How the News Media Activate Public Expression and Influence National Agendas

Gary King

Institute for Quantitative Social Science
Harvard University

IPSA-USP Summer School, University of São Paulo, Brazil, 1/16/2019

1 Based on joint work with Benjamin Schneer and Ariel White (Science 2017)
2 GaryKing.org
Introduction

Research Design

Results

Supporting Analyses

Implications
Statistical Problems: We Can’t Randomize

• Statistical Problems
  • Randomization: usually impossible
  • Endogeneity: media outlets compete for readers
  • Spillover: 1 intervention may affect all potential subjects

• Clever Research Designs (trying to approximate randomization)
  • New TV tower. Some behind hill, in radio shadow
  • Before/after studies of “surprise” media events
  • Roll out of Fox News to some towns and not others
  • Many others…

• But we still can’t randomize
  • Assumptions: better, but unavoidably dubious
    ⇝ “Profound biases,” > 600% difference from truth
  • Estimands: different, of sometimes questionable relevance
Statistical Problems: We Can’t Randomize

- Statistical Problems
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - **Randomization:** usually impossible
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - **Randomization**: usually impossible
  - **Endogeneity**: media outlets compete for readers
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - **Randomization**: usually impossible
  - **Endogeneity**: media outlets compete for readers
  - **Spillover**: 1 intervention may affect all potential subjects
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - Randomization: usually impossible
  - Endogeneity: media outlets compete for readers
  - Spillover: 1 intervention may affect all potential subjects

- **Clever Research Designs (trying to approximate randomization)**
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - Randomization: usually impossible
  - Endogeneity: media outlets compete for readers
  - Spillover: 1 intervention may affect all potential subjects

- **Clever Research Designs (trying to approximate randomization)**
  - New TV tower. Some behind hill, in radio shadow
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - **Randomization:** usually impossible
  - **Endogeneity:** media outlets compete for readers
  - **Spillover:** 1 intervention may affect all potential subjects

- **Clever Research Designs (trying to approximate randomization)**
  - New TV tower. Some behind hill, in radio shadow
  - Before/after studies of “surprise” media events

Assumptions: better, but unavoidably dubious

⇝ “Profound biases,” > 600% difference from truth

Estimands: different, of sometimes questionable relevance
Statistical Problems: We Can’t Randomize

• **Statistical Problems**
  - Randomization: usually impossible
  - Endogeneity: media outlets compete for readers
  - Spillover: 1 intervention may affect all potential subjects

• **Clever Research Designs (trying to approximate randomization)**
  - New TV tower. Some behind hill, in radio shadow
  - Before/after studies of “surprise” media events
  - Roll out of Fox News to some towns and not others

• Assumptions: better, but unavoidably dubious
  → "Profound biases," > 600% difference from truth

• Estimands: different, of sometimes questionable relevance
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - **Randomization:** usually impossible
  - **Endogeneity:** media outlets compete for readers
  - **Spillover:** 1 intervention may affect all potential subjects

- **Clever Research Designs (trying to approximate randomization)**
  - New TV tower. Some behind hill, in radio shadow
  - Before/after studies of “surprise” media events
  - Roll out of Fox News to some towns and not others
  - Many others...
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - Randomization: usually impossible
  - Endogeneity: media outlets compete for readers
  - Spillover: 1 intervention may affect all potential subjects

- **Clever Research Designs (trying to approximate randomization)**
  - New TV tower. Some behind hill, in radio shadow
  - Before/after studies of “surprise” media events
  - Roll out of Fox News to some towns and not others
  - Many others...

- **But we still can’t randomize**
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - **Randomization:** usually impossible
  - **Endogeneity:** media outlets compete for readers
  - **Spillover:** 1 intervention may affect all potential subjects

- **Clever Research Designs (trying to approximate randomization)**
  - New TV tower. Some behind hill, in radio shadow
  - Before/after studies of “surprise” media events
  - Roll out of Fox News to some towns and not others
  - Many others...

- **But we still can’t randomize**
  - **Assumptions:** better, but unavoidably dubious
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - Randomization: usually impossible
  - Endogeneity: media outlets compete for readers
  - Spillover: 1 intervention may affect all potential subjects

- **Clever Research Designs (trying to approximate randomization)**
  - New TV tower. Some behind hill, in radio shadow
  - Before/after studies of “surprise” media events
  - Roll out of Fox News to some towns and not others
  - Many others...

