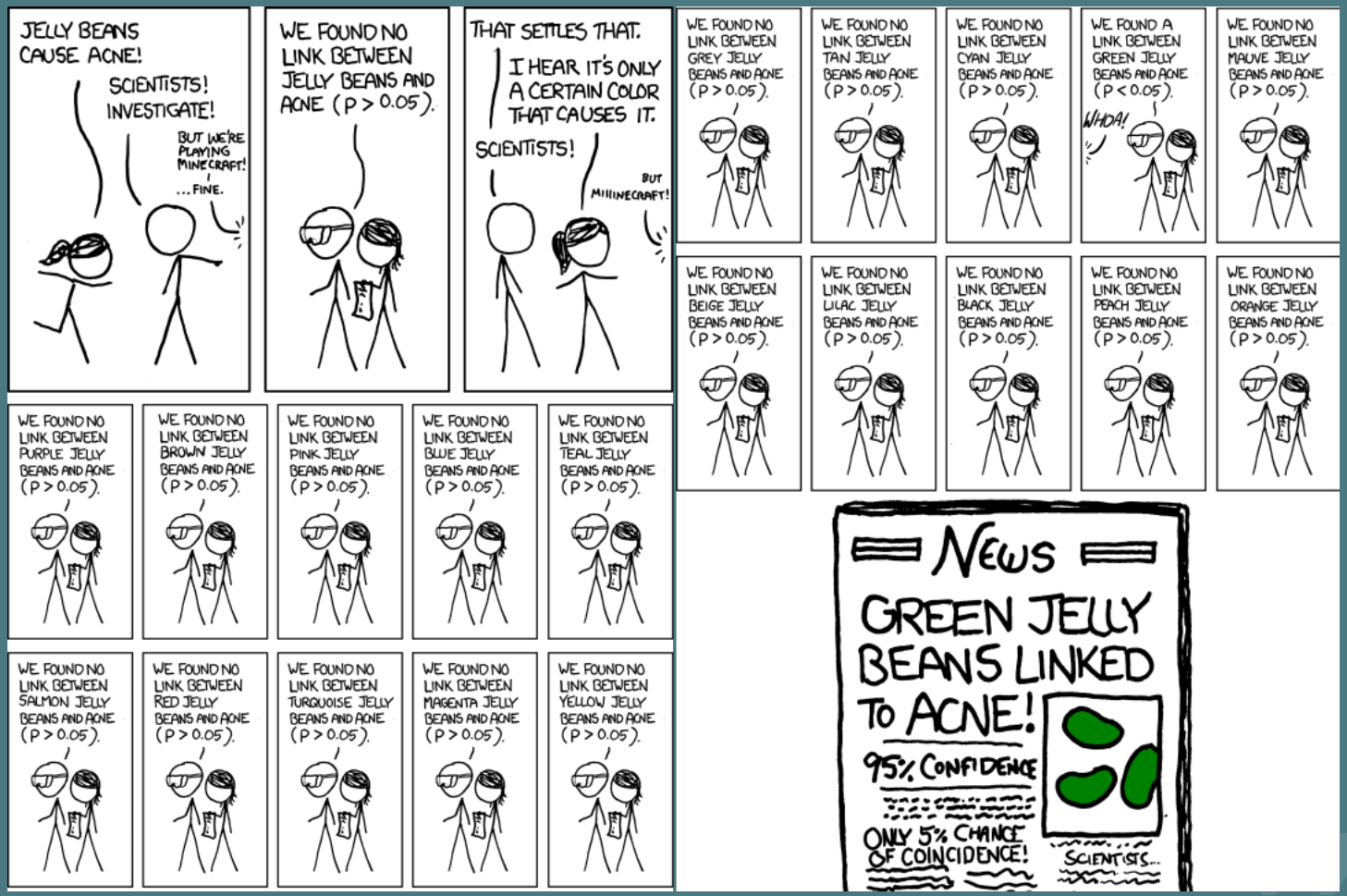


A *t*-contrast was used to test for regions with significant BOLD signal change during the photo condition compared to rest. The parameters for this comparison were $t(131) > 3.15$, p(uncorrected) < 0.001, 3 voxel extent

”

— You *always* find something: dead salmon

— You *always* find something: infinite billiard



— You *always* find something: N hypotheses

Experiments checklist: Control

Control all possible confounding factors

- Factors influencing your dep var **other** than the one you are studying.
- Their presence can make your results *spurious*

Controlling all confounding factors means reaching internal validity

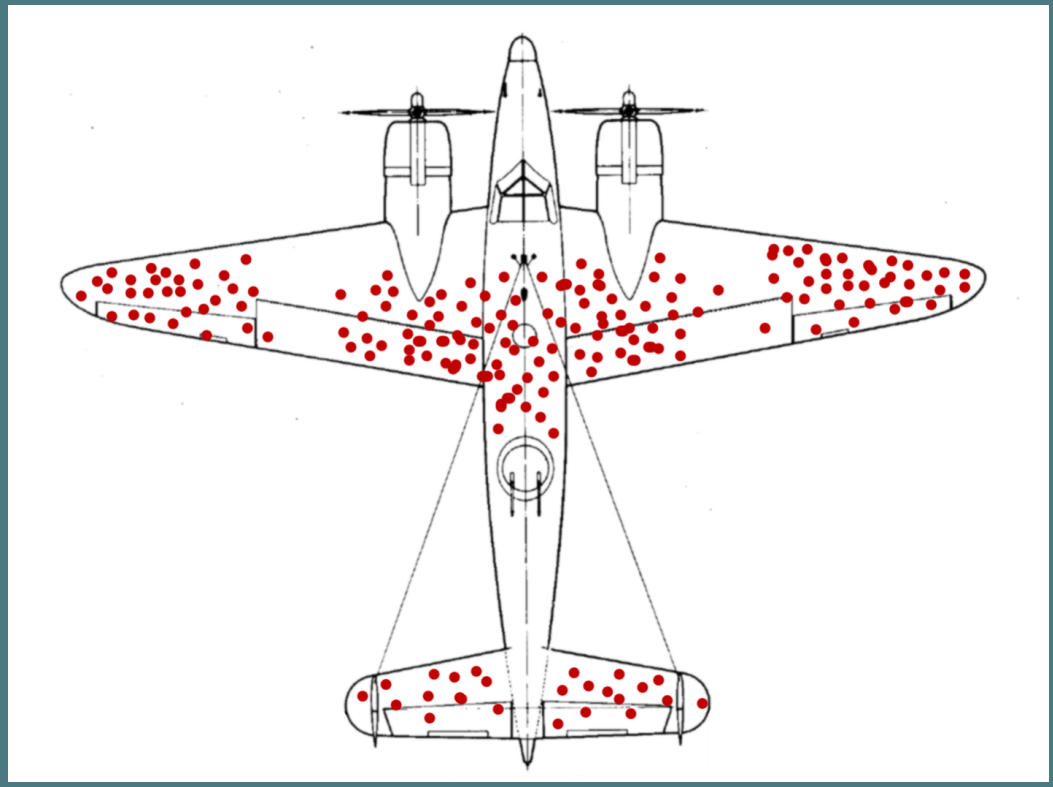
Internal validity is the basic minimum without which any experiment is uninterpretable: Did in fact the experimental treatments make a difference in the experimental instance?

