$$\sum_{i=1}^{n} y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} x_{i1} + \beta_2 \sum_{i=1}^{n} x_{i2} + \cdots \beta_p \sum_{i=1}^{n} x_{ip}$$

$$\sum_{i=1}^{n} x_{i1} y_i = \beta_0 \sum_{i=1}^{n} x_{i1} + \beta_1 \sum_{i=1}^{n} x_{i1}^2 + \beta_2 \sum_{i=1}^{n} x_{i1} x_{i2} + \cdots + \beta_p \sum_{i=1}^{n} x_{i1} x_{ip}$$

$$\sum_{i=1}^{n} x_{i2} y_i = \beta_0 \sum_{i=1}^{n} x_{i2} + \beta_1 \sum_{i=1}^{n} x_{i1} x_{i2} + \beta_2 \sum_{i=1}^{n} x_{i2}^{i} + \cdots + \beta_p \sum_{i=1}^{n} x_{i2} x_{ip}$$

$$\vdots$$

$$\sum_{i=1}^{n} x_{ip} y_i = \beta_0 \sum_{i=1}^{n} x_{ip} + \beta_1 \sum_{i=1}^{n} x_{i1} x_{ip} + \beta_2 \sum_{i=1}^{n} x_{i2} x_{ip} + \beta_2 \sum_{i=1}^{n} x_{i2} x_{ip} + \cdots + \beta_p \sum_{i=1}^{n} x_{ip}^2 .$$

"They Told Me There Was No Math": Quantitative Methods in Environmental Policy

Jason E. James, Attorney-Advisor, Illinois Pollution Control Board

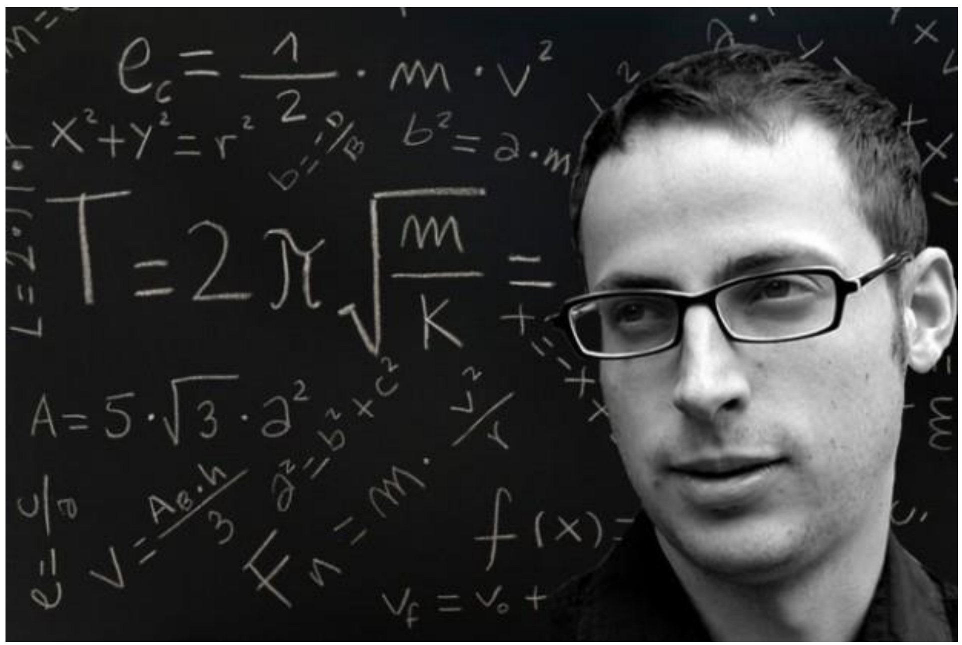
Presentation, August 25, 2016

Who needs statistics?

- "There are three types of lies—lies, damned lies, and statistics." – Benjamin Disraeli

Who needs statistics?

- "Who needs theory when you have so much information? But this is categorically the wrong attitude to take toward forecasting, especially in a field like economics where the data is so noisy." – Nate Silver



Why statistics?

- Data in public policy is here to stay
- Economics underlie the policies that the Board carries out through its adjudicatory cases and rulemakings
- Economists test theory with empirical (data-based) analysis using statistical methods
- Data-based policy is a common refrain, but in policy, data is meaningless (or worse) without statistical analysis
- When parties before the Board make claims that relate to economics, the Board as a panel of experts can critically evaluate these claims.

My background

- *Not* an economist or statistician
- Lawyer with a mathematics background and graduate coursework in economics and statistics
- Today's aim is to introduce some quantitative methods that policy professionals use to assess environmental policies
- Discussion will be done at a basic level—no knowledge of math assumed
- Naturally, all topics discussed get more complex as you get into the details
- When economic analysis pops up in Board matters—the Board should be able to identify and give a basic assessment of validity.

Outline for this hour

I. Economic theory predicts need for environmental regulation
II. Statistics interprets empirical data
III. Econometrics tests whether data supports predictions from economic theory

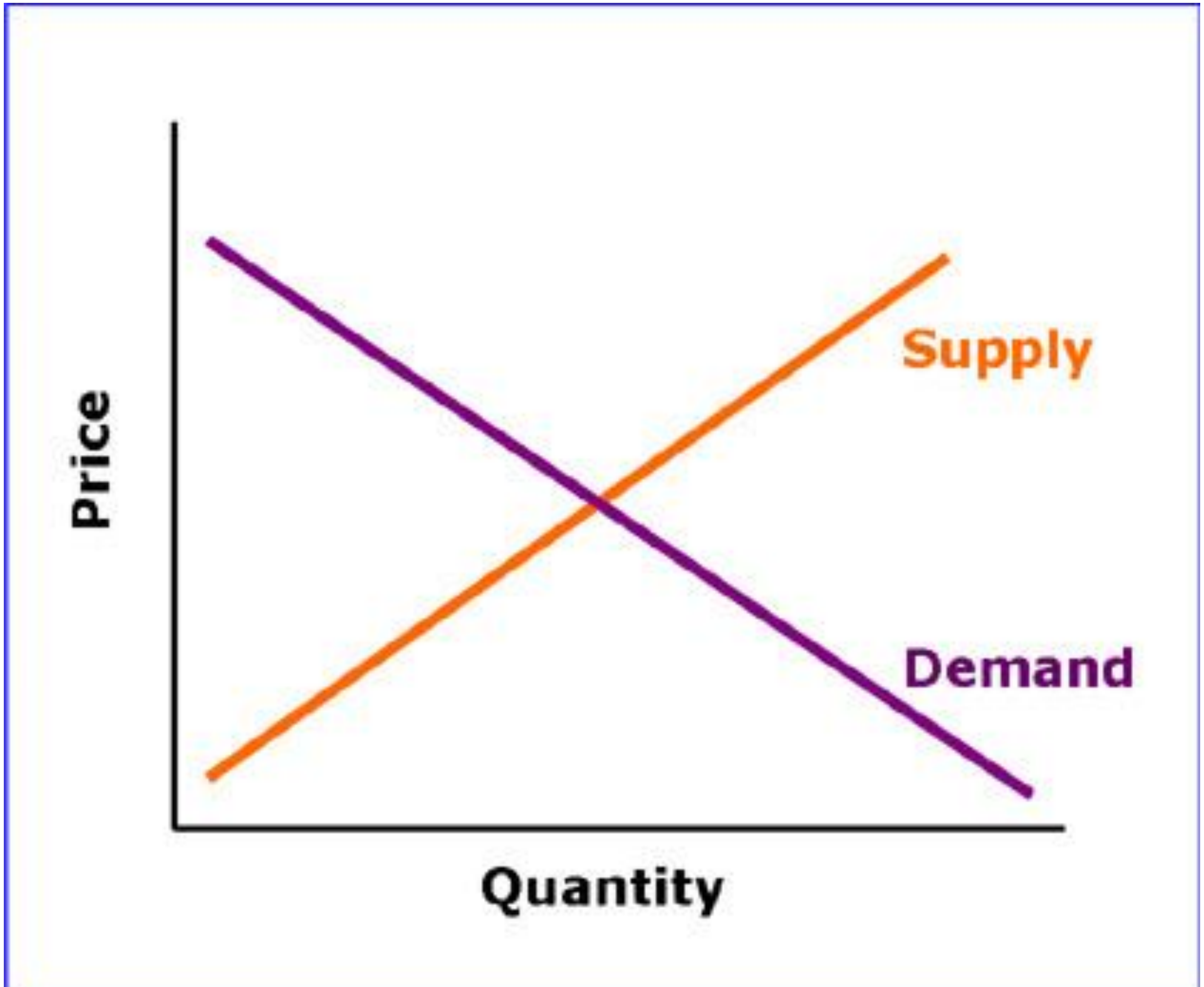
- Notes:
  - All content and opinions are solely my own—not the Board's and not Chairman Keenan's
  - All discussion concerning specific policies is simply demonstrative, I am not advocating for the merits of any specific study
  - As always, please interrupt me with questions and comments at any point