- **But we still can’t randomize**
  - Assumptions: better, but unavoidably dubious
    \[ \Rightarrow \text{“Profound biases,” } > 600\% \text{ difference from truth} \]
Statistical Problems: We Can’t Randomize

- **Statistical Problems**
  - **Randomization:** usually impossible
  - **Endogeneity:** media outlets compete for readers
  - **Spillover:** 1 intervention may affect all potential subjects

- **Clever Research Designs (trying to approximate randomization)**
  - New TV tower. Some behind hill, in radio shadow
  - Before/after studies of “surprise” media events
  - Roll out of Fox News to some towns and not others
  - Many others...

- **But we still can’t randomize**
  - **Assumptions:** better, but unavoidably dubious
    \( \rightsquigarrow \) “Profound biases,” > 600% difference from truth
  - **Estimands:** different, of sometimes questionable relevance
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets
  • Randomize news content and timing for each
  • Control collaboration to induce cross-outlet correlations

• Why is this plan so hard for media outlets?
  • Need to take actions few (if any) have ever before agreed to
  • Outlets are competitors: trying to scoop each other
  • Must share information with us (even if not with each other)
  • Need numerous agreements, technical infrastructure for large scale collaboration & data collection, extensive coordination, high levels of trust

• More specifically, to randomize
  • Journalists require: total control over what’s published & when
  • Scientists require: total control over what’s published & when
Political Problems: They Won’t Let Us Randomize

- What we’d do without constraints
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets

Introduction
Political Problems: They Won’t Let Us Randomize

- **What we’d do without constraints**
  - Sign up many news media outlets
  - Randomize news content and timing for each

---

Introduction
Political Problems: They Won’t Let Us Randomize

- **What we’d do without constraints**
  - Sign up many news media outlets
  - Randomize news content and timing for each
  - Control collaboration to induce cross-outlet correlations
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets
  • Randomize news content and timing for each
  • Control collaboration to induce cross-outlet correlations

• Why is this plan so hard for media outlets?
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets
  • Randomize news content and timing for each
  • Control collaboration to induce cross-outlet correlations

• Why is this plan so hard for media outlets?
  • Need to take actions few (if any) have ever before agreed to
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets
  • Randomize news content and timing for each
  • Control collaboration to induce cross-outlet correlations

• Why is this plan so hard for media outlets?
  • Need to take actions few (if any) have ever before agreed to
  • Outlets are competitors: trying to scoop each other
• **What we’d do without constraints**
  - Sign up many news media outlets
  - Randomize news content and timing for each
  - Control collaboration to induce cross-outlet correlations

• **Why is this plan so hard for media outlets?**
  - Need to take actions few (if any) have ever before agreed to
  - Outlets are competitors: trying to scoop each other
  - Must share information with us (even if not with each other)
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets
  • Randomize news content and timing for each
  • Control collaboration to induce cross-outlet correlations

• Why is this plan so hard for media outlets?
  • Need to take actions few (if any) have ever before agreed to
  • Outlets are competitors: trying to scoop each other
  • Must share information with us (even if not with each other)
  • Need numerous agreements,
Political Problems: They Won’t Let Us Randomize

• **What we’d do without constraints**
  - Sign up many news media outlets
  - Randomize news content and timing for each
  - Control collaboration to induce cross-outlet correlations

• **Why is this plan so hard for media outlets?**
  - Need to take actions few (if any) have ever before agreed to
  - Outlets are competitors: trying to scoop each other
  - Must share information with us (even if not with each other)
  - Need numerous agreements, technical infrastructure for large scale collaboration & data collection,
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets
  • Randomize news content and timing for each
  • Control collaboration to induce cross-outlet correlations

• Why is this plan so hard for media outlets?
  • Need to take actions few (if any) have ever before agreed to
  • Outlets are competitors: trying to scoop each other
  • Must share information with us (even if not with each other)
  • Need numerous agreements, technical infrastructure for large scale collaboration & data collection, extensive coordination,
Political Problems: They Won’t Let Us Randomize

• **What we’d do without constraints**
  - Sign up many news media outlets
  - Randomize news content and timing for each
  - Control collaboration to induce cross-outlet correlations