Biases introduced in your experiment by the fact that the subjects act in accordance to what they believe the experimenter wants of them, 'the right thing', 'the good thing'.

— Experimenter Demand Effect

Biases introduced in your experiment by the fact that the small (or null) monetary values present in the experiment elicit qualitatively different behavior than the real-world high stakes

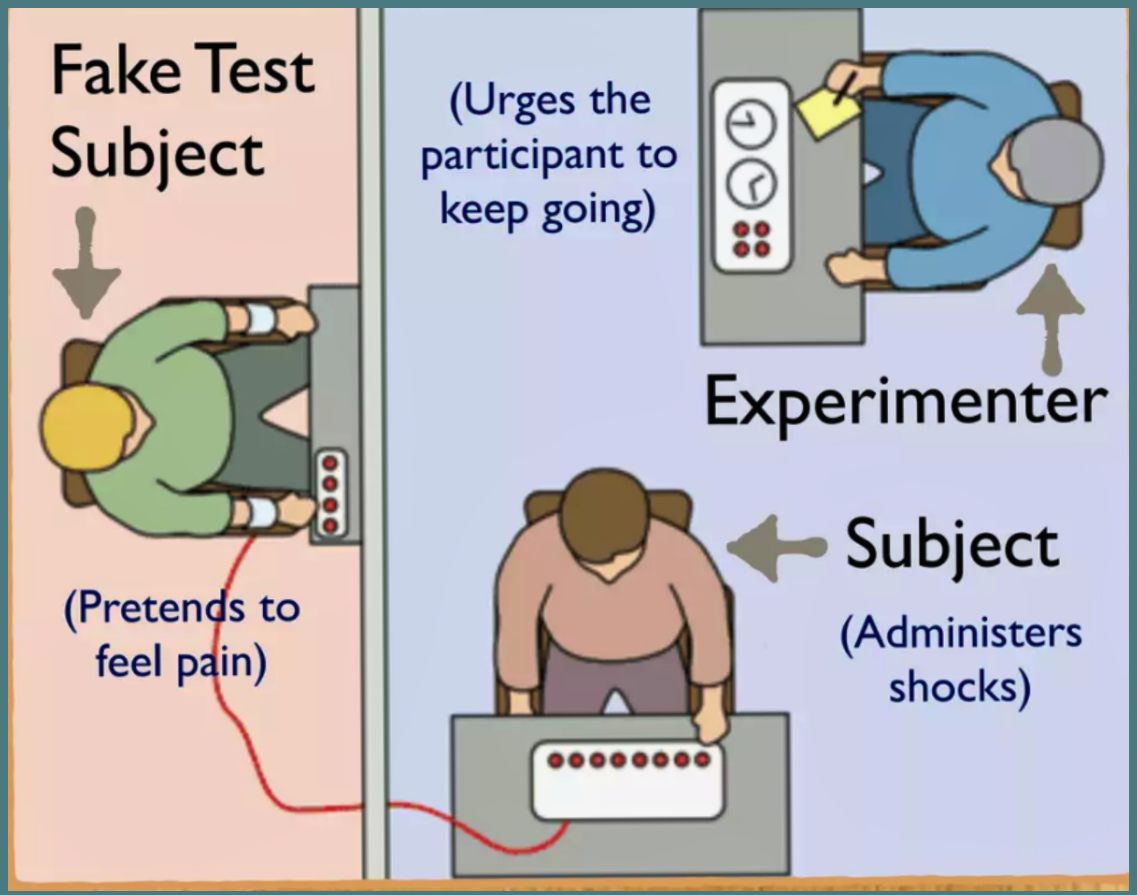
— Incentive effects



— Where would you add armor?

Selection bias is the selection of individuals, groups or data for analysis in such a way that proper randomization is not achieved, thereby ensuring that the sample obtained is not representative of the population intended to be analyzed.

— Sample selection effects – Wikipedia



— Credibility and interpretation: the Milgram study

Ceteris paribus: keep all things equal, but one

- This means *all* things
- Information; time of the day; room temperature; number of people; ...
- But also: awareness of being observed (placebos!)

But can you really control everything? No, but you can randomise

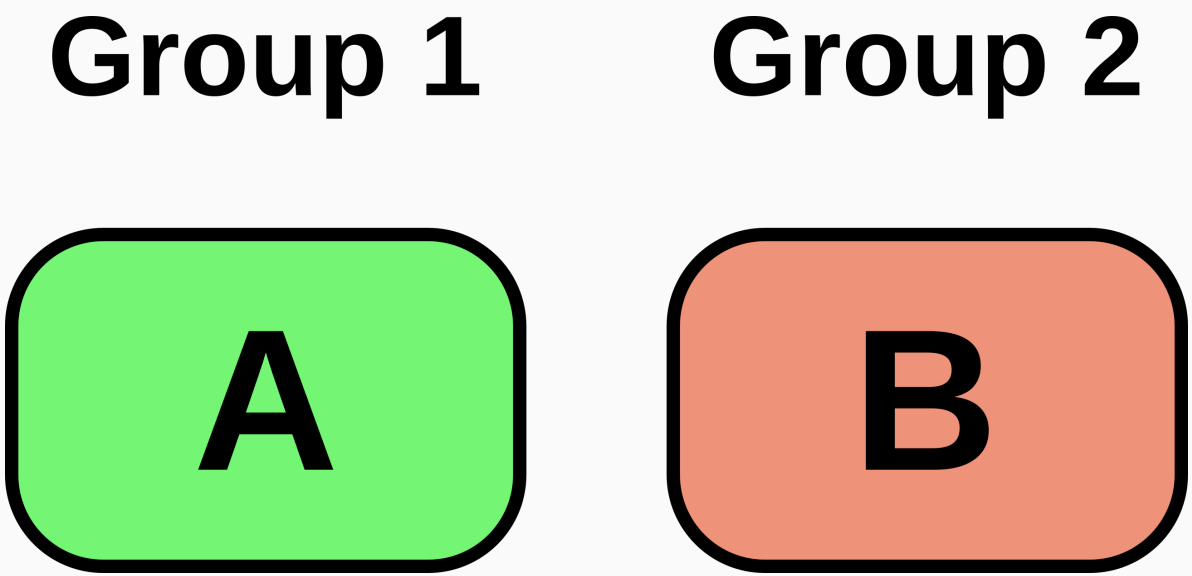
- Some factors can be controlled
- Most others are randomized: you let them vary *randomly* and rely on a high number of observations to smooth things out
- As $N \longrightarrow \infty$, random factors cancel each other out