I. Economic theory of environmental regulation

Why regulate?

- "Efficiency":
  - Resources are allocated so that nobody can be made better off without making someone else worse off.
  - Economists usually seek the most efficient outcome.
  - Many economists advocate for redistributing resources based on equity *after* setting a policy that leads to the most efficient outcome
- "Efficient markets":
  - In a competitive economy, the market equilibrium for distribution of resources is the most efficient outcome.

The law of supply and demand



Better if the Board never existed?

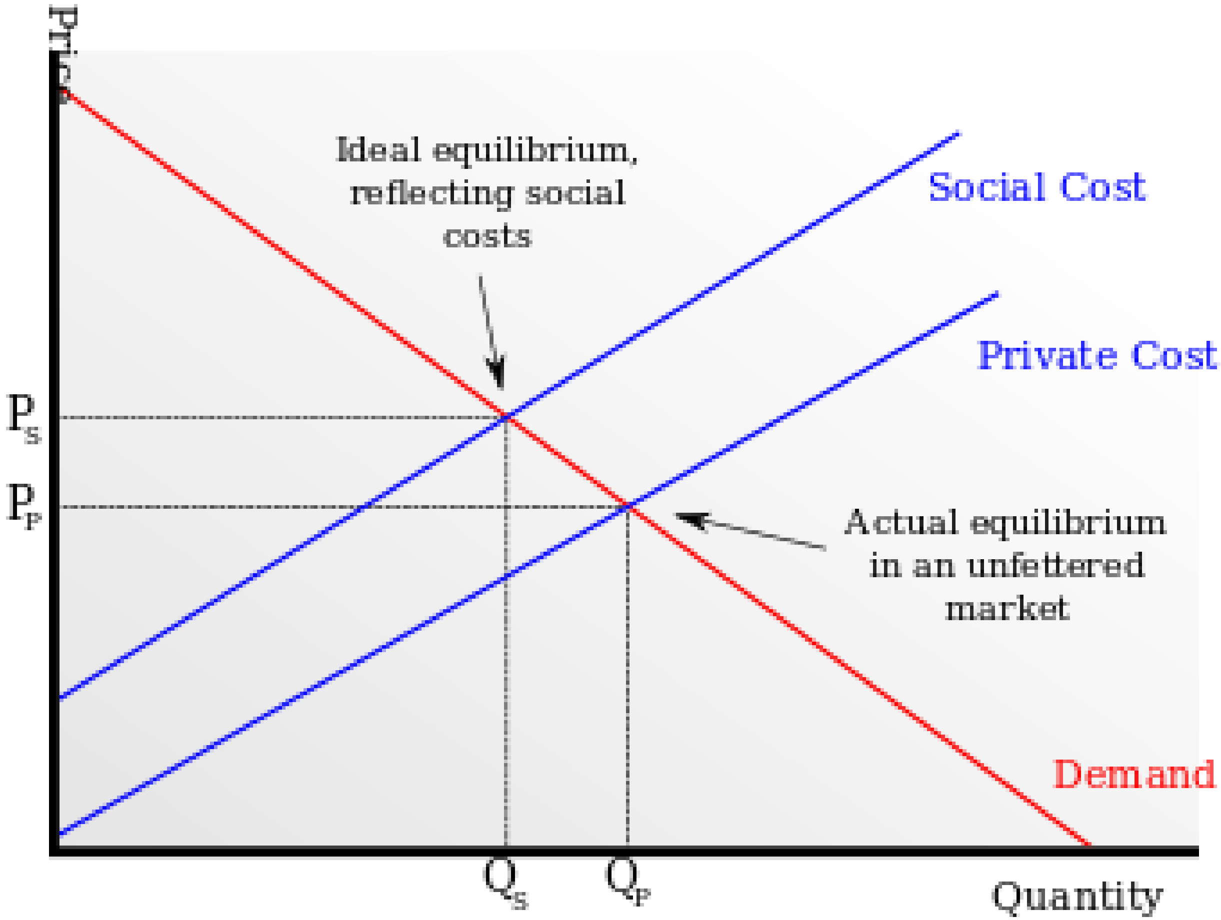
Why regulate?

- If competitive markets are efficient, then why should the government regulate?
- In a "competitive" economy:
  - There are well-defined, transferable, and secure property rights for all goods with all benefits or costs accruing to the property owner
  - Individual producers and consumers cannot influence market prices
  - Consumers and producers have complete information on current and future prices
  - There are no transaction costs to trade goods

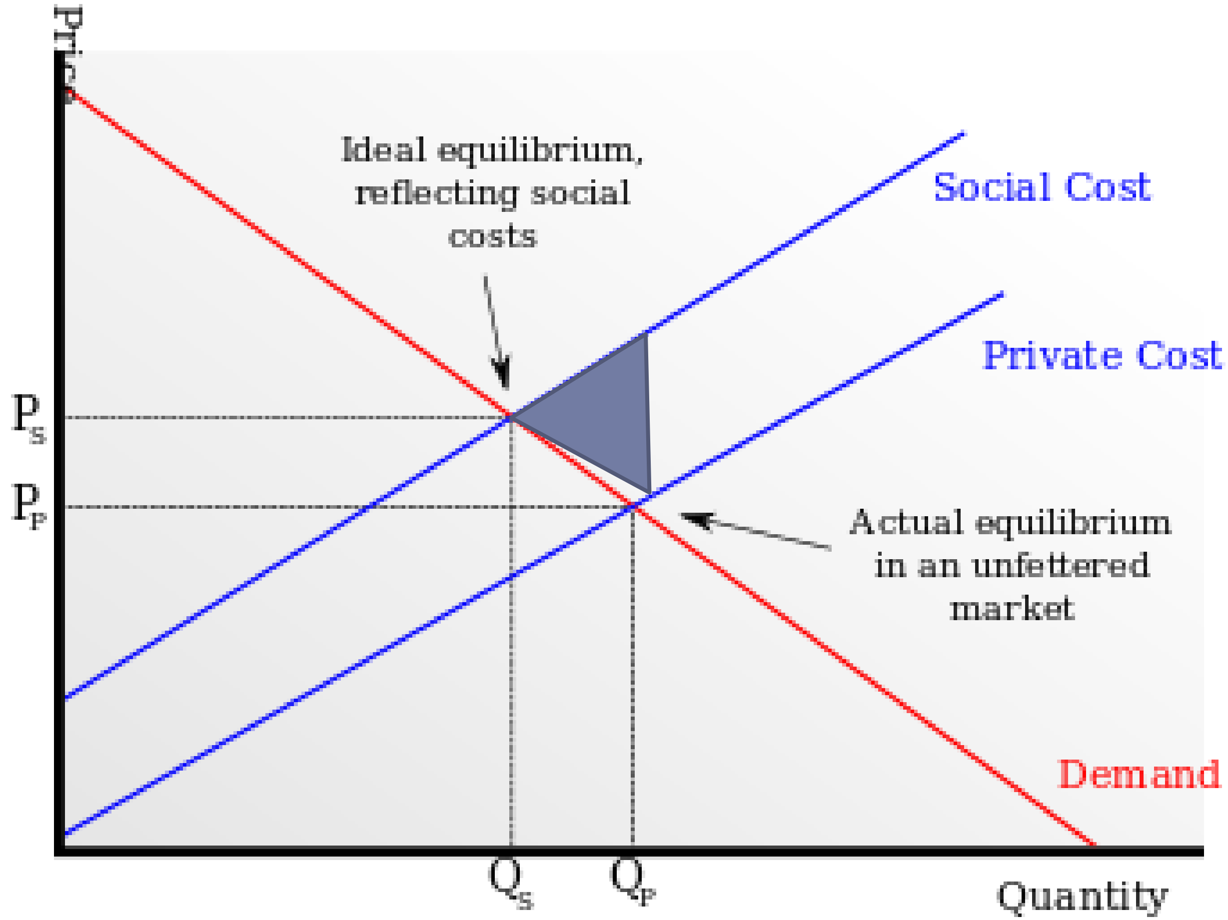
Why regulate?

