Progressive Media Impact: Metrics, Evaluation, and Improvement

Gary King

Institute for Quantitative Social Science
Harvard University

(talk at The Media Consortium, San Francisco, 10/13/2011)

Goals

Goals

- Measure public discourse

Goals

- Measure public discourse
- Estimate the causal effect of the progressive media on discourse

Goals

- Measure public discourse
- Estimate the causal effect of the progressive media on discourse
- Learn together how to make the progressive media more effective

Goals

- Measure public discourse
- Estimate the causal effect of the progressive media on discourse
- Learn together how to make the progressive media more effective
- Take advantage of recent dramatic scientific advances:

Goals

- Measure public discourse
- Estimate the causal effect of the progressive media on discourse
- Learn together how to make the progressive media more effective
- Take advantage of recent dramatic scientific advances:
    - Informative data now available on the public debate

Goals

- Measure public discourse
- Estimate the causal effect of the progressive media on discourse
- Learn together how to make the progressive media more effective
- Take advantage of recent dramatic scientific advances:
    - Informative data now available on the public debate
    - Measurement methods: meaning, not word counts

Goals

- Measure public discourse
- Estimate the causal effect of the progressive media on discourse
- Learn together how to make the progressive media more effective
- Take advantage of recent dramatic scientific advances:
  - Informative data now available on the public debate
  - Measurement methods: meaning, not word counts
  - Experimental designs: avoiding confounding factors

Goals

- Measure public discourse
- Estimate the causal effect of the progressive media on discourse
- Learn together how to make the progressive media more effective
- Take advantage of recent dramatic scientific advances:
  - Informative data now available on the public debate
  - Measurement methods: meaning, not word counts
  - Experimental designs: avoiding confounding factors
  - Causal effect estimation: less bias & inefficiency

Goals

- Measure public discourse
- Estimate the causal effect of the progressive media on discourse
- Learn together how to make the progressive media more effective
- Take advantage of recent dramatic scientific advances:
    - Informative data now available on the public debate
    - Measurement methods: meaning, not word counts
    - Experimental designs: avoiding confounding factors
    - Causal effect estimation: less bias & inefficiency
    - New ways of working together to improve impact

Who Frames the Public Debate?

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.
- Was this massive opinion change, or did the 9/11 frame change how we viewed the question?

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.
- Was this massive opinion change, or did the 9/11 frame change how we viewed the question?
- The frame: imposed by events, not the media

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.
- Was this massive opinion change, or did the 9/11 frame change how we viewed the question?
- The frame: imposed by events, not the media
- Public opinion polls: measuring what?

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.
- Was this massive opinion change, or did the 9/11 frame change how we viewed the question?
- The frame: imposed by events, not the media
- Public opinion polls: measuring what?

Support for Ban on "Partial Birth" Abortion

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.
- Was this massive opinion change, or did the 9/11 frame change how we viewed the question?
- The frame: imposed by events, not the media
- Public opinion polls: measuring what?

Support for Ban on "Partial Birth" Abortion

- Supporters: use "baby"

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.
- Was this massive opinion change, or did the 9/11 frame change how we viewed the question?
- The frame: imposed by events, not the media
- Public opinion polls: measuring what?

Support for Ban on "Partial Birth" Abortion

- Supporters: use "baby"
- Opponents: use "fetus"

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.
- Was this massive opinion change, or did the 9/11 frame change how we viewed the question?
- The frame: imposed by events, not the media
- Public opinion polls: measuring what?

Support for Ban on "Partial Birth" Abortion

- Supporters: use "baby"
- Opponents: use "fetus"
- In surveys, "baby" increases support for the ban by $\approx 30\%$!

Who Frames the Public Debate?

"Do you approve of how George W. Bush is handling his job?"

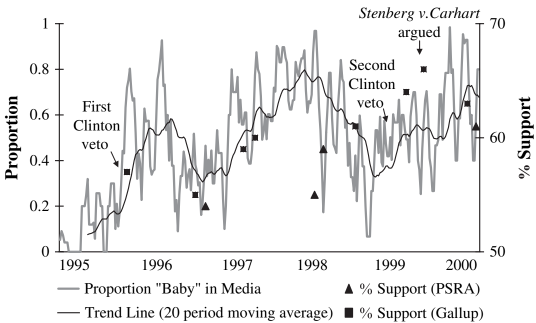
- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.
- Was this massive opinion change, or did the 9/11 frame change how we viewed the question?
- The frame: imposed by events, not the media
- Public opinion polls: measuring what?