Experiments checklist: Treatments

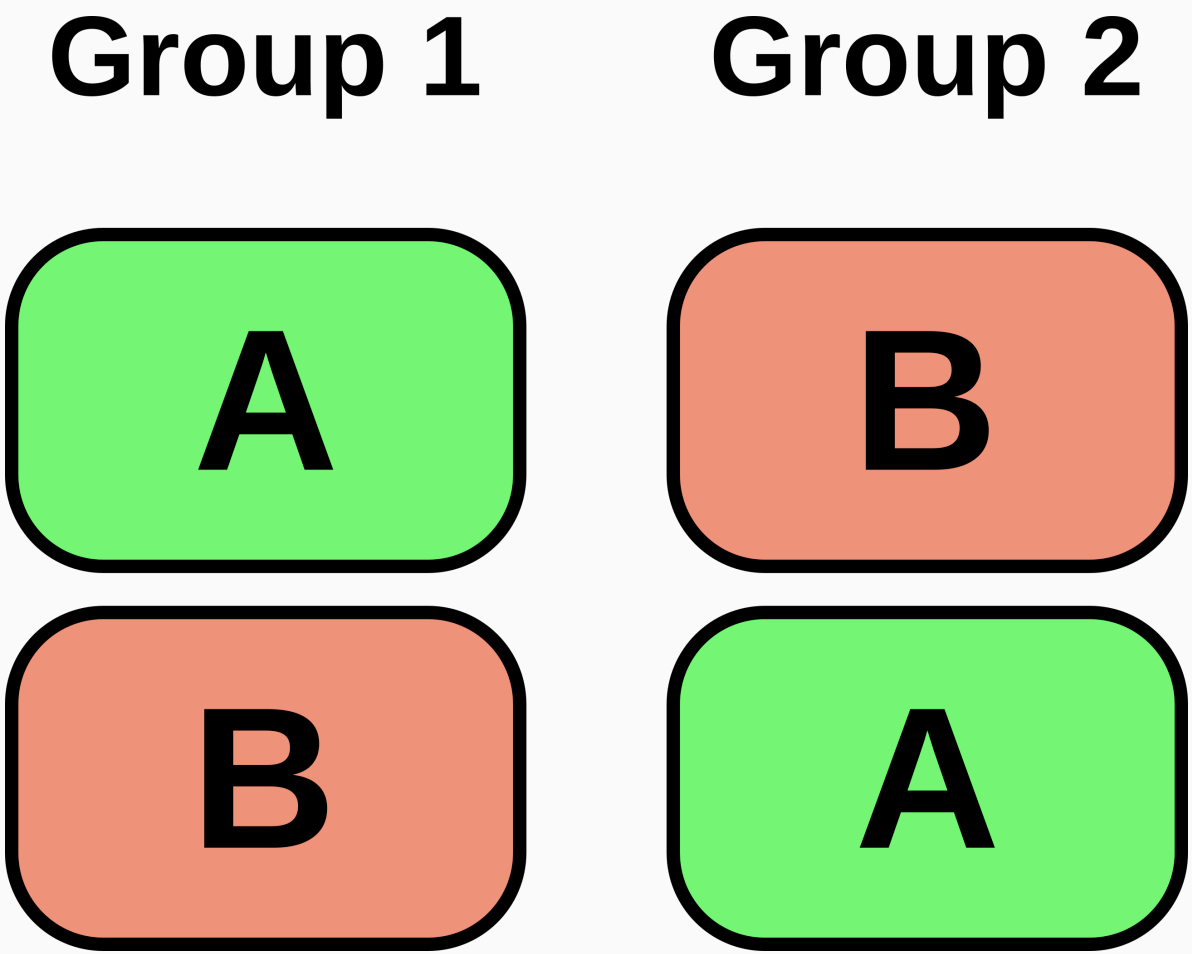
Systematically vary the one variable of interest

- These are what we call *Treatments*
- They can be varied within or between subjects
- Between: different subjects are subjected to different treatments (needs high N)
- Within: same subject is exposed to all treatments (needs less data, order effects)

Treatments: Between-subjects



Treatments: Within-subjects



Experiments checklist: Inference

Compare results with hypothesis: rejected or not?

- Statistical testing can tell us if two samples are different (in mean, median, distribution, count...)
- Regression analysis or ANOVA can take care of multivariate analyses (control for common confounds as age, gender...)

Hypotheses are not proved true; they are not proven wrong

- Your tests tell us that you fail to reject the hypothesis.
- This does not mean it is true; it just means it is not random...
- ...and it means the hypothesis is **supported by the data**.
- other data might prove it wrong; or it might not always apply; or...

Good scientists try as hard as they can to prove themselves wrong

Definition (External Validity)

In scientific research, external validity is the degree to which it is warranted to generalize results to other contexts.[source: Wikipedia]

How can you get high external validity?

The Case for Laboratory Experiments in Behavioural Public Policy (Peter D Lunn and Áine Ní Choisdealbha)

How to get external validity: get near to theory

How do you get external validity?

Get as far away from reality as possible and near to theory

- theory is generalizable by default
- so the closer your expeirment is to theory, the higher the chances of falsifying it
- if you fail, theory survives and you can apply it to further contextualized settings
- because imagine theory fails in *one* little setting
- then it can be because of some confounds, some external influence, while theory is still valid
- if the test is non-contextual then it is in principle stronger