- Pollution violates the first assumption:
  - A polluting facility, absent regulation, can affect the health and property of the nearby community.
  - These costs are incurred by the local community and not the facility's owner.
- This externality is a "market failure"
  - An unregulated economy with a market failure does not lead to an efficient outcome
  - Government regulation is needed to achieve the efficient outcome

Externality in supply and demand



Externality in supply and demand



Regulating despite uncertainty

- The economic theory of externalities is intuitive, but deciding "how much" protection is warranted is not easy
- The true and complete cost of most environmental externalities is uncertain
- Uncertainty in environmental policy makes it difficult to decide "how much" regulation
  - What policies are "good deals" (cost-effective) and which are bad (not cost-effective)?
- Statistics is the best tool we have to evaluate this uncertainty and make the best possible "bet"

Example: Climate Casino

Example: “Climate casino”

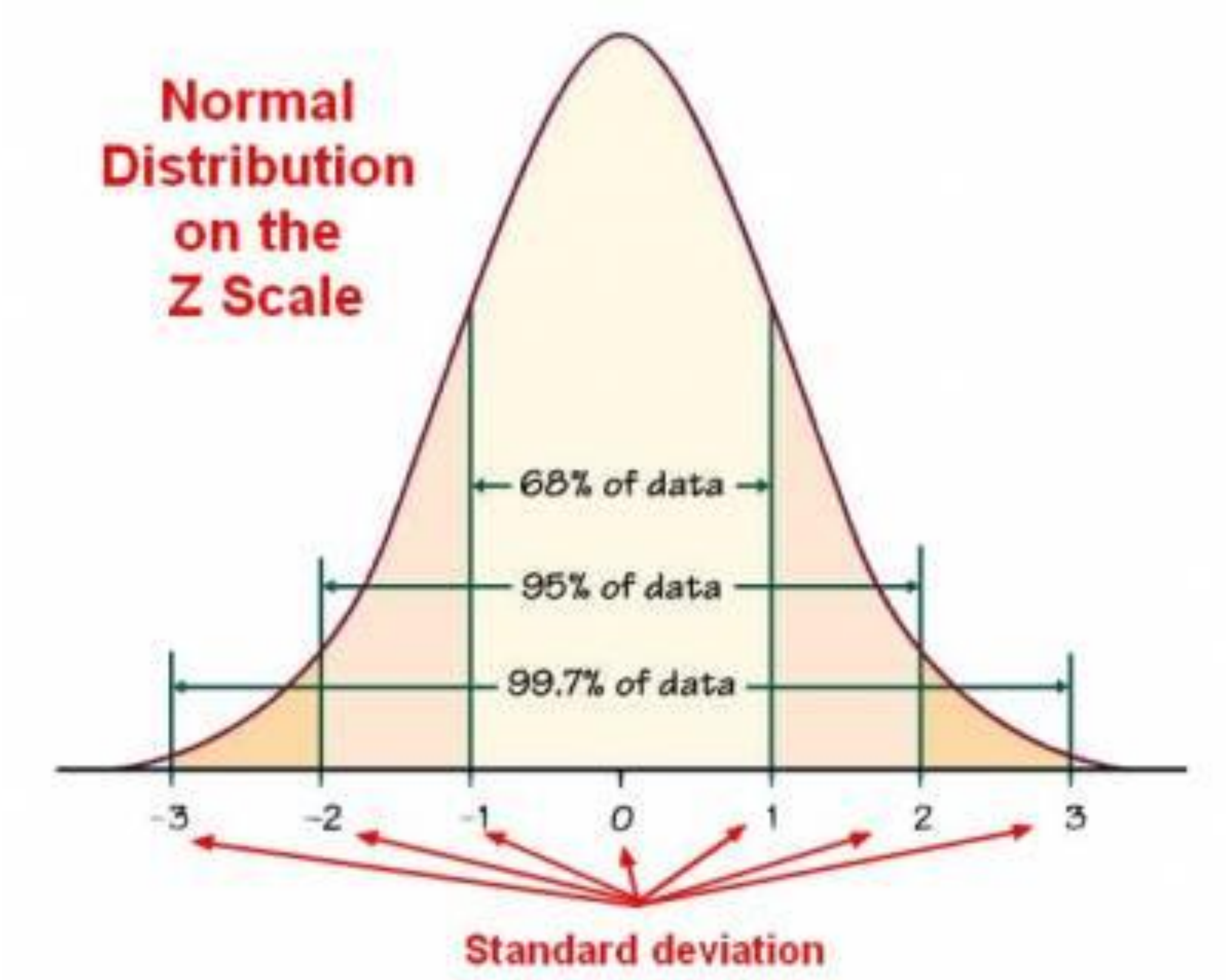
- Climate change:
  - We generally know quantity of GHG emissions
  - We know *direction* but do not know exact *magnitude* of:
    - Emissions raising atmospheric concentrations of GHG;
    - Atmospheric concentrations raising mean global temperatures;
    - Mean global temperatures and resulting local climactic effects;
    - Local climactic effects and ultimate total social cost of externality.
- Climate change regulation is necessarily a type of gamble
- Where should we place our bets?

- From “Climate Casino” by W. Nordhaus (2013)

Safest bet is to "buy" cheapest emissions reductions

- Making the best "investments" in environmental quality makes environmental policies politically feasible
- "Buying" environmental quality irrespective of price deteriorates political will to create a better environment.
- Political problem particularly acute in climate change:
  - Costs are concentrated in the here and now;
  - Benefits are dispersed worldwide and among generations.