Support for Ban on "Partial Birth" Abortion

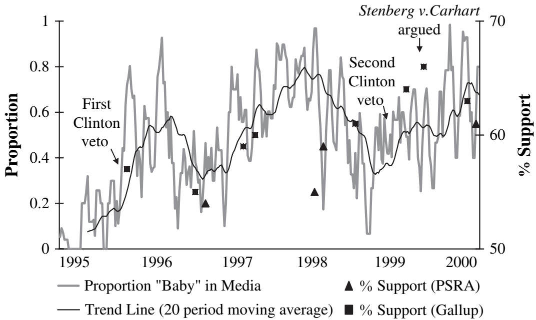
- Supporters: use "baby"
- Opponents: use "fetus"
- In surveys, "baby" increases support for the ban by $\approx$ 30%!
- Control the frame & you control the debate and policy outcome

“Partial Birth” Abortion Ban: Who controls the frame?

“Partial Birth” Abortion Ban: Who controls the frame?

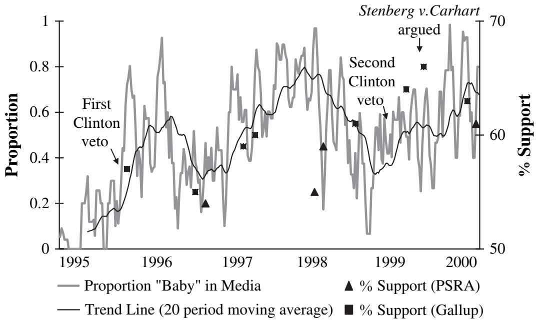


“Partial Birth” Abortion Ban: Who controls the frame?



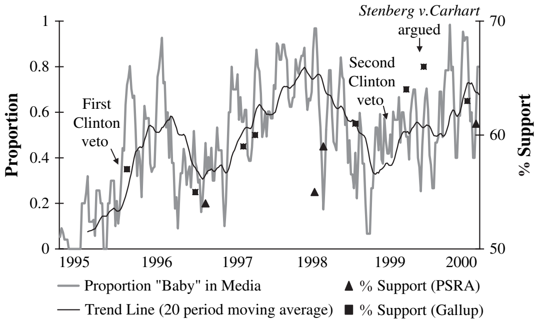
- When “baby” is used in the media, support for PBA ban increases

“Partial Birth” Abortion Ban: Who controls the frame?



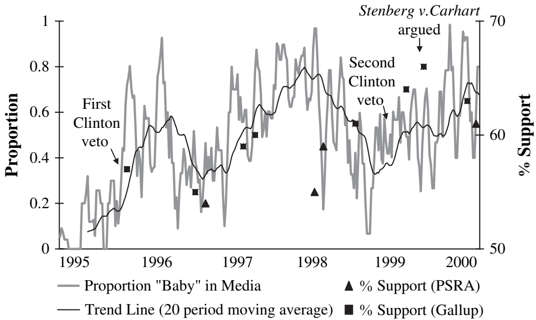
- When “baby” is used in the media, support for PBA ban increases
- Did the media influence public support or did public support influence the media?

"Partial Birth" Abortion Ban: Who controls the frame?



- When "baby" is used in the media, support for PBA ban increases
- Did the media influence public support or did public support influence the media? How can we tell?

"Partial Birth" Abortion Ban: Who controls the frame?



- When "baby" is used in the media, support for PBA ban increases
- Did the media influence public support or did public support influence the media? How can we tell?
- Do "baby" and "fetus" word counts even measure what we intend?

The Plan, 1: Collect New Data

The Plan, 1: Collect New Data

Data on Public Discourse

The Plan, 1: Collect New Data

Data on Public Discourse

- Until recently, measures of public discourse were inadequate: public opinion polls, content analyses of newspaper editorials; systematic, real time, informative data nonexistent

The Plan, 1: Collect New Data

Data on Public Discourse

- Until recently, measures of public discourse were inadequate: public opinion polls, content analyses of newspaper editorials; systematic, real time, informative data nonexistent
- Now:

The Plan, 1: Collect New Data

Data on Public Discourse

- Until recently, measures of public discourse were inadequate: public opinion polls, content analyses of newspaper editorials; systematic, real time, informative data nonexistent
- Now:
  - Hallway conversations appear in the 1.4B social media posts a week

The Plan, 1: Collect New Data

Data on Public Discourse

- Until recently, measures of public discourse were inadequate: public opinion polls, content analyses of newspaper editorials; systematic, real time, informative data nonexistent
- Now:
  - Hallway conversations appear in the 1.4B social media posts a week
  - Growing sources: blogs, Twitter, Facebook, forums, chat rooms, etc.