How to get external validity: get near to the real world & replicate

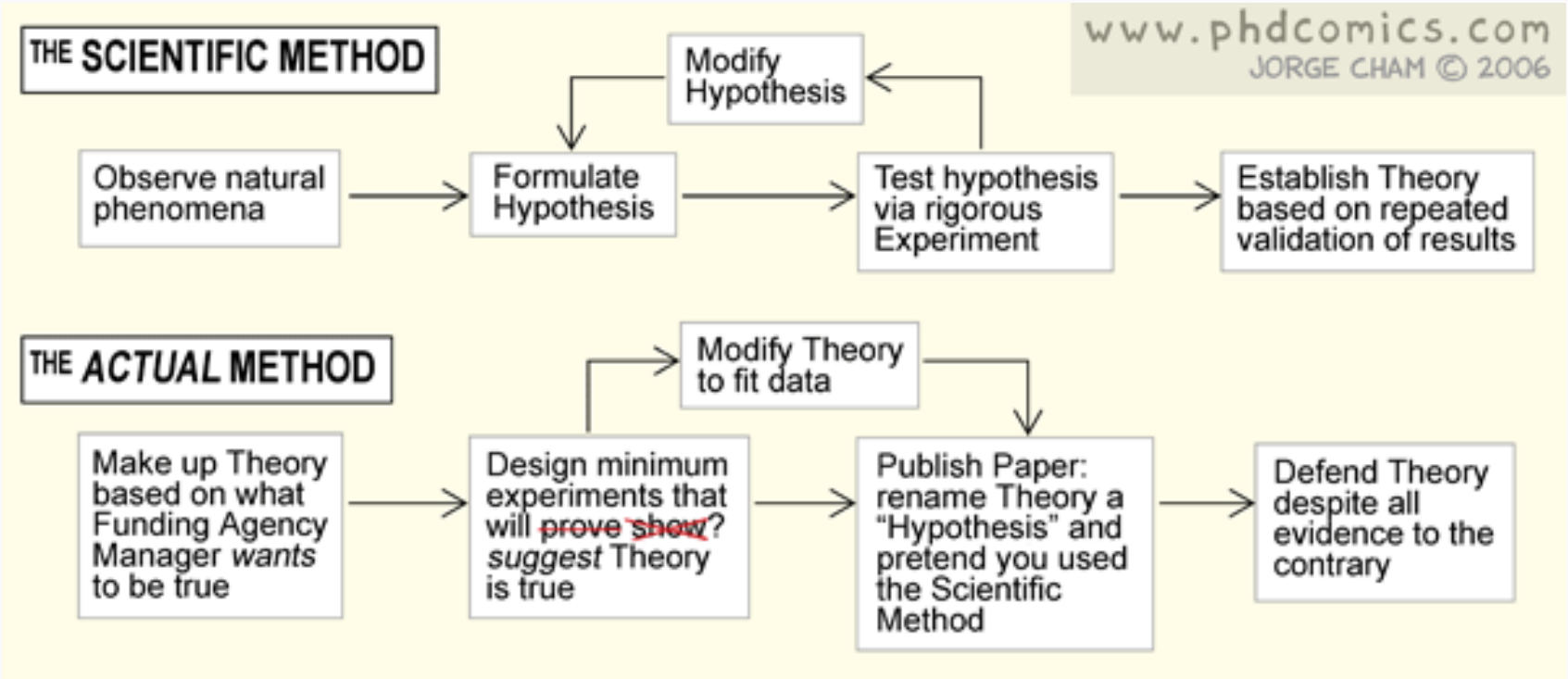
How do you get external validity?

Get as much real context as possible, and then replicate in several different contexts

- running an experiment in the field, within the target population, fully contextualized
- makes sure the result is valid in *that specific field*
- *but* leaves open the possibility that it will *not* replicate in *other* contexts
- from which we deduce the need for replication

That's the theory. But in practice?

Experiments: theory vs reality



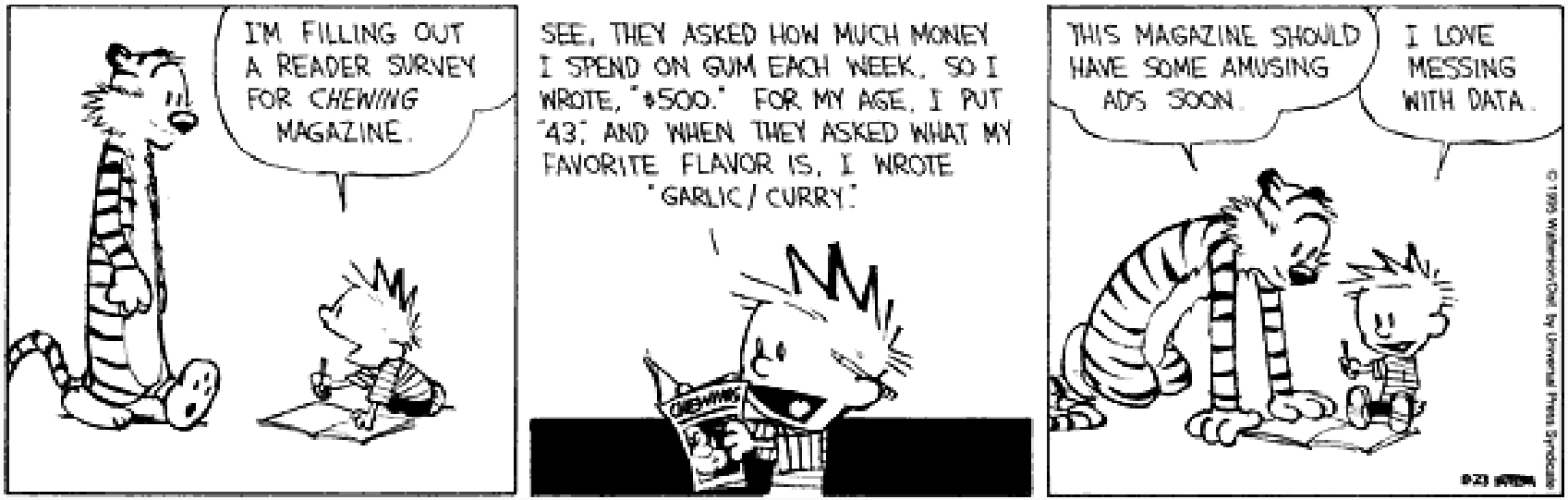
Some peculiarities of economic experiments

Experimental Economics in a nutshell

ExEc is the use of the experimental method in economics

- Usually relies on simple abstraction of real problems (k.i.s.s – *keep it simple, stupid!*)
- Usually does *not* trust subjects to tell the truth unless properly incentivized to do so
- Usually relies on giving full knowledge to participants
- Never lies and always tells subjects the truth, *never deceiving them*

Experimental Economics: incentives



Problem hypothetical bias

- Usually self-declared price $>$ incentivized price
- The second decision is called **incentive-compatible**
- i.e., you have real consequences, and your best strategy is to reveal your true preferences

We go at great length to create incentive-compatible mechanisms

Experimental Economics: common knowledge

Rules of the game are common knowledge

- As in the previous example, all was clear *before* we set up the sale
- No surprises, no changes, no confederates (i.e. false participants)
- Usually instructions are read aloud to make sure that not only everyone gets the same
- But that everybody knows that everyone has heard the same (common knowledge)

Common knowledge levels the playing field

Experimental Economics: credibility and the ban on deception

How to be sure that everyone is playing the same game?

- Experimental Economists never lie to subjects
- (this is not the case in psychology experiments)
- Deception (lying or in other ways inducing false beliefs in the subjects) is banned
- This is made to make sure that the experiment is **credible**
- Subject can trust the experimenter do to exactly as he said
- Subjects' beliefs are aligned
- Internal validity is enhanced

Lying is usually the fastest route; no deception means hard work, but more trustworthy result

Questions?