II. Statistics to analyze real-life data



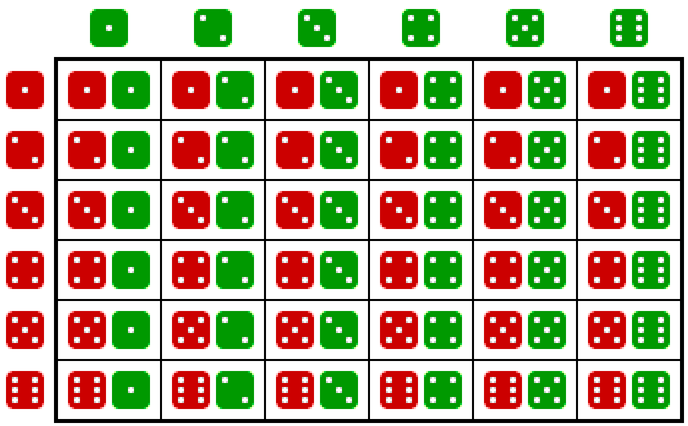
Now the numbers

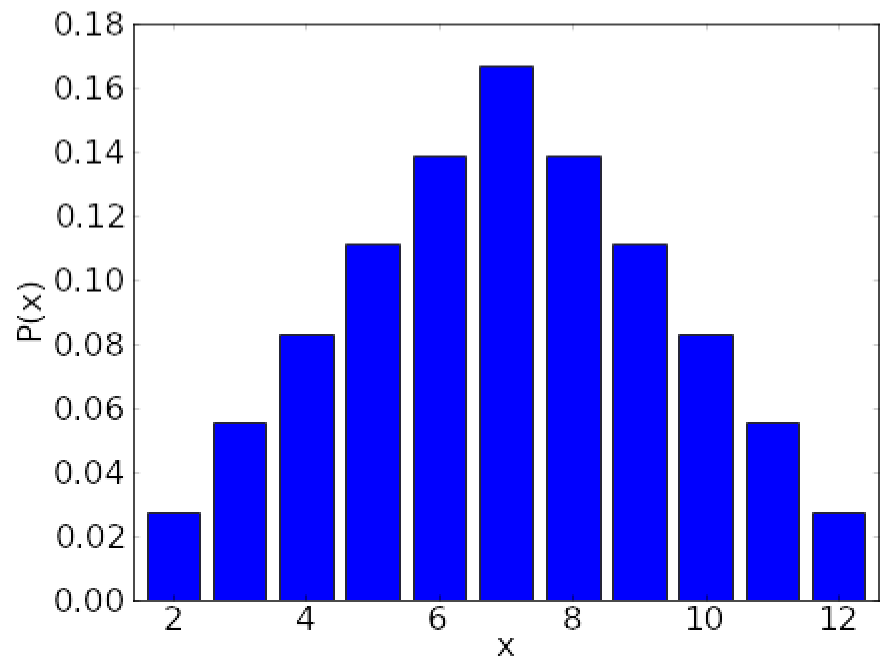
- We turn from economic theory into analysis of empirical data
- Economists use probability, statistics, and econometrics to evaluate policies that aim to address pollution

Probability of outcomes when rolling two 6-sided dice (2d6)

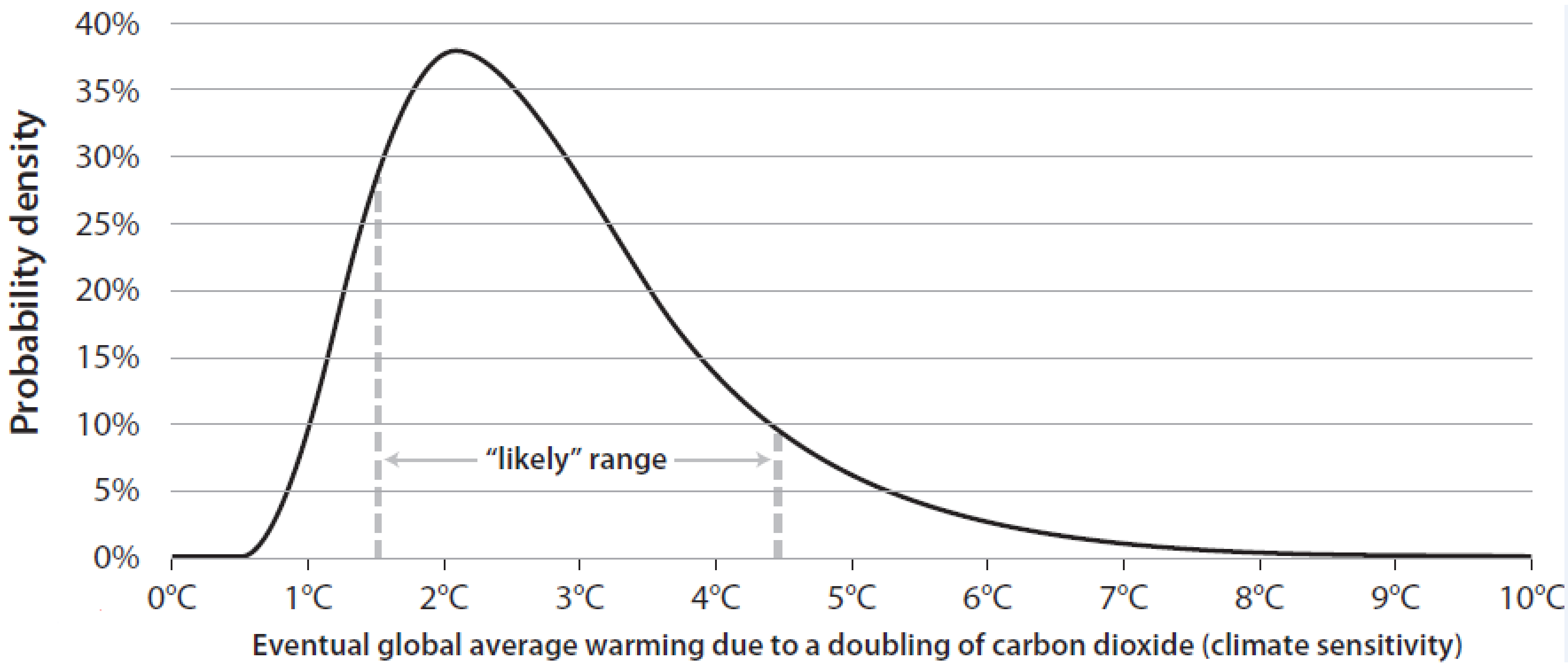
Outcome	Ways to Get Outcome	Number of Ways	Odds
2	(1,1)	1	1/36
3	(1,2), (2,1)	2	2/36
4	(1,3), (2,2), (3,1)	3	3/36
5	(1,4), (2,3), (3,2), (4,1)	4	4/36
6	(1,5), (2,4), (3,3), (4,2), (5,1)	5	5/36
7	(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)	6	6/36
8	(2,6), (3,5), (4, 4), (5,3), (6,2)	5	5/36
9	(3,6), (4,5), (5,4), (6,3)	4	4/36
10	(4,6), (5,5), (6,4)	3	3/36
11	(5,6), (6,5)	2	2/36
12	(6,6)	1	1/36

Probability of outcomes when rolling two 6-sided dice (2d6)