The Plan, 1: Collect New Data

Data on Public Discourse

- Until recently, measures of public discourse were inadequate: public opinion polls, content analyses of newspaper editorials; systematic, real time, informative data nonexistent
- Now:
  - Hallway conversations appear in the 1.4B social media posts a week
  - Growing sources: blogs, Twitter, Facebook, forums, chat rooms, etc.

Data on Media Outlets

Work together to find or create

The Plan, 1: Collect New Data

Data on Public Discourse

- Until recently, measures of public discourse were inadequate: public opinion polls, content analyses of newspaper editorials; systematic, real time, informative data nonexistent
- Now:
  - Hallway conversations appear in the 1.4B social media posts a week
  - Growing sources: blogs, Twitter, Facebook, forums, chat rooms, etc.

Data on Media Outlets

Work together to find or create

- Existing data on outputs (web traffic, donations, comments, emails, phone logs, etc.)

The Plan, 1: Collect New Data

Data on Public Discourse

- Until recently, measures of public discourse were inadequate: public opinion polls, content analyses of newspaper editorials; systematic, real time, informative data nonexistent
- Now:
  - Hallway conversations appear in the 1.4B social media posts a week
  - Growing sources: blogs, Twitter, Facebook, forums, chat rooms, etc.

Data on Media Outlets

Work together to find or create

- Existing data on outputs (web traffic, donations, comments, emails, phone logs, etc.)
- Existing data on content: e.g., tagged story databases, web artifacts, listener information, etc.

The Plan, 1: Collect New Data

Data on Public Discourse

- Until recently, measures of public discourse were inadequate: public opinion polls, content analyses of newspaper editorials; systematic, real time, informative data nonexistent
- Now:
    - Hallway conversations appear in the 1.4B social media posts a week
    - Growing sources: blogs, Twitter, Facebook, forums, chat rooms, etc.

Data on Media Outlets

Work together to find or create

- Existing data on outputs (web traffic, donations, comments, emails, phone logs, etc.)
- Existing data on content: e.g., tagged story databases, web artifacts, listener information, etc.
- New systematic data from texts of stories, audio-to-text, etc.

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:
  - Misleading word counts

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:
  - Misleading word counts
  - Impossible amounts of reading

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:
  - Misleading word counts
  - Impossible amounts of reading
- Now, new statistical methods for text analytics:

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:
  - Misleading word counts
  - Impossible amounts of reading
- Now, new statistical methods for text analytics:
  - Accurately summarize huge volumes of information

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:
  - Misleading word counts
  - Impossible amounts of reading
- Now, new statistical methods for text analytics:
  - Accurately summarize huge volumes of information
  - Better than humans alone, or computers alone, we amplify human intelligence

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:
  - Misleading word counts
  - Impossible amounts of reading
- Now, new statistical methods for text analytics:
  - Accurately summarize huge volumes of information
  - Better than humans alone, or computers alone, we amplify human intelligence
  - Works fast, in near real time

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:
  - Misleading word counts
  - Impossible amounts of reading
- Now, new statistical methods for text analytics:
  - Accurately summarize huge volumes of information
  - Better than humans alone, or computers alone, we amplify human intelligence
  - Works fast, in near real time
  - Can measure media frames (about politicians or policies) by volume, sentiment, topics, perspectives — or any categories we think of

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:
  - Misleading word counts
  - Impossible amounts of reading
- Now, new statistical methods for text analytics:
  - Accurately summarize huge volumes of information
  - Better than humans alone, or computers alone, we amplify human intelligence
  - Works fast, in near real time
  - Can measure media frames (about politicians or policies) by volume, sentiment, topics, perspectives — or any categories we think of
  - Measure as far back in time as feasible (6-24 months?)