Experiments in food economics: eliciting value

Value elicitation

How do we elicit value from subjects?

We sometimes need to know *how much people value stuff*. "Stuff" here can be products, policy options, their freedom, control over their data...

- value is the bedrock of demand;
- different value across people, space and time is the reason for trade;
- different values across options drives firms decisions
- value for policy options drives policy decisions

Value and price

A first experiment: **value** = **price**?

A simple Cola experiment

How much are you willing to pay for a coke can?

A simple Cola experiment

How much are you willing to pay for a coke can?

- Actual selling price will be drawn randomly $\sim U(0, 1)$.
- You are free to submit any offer – this is called **bid**
- If your price $\geq$ the hidden price, you **buy** at the hidden price
- If your price $\leq$ the hidden price, you do not buy.

- Write **your name** and **your willingness to pay** in the online form
- I'll reveal the price once *everyone has submitted their* **bid**

QR code

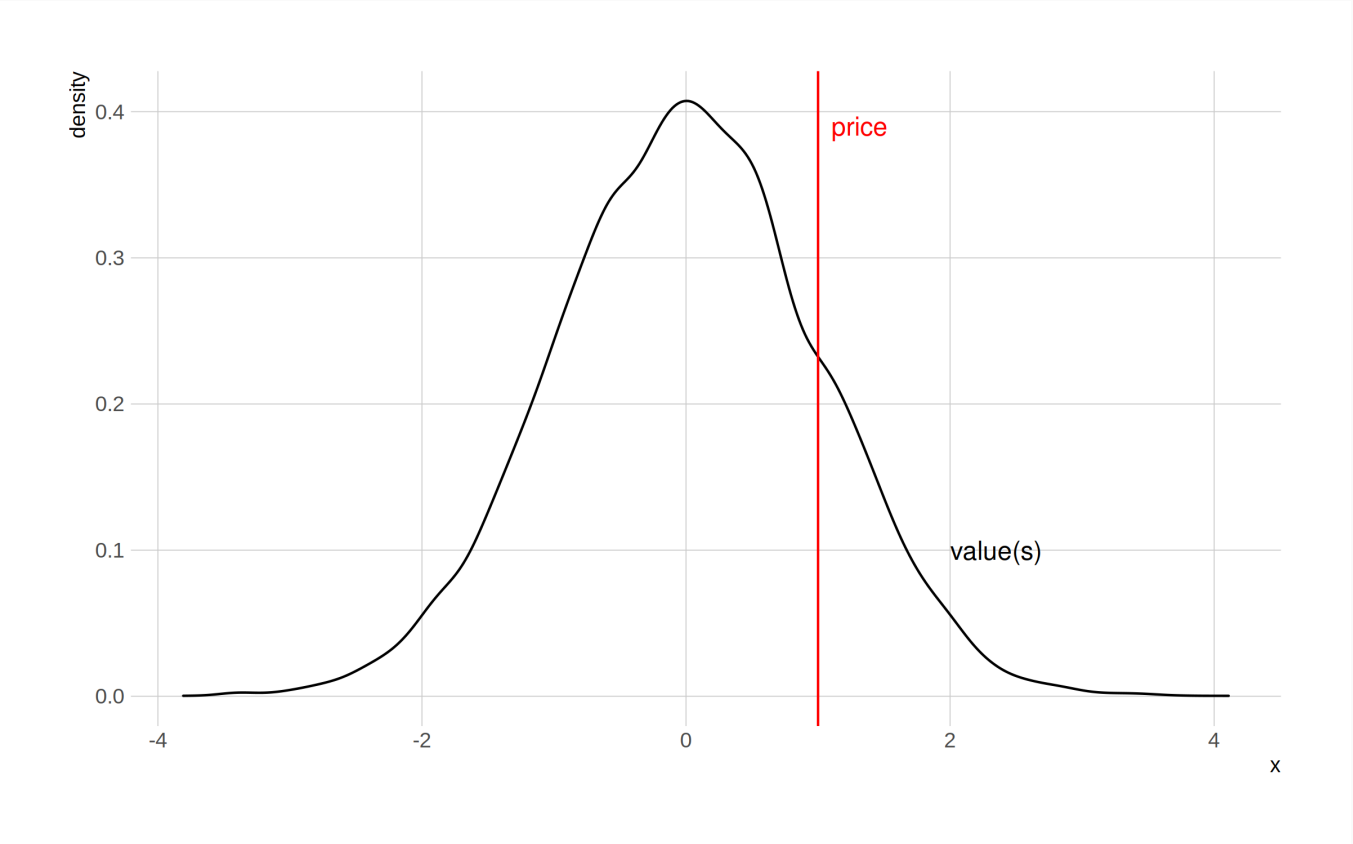
Experimental Economics: incentive-compatibility

If self-declared $\neq$ incentivized: hypothetical bias

- Usually self-declared price $>$ incentivized price
- A decision that results in *actual* consequences is called **incentive-compatible**
- i.e., real consequences + best strategy is to reveal your true preferences

We go at great length to create incentive-compatible mechanisms

Value and price



Value $\neq$ price

- Value belongs to consumers, price is set by firms given market conditions
- value can be higher, lower or equal to price
- it is the amount of satisfaction (=utility) you derive from a good
- it is (also) the personal amount of resources you wish to allocate to a good
- value is influenced by prices
- transactions occur if value $>$ price: then consumers have a surplus (value - price)

Value as willingness to pay

Value as Willingness to Pay (WTP)

We can operationalize value as willingness to pay (WTP)

- the amount of money a consumer wishes to allocate to a good **she does not yet have**
- Might depend on external conditions...
- ...good availability...
- ...information...
- ...existence of markets... etc.

Value elicitation: incentive compatibility

People could lie about their value, for a host of reason

- to get a personal advantage: *strategic* lying
- to foster their personal agenda: *political* lying
- to please the experimenter: *demand effect*
- unconsciously: *hypothetical bias* [remember?]
- just plain *error* or *inattention*

we need incentive compatible mechanisms to elicit value

BDM mechanism

Random price mechanism: Becker-DeGroot-Marschak (BDM)

- subjects privately and simultaneously submit a sealed bid
- the selling price is drawn from a (known) uniform distribution on a (known) support
- if bid $\geq$ price, then object is bought at price
- if bid $<$ price, then no transaction
- same mechanism used on earlier to for the coke can

Q: why the random price?