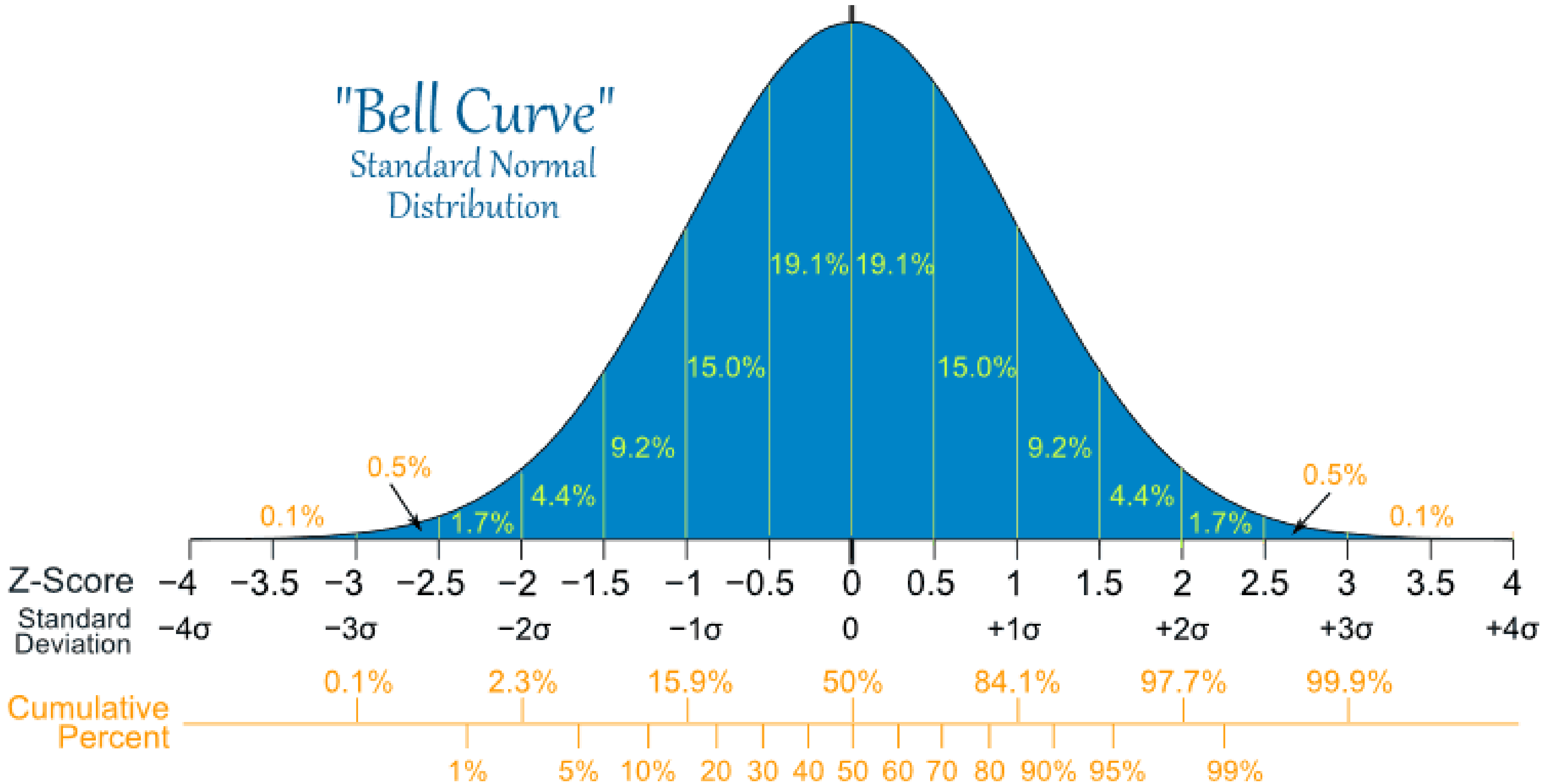


Probability in environmental policy



- From "Climate Shock" by Gernot Wagner and Martin Weitzman, economists for EDF

Standard normal distribution – The "Bell Curve"



$$\Phi(x) = \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}$$

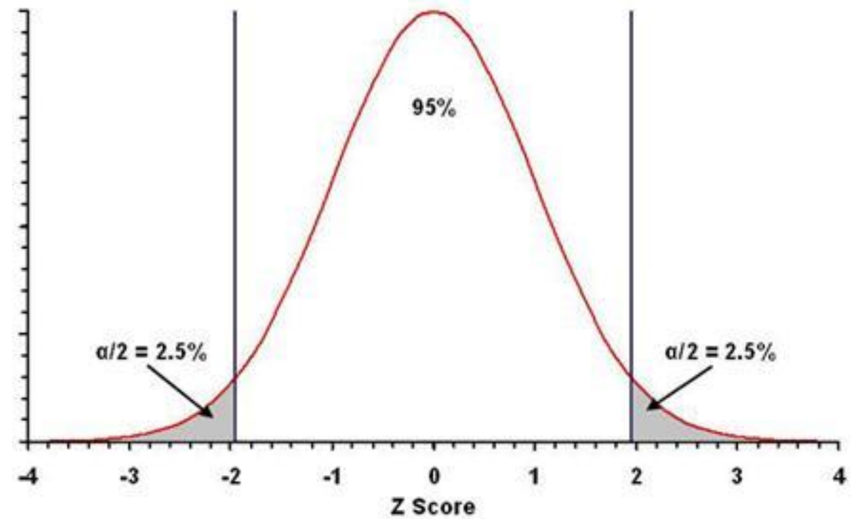
$$\int_{-\infty}^{\infty} \Phi(x)\, dx = 1$$

Statistical significance

- In law, we use "significant" in a qualitative sense.
- For economists, "significant" is a hard, quantitative measure.
- This concept is central to econometrics—the means by which economists infer causal relationships between a policy and its effect.
- Significance tests analyze data to see whether the result is "statistically significant"—whether the result means something or could have been created by random chance

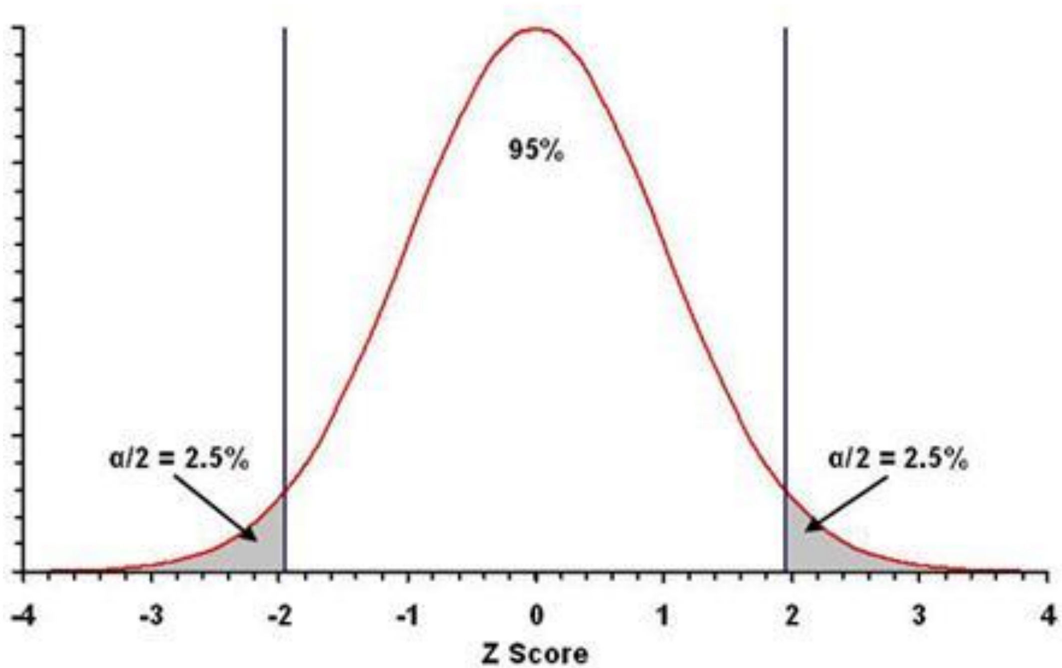
Example: Statistical significance

- Say we don't know whether or not a die is weighted to more frequently roll a '6' or is a 'fair' die.
- If mean values of rolls of the die significantly deviate from the expected mean value of rolls of a 'fair' die, we can infer that the die is weighted!
- In fact, if the mean value is about 2 standard deviations from the expected mean value, economists call this "significant"
- But we aren't certain—it's still possible, though very unlikely, for a fair die to roll a 6 a billion times in a row.