• **Why is this plan so hard for media outlets?**
  - Need to take actions few (if any) have ever before agreed to
  - Outlets are competitors: trying to scoop each other
  - Must share information with us (even if not with each other)
  - Need numerous agreements, technical infrastructure for large scale collaboration & data collection, extensive coordination, high levels of trust
Political Problems: They Won’t Let Us Randomize

- What we’d do without constraints
  - Sign up many news media outlets
  - Randomize news content and timing for each
  - Control collaboration to induce cross-outlet correlations

- Why is this plan so hard for media outlets?
  - Need to take actions few (if any) have ever before agreed to
  - Outlets are competitors: trying to scoop each other
  - Must share information with us (even if not with each other)
  - Need numerous agreements, technical infrastructure for large scale collaboration & data collection, extensive coordination, high levels of trust

- More specifically, to randomize
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets
  • Randomize news content and timing for each
  • Control collaboration to induce cross-outlet correlations

• Why is this plan so hard for media outlets?
  • Need to take actions few (if any) have ever before agreed to
  • Outlets are competitors: trying to scoop each other
  • Must share information with us (even if not with each other)
  • Need numerous agreements, technical infrastructure for large scale collaboration & data collection, extensive coordination, high levels of trust

• More specifically, to randomize
  • Journalists require:
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets
  • Randomize news content and timing for each
  • Control collaboration to induce cross-outlet correlations

• Why is this plan so hard for media outlets?
  • Need to take actions few (if any) have ever before agreed to
  • Outlets are competitors: trying to scoop each other
  • Must share information with us (even if not with each other)
  • Need numerous agreements, technical infrastructure for large scale collaboration & data collection, extensive coordination, high levels of trust

• More specifically, to randomize
  • Journalists require: total control over what’s published & when
Political Problems: They Won’t Let Us Randomize

• What we’d do without constraints
  • Sign up many news media outlets
  • Randomize news content and timing for each
  • Control collaboration to induce cross-outlet correlations

• Why is this plan so hard for media outlets?
  • Need to take actions few (if any) have ever before agreed to
  • Outlets are competitors: trying to scoop each other
  • Must share information with us (even if not with each other)
  • Need numerous agreements, technical infrastructure for large scale collaboration & data collection, extensive coordination, high levels of trust

• More specifically, to randomize
  • Journalists require: total control over what’s published & when
  • Scientists require:
Political Problems: They Won’t Let Us Randomize

- **What we’d do without constraints**
  - Sign up many news media outlets
  - Randomize news content and timing for each
  - Control collaboration to induce cross-outlet correlations

- **Why is this plan so hard for media outlets?**
  - Need to take actions few (if any) have ever before agreed to
  - Outlets are competitors: trying to scoop each other
  - Must share information with us (even if not with each other)
  - Need numerous agreements, technical infrastructure for large scale collaboration & data collection, extensive coordination, high levels of trust

- **More specifically, to randomize**
  - **Journalists require:** total control over what’s published & when
  - **Scientists require:** total control over what’s published & when
Our Approach:

• Build trust: 5 years of negotiating & communicating
• Develop incentive compatible research design: both get 100%, no compromises; ⇝ solve a political problem technologically
• Convince 48 media outlets to let us experiment on them
• Whenever possible, choose realism (even if inconvenient)
• Stick close to outlets’ standard operating procedures
• Embed treatment within ordinary routines
• ⇝ More expensive, logistically complicated, and time-consuming, but more generalizable
• Goal: Build platform to continue experiments
• A work of: political science
Our Approach: Let’s Randomize

• Build trust: 5 years of negotiating & communicating

• Develop incentive compatible research design: both get 100%, no compromises; ⇝ solve a political problem technologically

• Convince 48 media outlets to let us experiment on them

• Whenever possible, choose realism (even if inconvenient)

• Stick close to outlets’ standard operating procedures

• Embed treatment within ordinary routines

• More expensive, logistically complicated, and time-consuming, but more generalizable

• Goal: Build platform to continue experiments

• A work of: political science
Our Approach: Let’s Randomize

• **Build trust:** 5 years of negotiating & communicating
Our Approach: Let’s Randomize

- **Build trust:** 5 years of negotiating & communicating
- **Develop incentive compatible research design:**
Our Approach: Let’s Randomize