BDM mechanism – optimal strategy

BDM is incentive compatible: optimal strategy is to bid own real value

- your value is v_i, your bid is b_i, and the random drawn price is p
- if $b_i > v_i$, then:
  - if $b_i > p > v_i$, you **buy** at p that is higher than your value
- if $b_i < v_i$, then:
  - if $v_i > p > b_i$, you **do not** buy at a price that would have given you a gain
- hence by bidding $b_i \neq v_i$ you have either a **loss** or a foreg**one gain**
- the best you can do is bidding $b_i = v_i$

Auctions

BDM is a very basic kind of auction: other auctions are used

- sealed-bid *vs.* oral auctions
- first *vs.* second (or N^{th}) price auctions
- ascending or descending auctions

Sealed-bid 3^{rd} price auction: a milka chocolate bar

- you each submit a sealed bid here.
- the **two** highest bids buy the object...
- ...at the **third** highest price.

QR code

N^{th} price auction – optimal strategy

Optimal strategy is to bid your value

- ...if you bid higher, you risk buying at too high a price
- ...if you bid lower, you risk not buying at a good price
- (same reasoning as for the BDM)

- Q: would a **first** price sealed-bid auction be different?

- Q: why would you prefer auction over BDM?

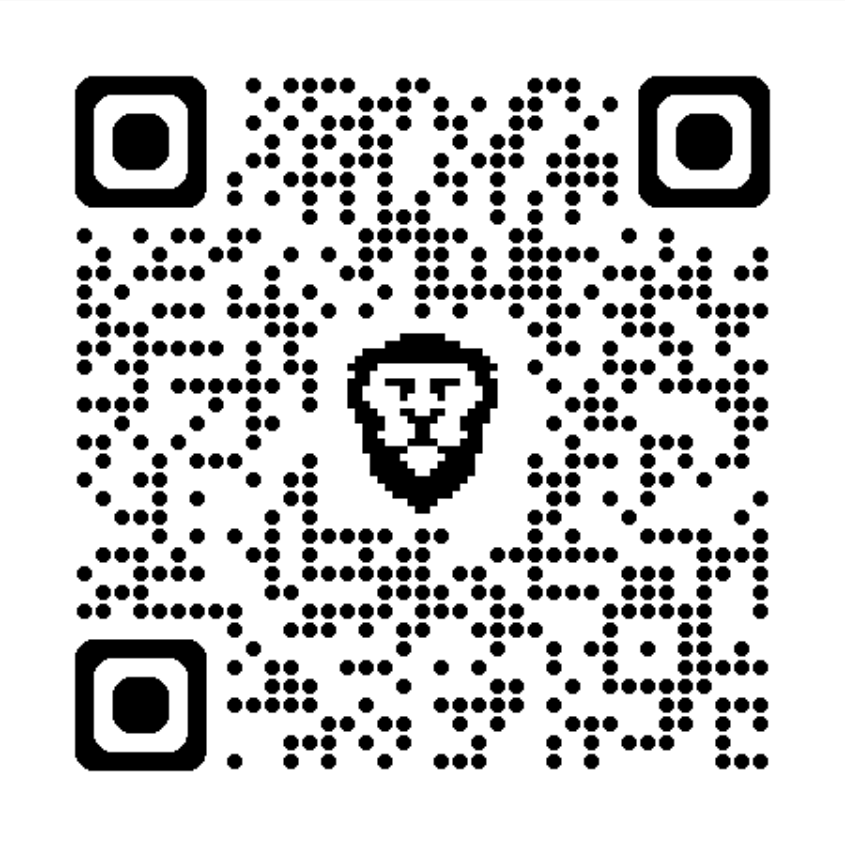
Common value auctions

First price auction for an oil extraction permit

- an oil field has a capacity estimated $\sim U(0, 50)$
- we set up a **first price** auction: the higher bid will get the extraction rights

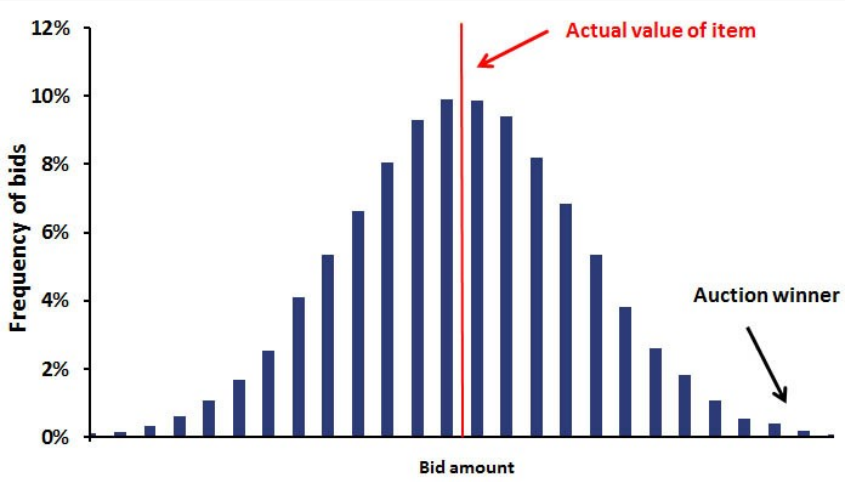
- we do this for real: 1 barrel $\Rightarrow$ 1 cent
- the oil well is worth between 0 and 50 cents
- **Now place your bids!**

QR code



Winner's curse

- this is a first price auction with a *common* but *unknown* value
- each participant has a heterogeneous estimate of the oil yield
- the one that will win the auction is the one most likely to overestimate it
- and hence the one less likely to make profits!



Winning an auction can ruin you!

Value as Willingness to Accept (WTA)

We can also operationalize value as willingness to accept (WTA)

- the amount of money a consumer wishes to receive to part with a good **that she has**
- Might depend on external conditions...
- ...good availability...
- ...information...
- ...existence of markets... etc.

WTP, WTA

Goods for which you might elicit WTP

- private goods (soda, cookies, electricity supplies...)
- public goods (a new park, a new social service...)