Example: Statistical significance

- Sample chosen no better than random chance—not a significant result



In the face of uncertainty, how to test a regulation's effects?

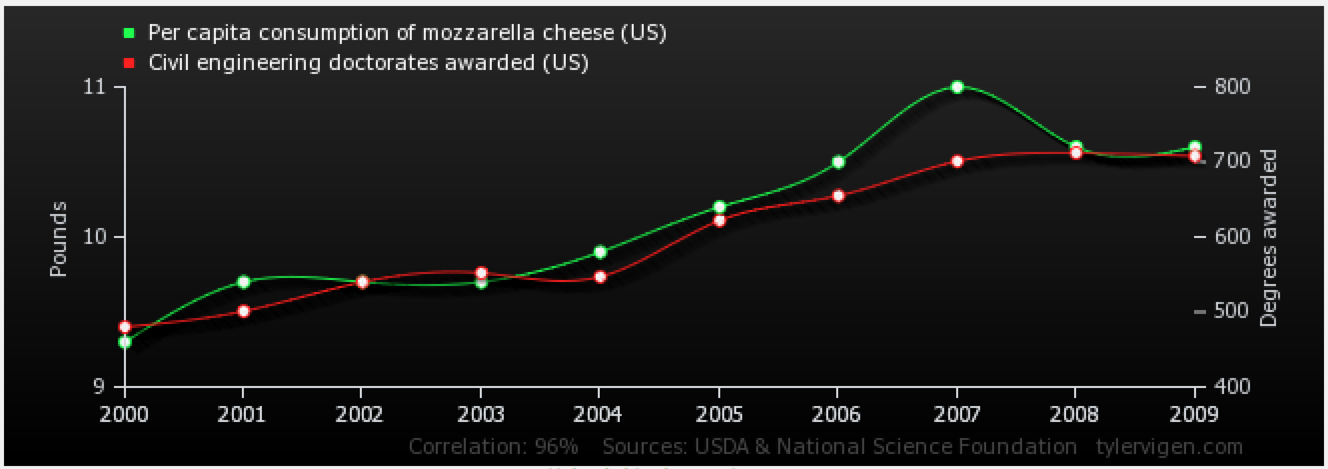
- These tools help us test whether we can say a policy is having an effect
- Another example:
  - The reading level of students in an elementary school closely correlates with students' shoe size
  - A very naïve policy-maker decides to implement a policy to increase students' shoe size with the end goal of helping students reading abilities.
  - But causation is not correlation!

Correlation is not causation!

Per capita consumption of mozzarella cheese (US)

correlates with

Civil engineering doctorates awarded (US)



Upload this chart to imgur

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Per capita consumption of mozzarella cheese (US) *Pounds (USDA)*	9.3	9.7	9.7	9.7	9.9	10.2	10.5	11	10.6	10.6
Civil engineering doctorates awarded (US) *Degrees awarded (National Science Foundation)*	480	501	540	552	547	622	655	701	712	708

Correlation: 0.958648

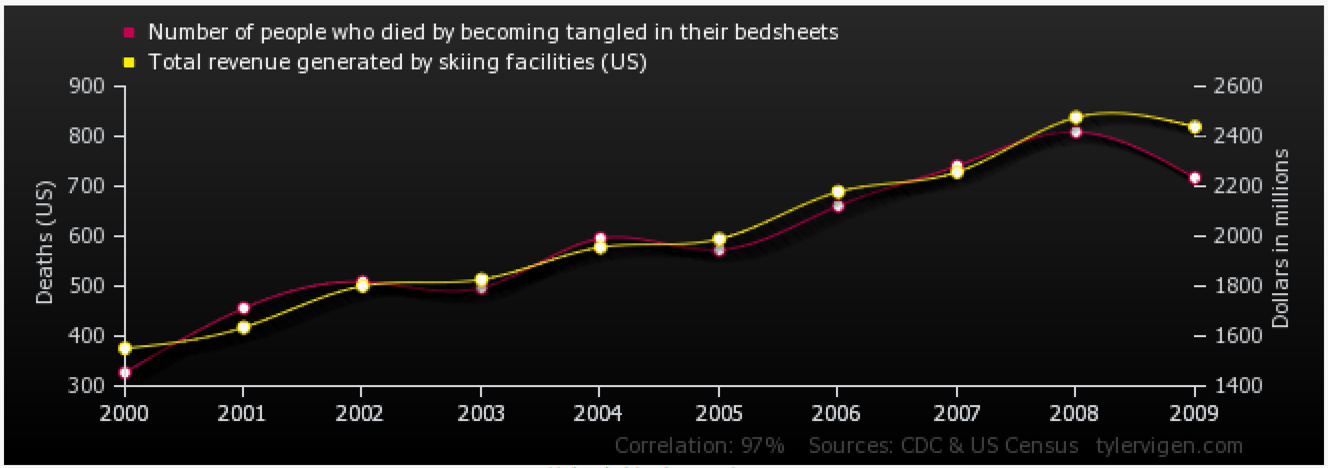
- From "Spurious Correlations" (tylervigen.com)

Correlation is not causation!

Number of people who died by becoming tangled in their bedsheets

correlates with

Total revenue generated by skiing facilities (US)



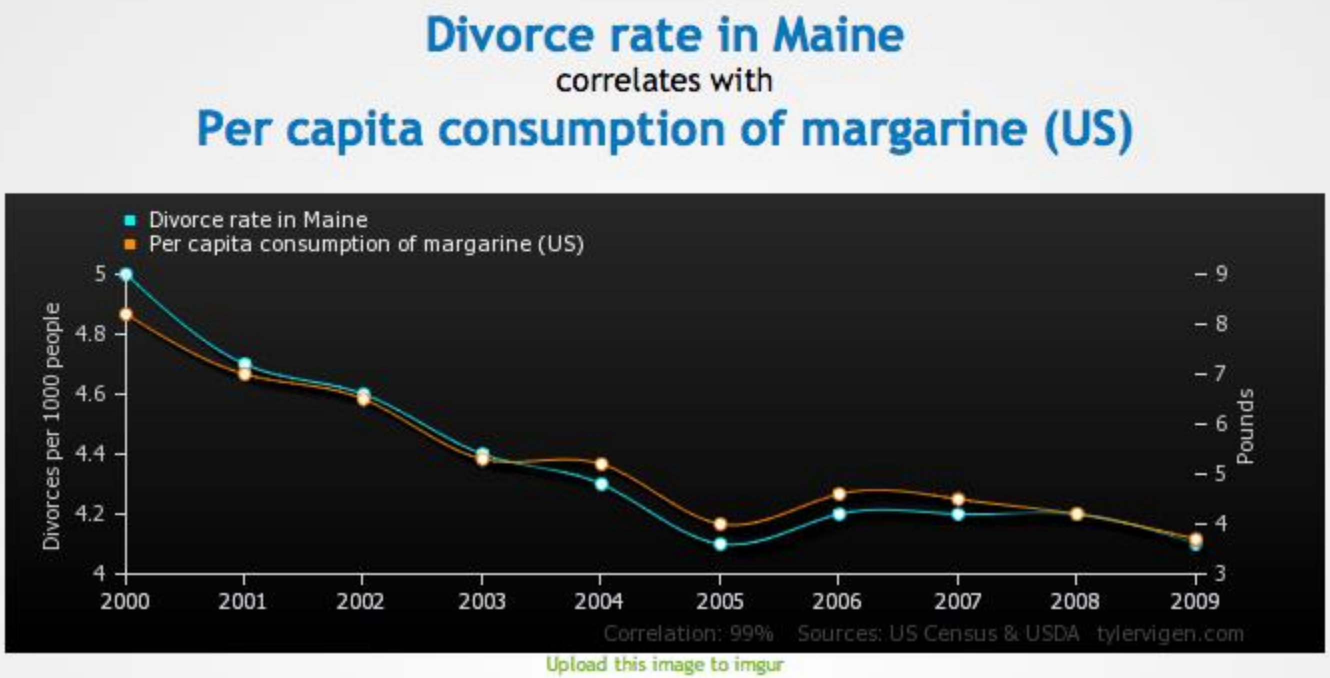
Upload this chart to imgur

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Number of people who died by becoming tangled in their bedsheets *Deaths (US) (CDC)*	327	456	509	497	596	573	661	741	809	717
Total revenue generated by skiing facilities (US) *Dollars in millions (US Census)*	1,551	1,635	1,801	1,827	1,956	1,989	2,178	2,257	2,476	2,438

Correlation: 0.969724

- From "Spurious Correlations" (tylervigen.com)

Correlation is not causation!