The Plan, 2: Adapt New Methods to Understand Billions of Social Media Posts, & News Stories

- Until recently, methods were inadequate:
  - Misleading word counts
  - Impossible amounts of reading
- Now, new statistical methods for text analytics:
  - Accurately summarize huge volumes of information
  - Better than humans alone, or computers alone, we amplify human intelligence
  - Works fast, in near real time
  - Can measure media frames (about politicians or policies) by volume, sentiment, topics, perspectives — or any categories we think of
  - Measure as far back in time as feasible (6-24 months?)
  - Design systematic measurement going forward

[Some detail on innovations in text analytics]

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit, (2) an attack on public sector workers

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit, (2) an attack on public sector workers, (3) an attack on the progressive movement generally

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
    - Read & interpret: infeasible
    - Sort into a few categories; track category percentages over time
    - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit, (2) an attack on public sector workers, (3) an attack on the progressive movement generally, or (4) due to an arrogant publicity-hungry Governor

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit, (2) an attack on public sector workers, (3) an attack on the progressive movement generally, or (4) due to an arrogant publicity-hungry Governor
- How to sort billions of social media posts into categories?

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit, (2) an attack on public sector workers, (3) an attack on the progressive movement generally, or (4) due to an arrogant publicity-hungry Governor
- How to sort billions of social media posts into categories?
  - Classify by hand: infeasible for more than a sample at one point in time

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit, (2) an attack on public sector workers, (3) an attack on the progressive movement generally, or (4) due to an arrogant publicity-hungry Governor
- How to sort billions of social media posts into categories?
  - Classify by hand: infeasible for more than a sample at one point in time
  - Guess what words imply which categories: inaccurate

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit, (2) an attack on public sector workers, (3) an attack on the progressive movement generally, or (4) due to an arrogant publicity-hungry Governor
- How to sort billions of social media posts into categories?
  - Classify by hand: infeasible for more than a sample at one point in time
  - Guess what words imply which categories: inaccurate
  - Use "machine learning" methods with hand-coded training set to automatically classify: at best 60-70% accuracy (useful for Google searches, useless for category percentages)

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit, (2) an attack on public sector workers, (3) an attack on the progressive movement generally, or (4) due to an arrogant publicity-hungry Governor
- How to sort billions of social media posts into categories?
  - Classify by hand: infeasible for more than a sample at one point in time
  - Guess what words imply which categories: inaccurate
  - Use "machine learning" methods with hand-coded training set to automatically classify: at best 60-70% accuracy (useful for Google searches, useless for category percentages)
  - Use new statistical method with hand-coded training set to estimate category percentages (without individual classifications): Extremely accurate

[Some detail on innovations in text analytics]

- How can humans understand large numbers of social media posts (or news stories)?
  - Read & interpret: infeasible
  - Sort into a few categories; track category percentages over time
  - Example of categories: Wisconsin budget standoff is (1) a responsible effort to fix budget deficit, (2) an attack on public sector workers, (3) an attack on the progressive movement generally, or (4) due to an arrogant publicity-hungry Governor
- How to sort billions of social media posts into categories?
  - Classify by hand: infeasible for more than a sample at one point in time
  - Guess what words imply which categories: inaccurate
  - Use "machine learning" methods with hand-coded training set to automatically classify: at best 60-70% accuracy (useful for Google searches, useless for category percentages)
  - Use new statistical method with hand-coded training set to estimate category percentages (without individual classifications): Extremely accurate
  - $\rightsquigarrow$ Finding the needle in the haystack (Google search) $\neq$ characterizing the haystack, and so new methods were required

The Plan, 3: Use New Methods of Causal Inference for Observational Data

The Plan, 3: Use New Methods of Causal Inference for Observational Data

- Until recently: Estimating causal effects from observational data was very risky and often illusory

The Plan, 3: Use New Methods of Causal Inference for Observational Data

- Until recently: Estimating causal effects from observational data was very risky and often illusory
- Now, new methods:

The Plan, 3: Use New Methods of Causal Inference for Observational Data

- Until recently: Estimating causal effects from observational data was very risky and often illusory
- Now, new methods:
  - Give strong hints about causal relationships

The Plan, 3: Use New Methods of Causal Inference for Observational Data

- Until recently: Estimating causal effects from observational data was very risky and often illusory
- Now, new methods:
  - Give strong hints about causal relationships
  - Reveal all hidden assumptions (ignorability, interference, etc.)