- **Build trust**: 5 years of negotiating & communicating
- **Develop incentive compatible research design**: both get 100%, no compromises;
Our Approach: Let’s Randomize

• **Build trust**: 5 years of negotiating & communicating
• **Develop incentive compatible research design**: both get 100%, no compromises; \( \rightsquigarrow \) solve a political problem technologically
Our Approach: Let’s Randomize

- **Build trust**: 5 years of negotiating & communicating
- **Develop incentive compatible research design**: both get 100%, no compromises; \(\rightsquigarrow\) solve a political problem technologically
- **Convince 48 media outlets to let us experiment on them**
Our Approach: Let’s Randomize

- **Build trust:** 5 years of negotiating & communicating
- **Develop incentive compatible research design:** both get 100%, no compromises; ⇝ solve a political problem technologically
- Convince 48 media outlets to let us experiment on them
- Whenever possible, choose realism (even if inconvenient)
Our Approach: Let’s Randomize

- **Build trust**: 5 years of negotiating & communicating
- **Develop incentive compatible research design**: both get 100%, no compromises; \( \rightsquigarrow \) solve a political problem technologically
- Convince 48 media outlets to let us experiment on them
- Whenever possible, choose realism (even if inconvenient)
  - Stick close to outlets’ standard operating procedures
Our Approach: Let’s Randomize

- **Build trust**: 5 years of negotiating & communicating
- **Develop incentive compatible research design**: both get 100%, no compromises; $\leadsto$ solve a political problem technologically
- **Convince 48 media outlets to let us experiment on them**
- **Whenever possible, choose realism (even if inconvenient)**
  - Stick close to outlets’ standard operating procedures
  - Embed treatment within ordinary routines
Our Approach: Let’s Randomize

• **Build trust:** 5 years of negotiating & communicating

• **Develop incentive compatible research design:** both get 100%, no compromises; ⇝ solve a political problem technologically

• **Convince 48 media outlets to let us experiment on them**

• **Whenever possible, choose realism (even if inconvenient)**
  
  • Stick close to outlets’ standard operating procedures
  • Embed treatment within ordinary routines
  • ⇝ More expensive, logistically complicated, and time-consuming, but more generalizable

Goal:
Build platform to continue experiments

A work of: political science
Our Approach: Let’s Randomize

- **Build trust:** 5 years of negotiating & communicating
- **Develop incentive compatible research design:** both get 100%, no compromises; ⇝ solve a political problem technologically
- **Convince 48 media outlets to let us experiment on them**
- **Whenever possible, choose realism (even if inconvenient)**
  - Stick close to outlets’ standard operating procedures
  - Embed treatment within ordinary routines
  - ⇝ More expensive, logistically complicated, and time-consuming, but more generalizable
- **Goal:** Build platform to continue experiments
Our Approach: Let’s Randomize

- **Build trust**: 5 years of negotiating & communicating
- **Develop incentive compatible research design**: both get 100%, no compromises; $\implies$ solve a political problem technologically
- **Convince 48 media outlets to let us experiment on them**
- **Whenever possible, choose realism (even if inconvenient)**
  - Stick close to outlets’ standard operating procedures
  - Embed treatment within ordinary routines
  - $\implies$ More expensive, logistically complicated, and time-consuming, but more generalizable
- **Goal**: Build **platform** to continue experiments
- **A work of**: political science
Our Approach: Let’s Randomize

- **Build trust**: 5 years of negotiating & communicating
- **Develop incentive compatible research design**: both get 100%, no compromises; ⇝ solve a political problem technologically
- **Convince 48 media outlets to let us experiment on them**
- **Whenever possible, choose realism (even if inconvenient)**
  - Stick close to outlets’ standard operating procedures
  - Embed treatment within ordinary routines
  - ⇝ More expensive, logistically complicated, and time-consuming, but more generalizable
- **Goal**: Build platform to continue experiments
- **A work of**: political science
Our Approach: Let’s Randomize

- **Build trust**: 5 years of negotiating & communicating
- **Develop incentive compatible research design**: both get 100%, no compromises; \( \rightsquigarrow \) solve a political problem technologically
- **Convince 48 media outlets to let us experiment on them**
- **Whenever possible, choose realism (even if inconvenient)**
  - Stick close to outlets’ standard operating procedures
  - Embed treatment within ordinary routines
  - \( \rightsquigarrow \) More expensive, logistically complicated, and time-consuming, but more generalizable
- **Goal**: Build platform to continue experiments
- **A work of**: political science
Define Outcome Variable: Types of News Media Effects