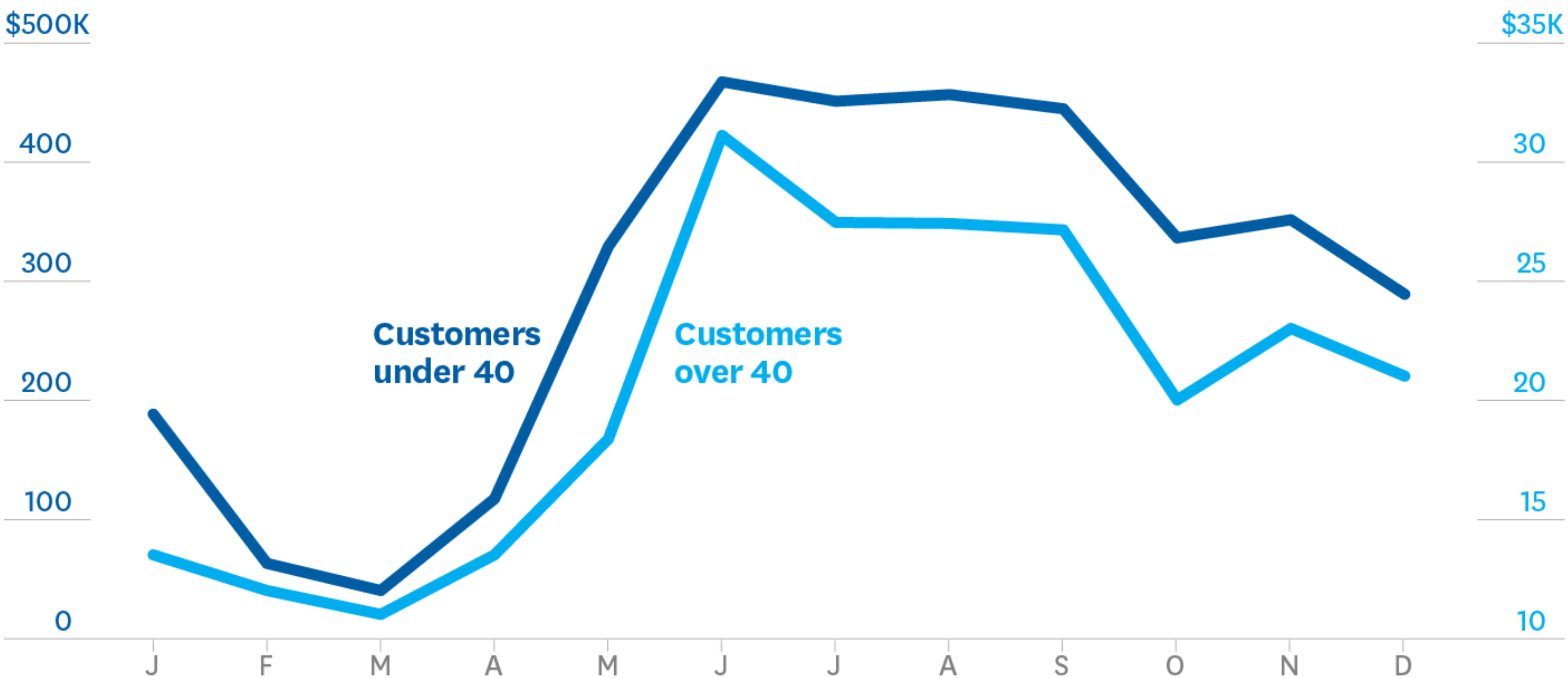


	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Divorce rate in Maine *Divorces per 1000 people (US Census)*	5	4.7	4.6	4.4	4.3	4.1	4.2	4.2	4.2	4.1
Per capita consumption of margarine (US) *Pounds (USDA)*	8.2	7	6.5	5.3	5.2	4	4.6	4.5	4.2	3.7

Correlation: 0.992558

From “Spurious Correlations” (tylervigen.com)