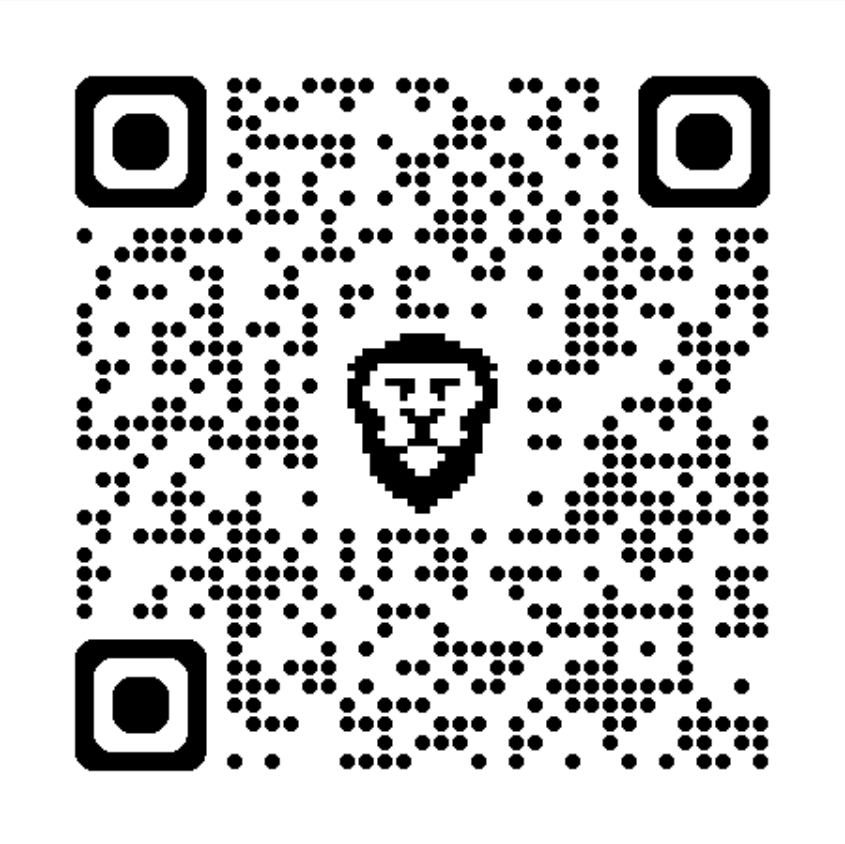
Goods for which you might elicit WTA

- private goods (selling your car, changing electricity provider...)
- public goods (expected environmental deterioration, a new development on an existing park...)

A simple experiment

- half of you receive salty crisps
- half of you receive sweet madeleines
- please indicate in a sealed bid:
  - for how much would you sell the good you own
  - for how much would you buy the good you do not own
- we will then randomly match you, and if prices match (bid>ask) the transaction is made
- according to the usual rules

QRcode



WTP vs. WTA: experiments

Students in every other seat were given university mugs. Then reported how much they would be willing to sell the mug for.

Students who did not get a mug reported the price they would be willing to pay to get one.

What happened?

a) The students with mugs priced them higher.
b) The students with no mugs priced them higher.
c) Both sets of students priced them about the same

WTP vs. WTA: experiments

Students with the mugs were willing to sell them, on average, for

$4.50

Students with no mugs were willing to buy them, on average, for

$2.25

Kahneman, D. (UC Berkley), Knetsch, J. (Simon Fraser U), Thaler, R. (Cornell), 1990, Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy, 98*(6), 1325-1348.

WTP vs. WTA: the endowment effect

this is the Endowment effect: you value a good more because it's yours

"Have you ever noticed that their stuff is shit and your shit is stuff?"

– George Carlin, A Place for My Stuff

WTP vs. WTA: theory

In theory (on average, in large samples) $WTP = WTA$

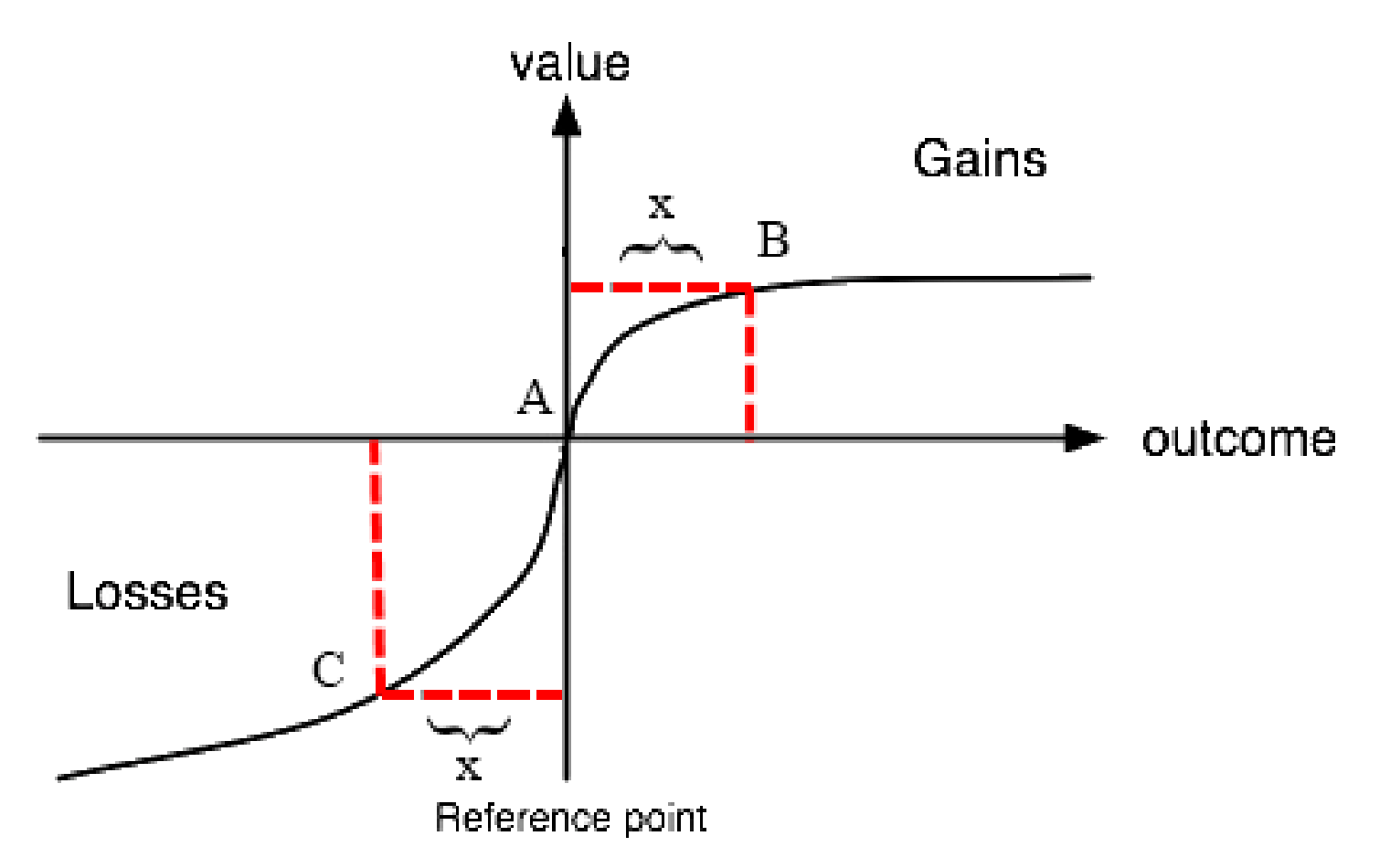
- tastes differ, so some people might prefer X to Y and others Y to X
- but not *systematically so* [here you are $\sim$16, way too small]
- if the two goods are equally desirable, we should see $\sim$ the same evaluations, and roughly an even number of transactions
- if desirability is asymmetric, we should see asymmetry in transaction, but still no systematic difference between the WTP of one group and the WTA of the other.