• Individual-level Effects
  - Outcome variable: individual knowledge and opinion
  - Effects: persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - Measurement: survey research

• Collective Effects: Impact on the national conversation
  - Outcome variable: activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - Classic definition of public opinion, predating survey research
  - Measurement
    - Previously: hallway conversations, "water-cooler events", soapbox speeches in public squares, editorials, etc.
    - Now: 750M public social media posts/day
  - Target population: different than survey research!
  - Surveys: pop quizzes of everyone, even uninformed & inactive
  - Social media: counts only activated opinion

• Democracies: Can ignore individuals, but collective expression sets agendas
• Autocracies: Ignore criticism, but censor expression about collective action
Define Outcome Variable: Types of News Media Effects

- Individual-level Effects

  - Outcome variable: individual knowledge and opinion
  - Effects: persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - Measurement: survey research

- Collective Effects: Impact on the national conversation

  - Outcome variable: activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - Classic definition of public opinion, predating survey research
  - Measurement:
    - Previously: hallway conversations, "water-cooler events", soapbox speeches in public squares, editorials, etc.
    - Now: 750M public social media posts/day
  - Target population: different than survey research!
    - Surveys: pop quizzes of everyone, even uninformed & inactive
    - Social media: counts only activated opinion
  - Democracies: can ignore individuals, but collective expression sets agendas
  - Autocracies: ignore criticism, but censor expression about collective action
Define Outcome Variable: Types of News Media Effects

- Individual-level Effects

- Collective Effects: Impact on the national conversation
Define Outcome Variable: Types of News Media Effects

- **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion

- **Collective Effects: Impact on the national conversation**
Define Outcome Variable: Types of News Media Effects

• **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion
  - **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing,

• **Collective Effects:** Impact on the national conversation
Define Outcome Variable: Types of News Media Effects

- **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion
  - **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - **Measurement:** survey research

- **Collective Effects:** Impact on the national conversation

- Previously: hallway conversations, “water-cooler events”, soapbox speeches in public squares, editorials, etc.
  - **Now:** 750 million public social media posts/day

- Target population: different than survey research!
  - Surveys: pop quizzes of everyone, even uninformed & inactive
  - Social media: counts only activated opinion

- Democracies: Can ignore individuals, but collective expression sets agendas
- Autocracies: Ignore criticism, but censor expression about collective action
Define Outcome Variable: Types of News Media Effects

• **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion
  - **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - **Measurement:** survey research

• **Collective Effects:** Impact on the national conversation
  - **Outcome variable:** activated public opinion, views of all those trying to express themselves publicly about policy and politics
Define Outcome Variable: Types of News Media Effects

• **Individual-level Effects**
  • **Outcome variable:** individual knowledge and opinion
  • **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  • **Measurement:** survey research

• **Collective Effects: Impact on the national conversation**
  • **Outcome variable:** activated public opinion, views of all those trying to express themselves publicly about policy and politics
  • **Classic definition of public opinion,** predating survey research
Define Outcome Variable: Types of News Media Effects

- **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion
  - **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - **Measurement:** survey research

- **Collective Effects: Impact on the national conversation**
  - **Outcome variable:** activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - **Classic definition of public opinion,** predating survey research
  - **Measurement**
Define Outcome Variable: Types of News Media Effects

- **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion
  - **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - **Measurement:** survey research

- **Collective Effects: Impact on the national conversation**
  - **Outcome variable:** activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - **Classic definition of public opinion,** predating survey research
  - **Measurement**
    - **Previously:** hallway conversations, “water-cooler events”, soapbox speeches in public squares, editorials, etc.