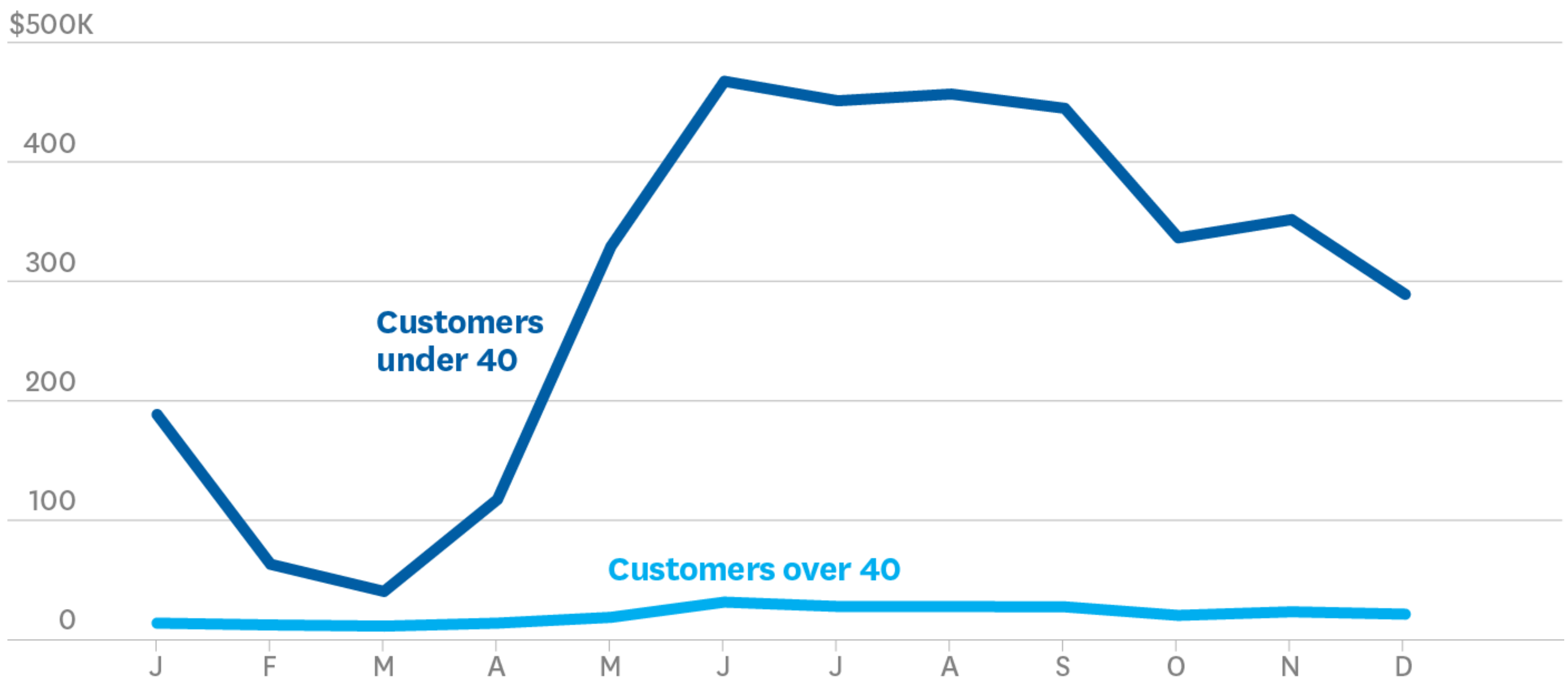
Both spurious and pernicious



FROM "BEWARE SPURIOUS CORRELATIONS," JUNE 2015

© HBR.ORG

Both spurious and pernicious



FROM "BEWARE SPURIOUS CORRELATIONS," JUNE 2015

© HBR.ORG

III. Econometrics to test economic theory using data

Econometrics – drawing causal inference

- Applying the statistical method of linear regression to a set of data
- Essentially attempts to simulating a randomized controlled test, as in the field of science
- Hopes to show how much one factor "causes" an effect—a stronger relationship than correlation
- For example: what effect does a worker's education, experience, and tenure have on his or her wage?

Example: effect of soda tax

ARCH | The New York Times | SUBSCRIBE NOW | LOG

ce That Soda ... da Drinking

ECONOMIC TRENDS
The Housing Market Is Finally Starting to Look Healthy

THE NEW HEALTH CARE
The EpiPen, a Case Study in Health System Dysfunction

THE 2016 RACE
What Are the Chances That Democrats Retake the House?

The Upshot

FOLLOW US:
GET THE UPSHOT IN YOUR INBOX

SCALING BACK

More Evidence That Soda Taxes Cut Soda Drinking

Margot Sanger-Katz @sangerkatz AUG. 25, 2016

3

It may seem obvious that taxing sugary drinks causes people to drink less of them. But that's actually controversial.

Now a new study out of Berkeley, Calif., adds to the evidence that our intuition is right.

Researchers followed residents of several low-income communities in Berkeley, San Francisco and Oakland around the time that Berkeley voters passed the country's first big soda tax in 2014. The study found that, in the four months after the tax took effect last year, self-reported consumption of sugary drinks fell by 21 percent in the Berkeley neighborhoods, but rose by 4 percent in the other two cities.

The study, published in The American Journal of Public Health on Tuesday, also found that the Berkeley residents reported drinking more water, a sign that they were replacing sugar-sweetened beverages with something healthier.

The research was conducted using in-person surveys of neighborhood residents, a method with some problems because people are not always accurate in describing their diets. But the study is the first to assess soda

Example: effect of soda tax

TABLE 2—Beverage Consumption and Pre- to Posttax Change (%) in Consumption in Berkeley, CA, Versus Comparison Cities (Oakland and San Francisco, CA) Among 2679 Participants

Consumption (Times/Day)	Berkeley, CA (n = 873) Unadjusted Pretax, Mean ±SD	Berkeley, CA (n = 873) Unadjusted Posttax,[a] Mean ±SD	Berkeley, CA (n = 873) Unadjusted Absolute Difference	Berkeley, CA (n = 873) Adjusted[b] Percent Change[c]	Comparison Cities (n = 1806) Unadjusted Pretax, Mean ±SD	Comparison Cities (n = 1806) Unadjusted Posttax,[a] Mean ±SD	Comparison Cities (n = 1806) Unadjusted Absolute Difference	Comparison Cities (n = 1806) Adjusted[b] Percent Change[c]	Ratio of Post- to Pretax Consumption in Berkeley Relative to Comparison Cities (n = 2679), B[b] (95% CI)
SSBs	1.25 ±2.25	0.97 ±1.66	−0.28	−21	1.29 ±1.76	1.26 ±2.09	−0.03	+4	0.76 (0.58, 0.995)
Regular soda	0.47 ±1.40	0.34 ±0.86	−0.13	−26	0.44 ±0.79	0.47 ±1.11	+0.03	+10	0.67 (0.45, 1.00)
Sports drinks	0.18 ±0.49	0.12 ±0.42	−0.06	−36	0.18 ±0.45	0.17 ±0.56	−0.01	+21	0.53 (0.31, 0.91)
Energy drinks	0.09 ±0.51	0.05 ±0.24	−0.04	−29	0.07 ±0.28	0.07 ±0.32	0.00	−14	0.83 (0.38, 1.82)
Fruit drinks	0.28 ±0.57	0.26 ±0.65	−0.03	−13	0.39 ±0.79	0.34 ±0.81	−0.06	−12	0.99 (0.69, 1.44)
Sweetened coffee or tea	0.23 ±0.57	0.21 ±0.61	−0.02	−13	0.21 ±0.56	0.21 ±0.59	0.00	+22	0.71 (0.44, 1.15)
Water[d]	3.50 ±3.24	5.84 ±10.38	+2.33	+63	3.98 ±3.12	4.69 ±3.53	+0.70	+19	1.37 (1.14, 1.64)

Note. CI = confidence interval; SSB = sugar-sweetened beverage.

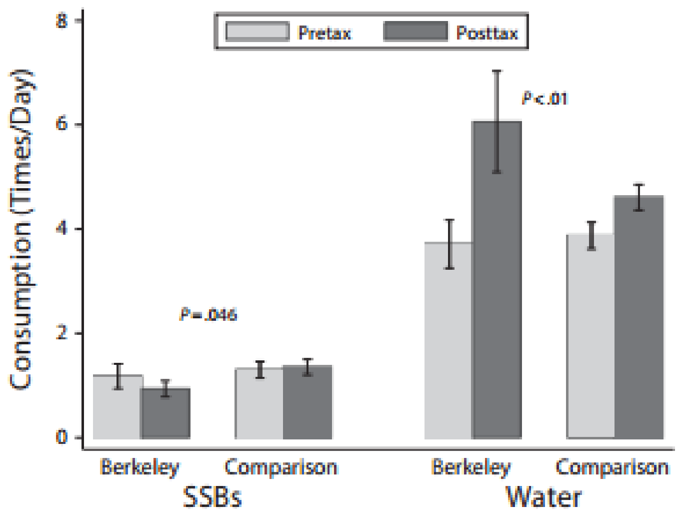
[a]Posttax data were collected approximately 12 months after pretax data collection, 8 months after elections, and 4 months after implementation of the tax.

[b]Adjusted for gender, age, education, race/ethnicity, language, and neighborhood in which the survey was conducted. Generalized linear models were used with a γ distribution, log link, and robust standard errors.

[c]From adjusted within-city ratio of post- to pretax consumption.

[d]Sample sizes for water included 2437—786 in Berkeley and 1651 in comparison cities.

Example: effect of soda tax



Note. Adjusted means and 95% confidence intervals were obtained by using the margins command in Stata/IC version 13.1 (StataCorp LP, College Station, TX) after running generalized linear models adjusting for neighborhood, gender, age, education, race/ethnicity, and language. *P* values shown are for the difference between Berkeley and comparison cities in change in consumption and come from the generalized linear models.

FIGURE 2—Adjusted Mean Consumption of Sugar-Sweetened Beverages (SSBs) and Water Before and After the Tax in Berkeley, CA, and Comparison Cities (Oakland and San Francisco, CA)

Finish Line

Conclusions

- Though the Board does not employ an economist, we can do more than just uncritically accept economic evidence at face value.
- First, consider the source of data.
- When presented with economic evidence, consider whether a correlation or causal inference is being presented.
- If evidence is purportedly causal, consider how that conclusion was made
- If it's just a correlation, consider what other contributing factors could come into play.

Questions?