Why the endowment effect?

Loss aversion and reference dependence

- people feel losses more than equivalent gain
- (would accept to play a lottery in which you have 50-50 chances of losing 100 and gaining 120? and 100 and 100?)
- this is called *loss aversion*
- moreover, losses are not evaluated around zero, but around the *status quo*
- if you get a raise in your wage, you get used to it; and less money will be felt as a loss

Loss aversion and reference dependence



Mainly for this contribution Daniel Kahneman and Amos Tversky got the 2002 Nobel Prize in Economics!

More complex evaluation: a burger with fries

- we deal with a **tasty burger** with **fries side**.
- please write down somewhere your WTP for this meal now.
- no incentives this time, but let's pretend as if they were there.

Social value and social norms

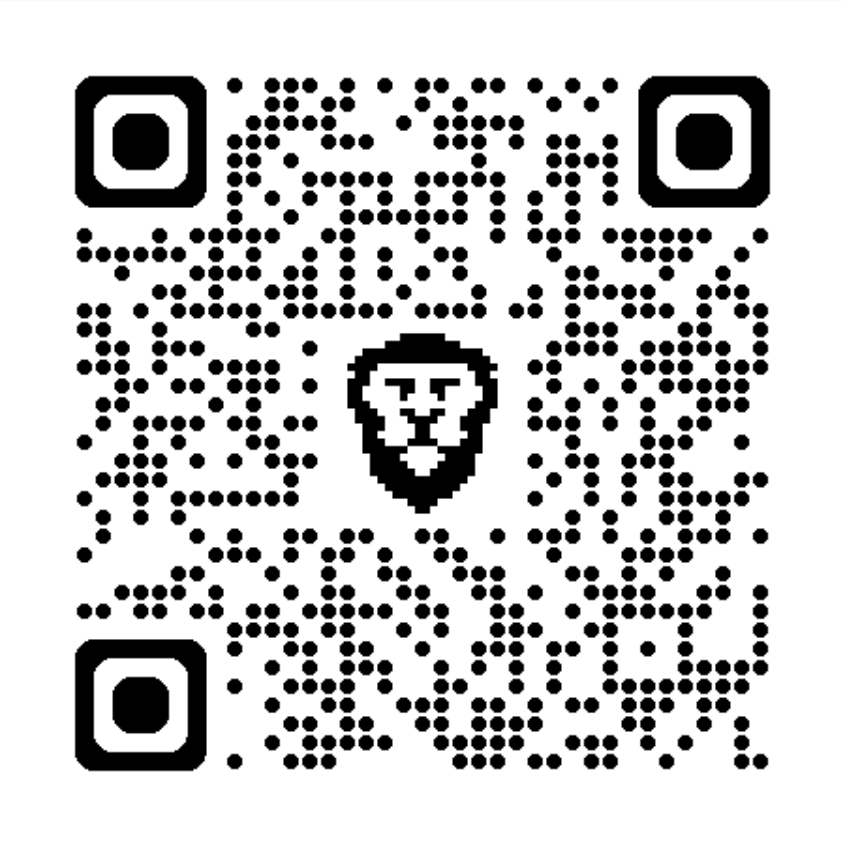
Now imagine you want to elicit social norms on value

- that is, **not** the individual value
- but the value that an individual *thinks* the *others* have
- is it possible to incentivize this?

Social norm elicitation

- we deal with a **tasty burger** with **fries side**.
- your task is to *guess* how much *the average bid* of **everyone in the room** is
- we will compute the **average** of all your guesses, and that is the target.
- the person that **gets nearest to the average** wins!

QRcode



A Beauty contest

SUNDAY MORNING, MAY 14, 1933 SAN ANTONIO EXPRESS D 9

Here Are Ten More Winners in the Court of Honor Competition

JULIA ROSE WIEDENMANN.

JANIE GRAY PAYNANS.

WANDA STAER.

BELLE SAPER.

THORA ELIZABETH MERCHANT.

CHRISTINE HUGHES.

VIVIAN FERGUSON.

HAZEL MAE GRENINGER.

LUCILLE RADER.

VIOLA DELL.

Queen and Ladies-in-Waiting Will Be Announced at Dinner Night of May 26 When Entire Court Will Be Guests of Honor.

San Antonio Express brewcrift presents the second group of 10 winners in the Court of Honor Competition, the first 10 having been presented last Sunday. Another group of 10 will be shown in the Sunday Express May 21, and the final 21 winners including the queen and Indianola waiting will appear in San Antonio Express May 28.

President Roosevelt's announcement that he would open the Century of Progress Exposition on May 27 instead of June 1, caused a quick shuffle of plans in the tentative schedule for entertainment of the court. Arrivals of the girls who will go to Chicago from other American cities and from Europe was pushed forward one day—from May 26 to May 25.

Then instead of a leisurely preparation of their costumes over a period of four days, they will hurry through Chicago and be ready on Friday, May 26, for the review by the Judges, that will place any of their number on the throne. The election of the queen will be held on May 26 instead of May 27, as originally planned.

That night, on the eve of the first question, the coronation dinner will be held with representative company. Attending each girl will be an honor student from Northwestern or the University of Chicago, and [illegible] themselves an important part of the tremendous pageant of the opening night.

At 1 o'clock in the afternoon, they will find themselves dressed for the significant ceremony, on the temporary rostrum to be built in Soldiers' Field for the brunette race who will honor the city and the occasion by their presence. In the same white organdy dresses in which they paraded before the Judges the preceding evening, the girls will participate in the ceremony, as the formal program are given out that will throw wide the grounds for a full five months [illegible]

THIS WEEK'S WINNERS

Janie Gray Paynans, 24 Virginia St., Dorchester, Mass.
Thora Elizabeth Merchant, 370 Funston Ave., San Francisco, Calif.
Wanda Staer, 3205 N. E. 42nd Ave., Portland, Ore.
Christine Hughes, South County, North Holston, Va.
Vivian Ferguson, 1305 Cumberland, Little Rock, Ark.
Julia Rose Wiedenmann, 837 Archbald, Kansas City, Mo.
Belle Saper, 1352 S. Avers Ave., Chicago, Ill.
Hazel Mae Greninger, 349 N. Felton St., Philadelphia, Pa.
Lucille Rader, Ennis, Tex.
Viola Dell, 3346 Carroll Ave., Chicago, Ill.

Beauty contest: why and optimal strategy

- beauty contest allows subjects to express beliefs over other subjects
- i.e. a proxy of the *social norm* (what I think others usually do $\sim$ what I think ought to be done)
- the optimal strategy is to state one's true belief (see auctions or BDM)
- BC allow to assess public awareness and awareness of public awareness about a topic

Questions?