The Plan, 3: Use New Methods of Causal Inference for Observational Data

- Until recently: Estimating causal effects from observational data was very risky and often illusory
- Now, new methods:
  - Give strong hints about causal relationships
  - Reveal all hidden assumptions (ignorability, interference, etc.)
  - Help inform what experiments to run

The Plan, 3: Use New Methods of Causal Inference for Observational Data

- Until recently: Estimating causal effects from observational data was very risky and often illusory
- Now, new methods:
  - Give strong hints about causal relationships
  - Reveal all hidden assumptions (ignorability, interference, etc.)
  - Help inform what experiments to run
- We'll use new methods to:

The Plan, 3: Use New Methods of Causal Inference for Observational Data

- Until recently: Estimating causal effects from observational data was very risky and often illusory
- Now, new methods:
  - Give strong hints about causal relationships
  - Reveal all hidden assumptions (ignorability, interference, etc.)
  - Help inform what experiments to run
- We'll use new methods to:
  - Estimate the effect of progressive media on use of media frames, & sentiment about them, in public discourse

The Plan, 3: Use New Methods of Causal Inference for Observational Data

- Until recently: Estimating causal effects from observational data was very risky and often illusory
- Now, new methods:
  - Give strong hints about causal relationships
  - Reveal all hidden assumptions (ignorability, interference, etc.)
  - Help inform what experiments to run
- We'll use new methods to:
  - Estimate the effect of progressive media on use of media frames, & sentiment about them, in public discourse
  - Develop strong testable hypotheses about what do next

The Plan, 3: Use New Methods of Causal Inference for Observational Data

- Until recently: Estimating causal effects from observational data was very risky and often illusory
- Now, new methods:
  - Give strong hints about causal relationships
  - Reveal all hidden assumptions (ignorability, interference, etc.)
  - Help inform what experiments to run
- We'll use new methods to:
  - Estimate the effect of progressive media on use of media frames, & sentiment about them, in public discourse
  - Develop strong testable hypotheses about what do next
  - We'll do everything that can be done this way before turning to experiments (and using your time!)

The Plan, 4: Use New Experimental Designs

The Plan, 4: Use New Experimental Designs

- Observational analysis: some evidence of what might have worked in the past

The Plan, 4: Use New Experimental Designs

- Observational analysis: some evidence of what might have worked in the past
- Experiments: gold standard for evaluating causal claims

The Plan, 4: Use New Experimental Designs

- Observational analysis: some evidence of what might have worked in the past
- Experiments: gold standard for evaluating causal claims
- Experiments benefits from: investigator control (usually randomized) over the "treatment" (drugs, stories)

The Plan, 4: Use New Experimental Designs

- Observational analysis: some evidence of what might have worked in the past
- Experiments: gold standard for evaluating causal claims
- Experiments benefits from: investigator control (usually randomized) over the "treatment" (drugs, stories)
- What randomized treatments get us: a variable unrelated to *all* confounders

The Plan, 4: Use New Experimental Designs

- Observational analysis: some evidence of what might have worked in the past
- Experiments: gold standard for evaluating causal claims
- Experiments benefits from: investigator control (usually randomized) over the “treatment” (drugs, stories)
- What randomized treatments get us: a variable unrelated to *all* confounders
- How experiments fail:

The Plan, 4: Use New Experimental Designs

- Observational analysis: some evidence of what might have worked in the past
- Experiments: gold standard for evaluating causal claims
- Experiments benefits from: investigator control (usually randomized) over the "treatment" (drugs, stories)
- What randomized treatments get us: a variable unrelated to *all* confounders
- How experiments fail:
  - Insensitivity to research subjects, political interests, or local context

The Plan, 4: Use New Experimental Designs

- Observational analysis: some evidence of what might have worked in the past
- Experiments: gold standard for evaluating causal claims
- Experiments benefits from: investigator control (usually randomized) over the "treatment" (drugs, stories)
- What randomized treatments get us: a variable unrelated to *all* confounders
- How experiments fail:
  - Insensitivity to research subjects, political interests, or local context
  - Examples: Drug trials; Mexico's Anti-Poverty program

The Plan, 4: Use New Experimental Designs

- Observational analysis: some evidence of what might have worked in the past
- Experiments: gold standard for evaluating causal claims
- Experiments benefits from: investigator control (usually randomized) over the "treatment" (drugs, stories)
- What randomized treatments get us: a variable unrelated to *all* confounders
- How experiments fail:
  - Insensitivity to research subjects, political interests, or local context
  - Examples: Drug trials; Mexico's Anti-Poverty program
  - $\rightsquigarrow$ solutions always exist; we'll find them together

The Plan, 4: Use New Experimental Designs

- Observational analysis: some evidence of what might have worked in the past
- Experiments: gold standard for evaluating causal claims
- Experiments benefits from: investigator control (usually randomized) over the "treatment" (drugs, stories)
- What randomized treatments get us: a variable unrelated to *all* confounders
- How experiments fail:
  - Insensitivity to research subjects, political interests, or local context
  - Examples: Drug trials; Mexico's Anti-Poverty program
  - $\rightsquigarrow$ solutions always exist; we'll find them together
- New experimental designs: let randomization survive disruption

The Plan, 5: Implementation Strategy

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:
  - Offer optional small grants to media outlets to encourage covering particular stories (everyone will get a chance, no requirements, all decision making over content remains with organization)

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:
    - Offer optional small grants to media outlets to encourage covering particular stories (everyone will get a chance, no requirements, all decision making over content remains with organization)
    - Find otherwise arbitrary decisions to randomize (e.g., a story we're indifferent to being above or below the fold; the order of stories to present)

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:
  - Offer optional small grants to media outlets to encourage covering particular stories (everyone will get a chance, no requirements, all decision making over content remains with organization)
  - Find otherwise arbitrary decisions to randomize (e.g., a story we're indifferent to being above or below the fold; the order of stories to present)
  - Example: Mexico's Health Insurance evaluation

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:
  - Offer optional small grants to media outlets to encourage covering particular stories (everyone will get a chance, no requirements, all decision making over content remains with organization)
  - Find otherwise arbitrary decisions to randomize (e.g., a story we're indifferent to being above or below the fold; the order of stories to present)
  - Example: Mexico's Health Insurance evaluation
- Our scarcest resource: your time (so observational work first)

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:
  - Offer optional small grants to media outlets to encourage covering particular stories (everyone will get a chance, no requirements, all decision making over content remains with organization)
  - Find otherwise arbitrary decisions to randomize (e.g., a story we're indifferent to being above or below the fold; the order of stories to present)
  - Example: Mexico's Health Insurance evaluation
- Our scarcest resource: your time (so observational work first)
- Experiments will

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:
  - Offer optional small grants to media outlets to encourage covering particular stories (everyone will get a chance, no requirements, all decision making over content remains with organization)
  - Find otherwise arbitrary decisions to randomize (e.g., a story we're indifferent to being above or below the fold; the order of stories to present)
  - Example: Mexico's Health Insurance evaluation
- Our scarcest resource: your time (so observational work first)
- Experiments will start small (probably individual media outlets)

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:
  - Offer optional small grants to media outlets to encourage covering particular stories (everyone will get a chance, no requirements, all decision making over content remains with organization)
  - Find otherwise arbitrary decisions to randomize (e.g., a story we're indifferent to being above or below the fold; the order of stories to present)
  - Example: Mexico's Health Insurance evaluation
- Our scarcest resource: your time (so observational work first)
- Experiments will start small (probably individual media outlets), be sequential (each will build on the previous)

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:
  - Offer optional small grants to media outlets to encourage covering particular stories (everyone will get a chance, no requirements, all decision making over content remains with organization)
  - Find otherwise arbitrary decisions to randomize (e.g., a story we're indifferent to being above or below the fold; the order of stories to present)
  - Example: Mexico's Health Insurance evaluation
- Our scarcest resource: your time (so observational work first)
- Experiments will start small (probably individual media outlets), be sequential (each will build on the previous), and grow in extent (larger experiments, with more outlets, as we learn more)

The Plan, 5: Implementation Strategy

- Examples of how to experiment without infringing on editorial discretion:
  - Offer optional small grants to media outlets to encourage covering particular stories (everyone will get a chance, no requirements, all decision making over content remains with organization)
  - Find otherwise arbitrary decisions to randomize (e.g., a story we're indifferent to being above or below the fold; the order of stories to present)
  - Example: Mexico's Health Insurance evaluation
- Our scarcest resource: your time (so observational work first)
- Experiments will start small (probably individual media outlets), be sequential (each will build on the previous), and grow in extent (larger experiments, with more outlets, as we learn more)
- Our goal: Turn evidence of what has worked into increasingly effective strategies

For More Information

GKing.Harvard.edu