- **Target population:** different than survey research!
- Surveys: pop quizzes of everyone, even uninformed & inactive
- Social media: counts only activated opinion
- **Democracies:** Can ignore individuals, but collective expression sets agendas
- **Autocracies:** Ignore criticism, but censor expression about collective action
Define Outcome Variable: Types of News Media Effects

- **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion
  - **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - **Measurement:** survey research

- **Collective Effects: Impact on the national conversation**
  - **Outcome variable:** activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - **Classic definition of public opinion,** predating survey research
  - **Measurement**
    - **Previously:** hallway conversations, “water-cooler events”, soapbox speeches in public squares, editorials, etc.
    - **Now:** 750M *public* social media posts/day
Define Outcome Variable: Types of News Media Effects

- **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion
  - **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - **Measurement:** survey research

- **Collective Effects: Impact on the national conversation**
  - **Outcome variable:** activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - **Classic definition of public opinion,** predating survey research
  - **Measurement**
    - **Previously:** hallway conversations, “water-cooler events”, soapbox speeches in public squares, editorials, etc.
    - **Now:** 750M public social media posts/day
  - **Target population:** different than survey research!
Define Outcome Variable: Types of News Media Effects

• **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion
  - **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - **Measurement:** survey research

• **Collective Effects: Impact on the national conversation**
  - **Outcome variable:** activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - **Classic definition of public opinion,** predating survey research
  - **Measurement**
    - **Previously:** hallway conversations, “water-cooler events”, soapbox speeches in public squares, editorials, etc.
    - **Now:** 750M *public* social media posts/day
  - **Target population:** different than survey research!
    - **Surveys:** pop quizzes of everyone, even uninformed & inactive
Define Outcome Variable: Types of News Media Effects

- **Individual-level Effects**
  - *Outcome variable*: individual knowledge and opinion
  - *Effects*: Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - *Measurement*: survey research

- **Collective Effects: Impact on the national conversation**
  - *Outcome variable*: activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - *Classic definition of public opinion*, predating survey research
  - *Measurement*
    - *Previously*: hallway conversations, “water-cooler events”, soapbox speeches in public squares, editorials, etc.
    - *Now*: 750M *public* social media posts/day
  - *Target population*: different than survey research!
    - *Surveys*: pop quizzes of everyone, even uninformed & inactive
    - *Social media*: counts only activated opinion
Define Outcome Variable: Types of News Media Effects

- **Individual-level Effects**
  - **Outcome variable:** individual knowledge and opinion
  - **Effects:** Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - **Measurement:** survey research

- **Collective Effects: Impact on the national conversation**
  - **Outcome variable:** activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - **Classic definition of public opinion,** predating survey research
  - **Measurement**
    - **Previously:** hallway conversations, “water-cooler events”, soapbox speeches in public squares, editorials, etc.
    - **Now:** 750M *public* social media posts/day
  - **Target population:** different than survey research!
    - **Surveys:** pop quizzes of everyone, even uninformed & inactive
    - **Social media:** counts only activated opinion
  - **Democracies:** Can ignore individuals, but collective expression sets agendas
Define Outcome Variable: Types of News Media Effects

- **Individual-level Effects**
  - **Outcome variable**: individual knowledge and opinion
  - **Effects**: Persuasion, attitude formation, diffusion, gatekeeping, priming, issue framing, etc.
  - **Measurement**: survey research

- **Collective Effects: Impact on the national conversation**
  - **Outcome variable**: activated public opinion, views of all those trying to express themselves publicly about policy and politics
  - **Classic definition of public opinion**, predating survey research
  - **Measurement**
    - Previously: hallway conversations, “water-cooler events”, soapbox speeches in public squares, editorials, etc.
    - Now: 750M public social media posts/day
  - **Target population**: different than survey research!
    - **Surveys**: pop quizzes of everyone, even uninformed & inactive
    - **Social media**: counts only activated opinion
  - **Democracies**: Can ignore individuals, but collective expression sets agendas
  - **Autocracies**: Ignore criticism, but censor expression about collective action
Setup

- Signup 48 small media outlets (& > 12 others just for info)
- 17 for trial runs, 33 in experiment, 2 in both
- Median size: The Progressive, 50,000 subscribers
- Examples:
  - Establish 11 broad policy areas
  - Rules: (a) major national importance; (b) interest to outlets
  - race, immigration, jobs, abortion, climate, food policy, water, education policy, refugees, domestic energy production, and reproductive rights
- Using 11 rather than 1: more representative; larger $n$ needed
Setup

• Signup 48 small media outlets (& > 12 others just for info)
Setup

• Signup 48 small media outlets (& > 12 others just for info)
  • 17 for trial runs, 33 in experiment, 2 in both
Setup

• Signup 48 small media outlets (& > 12 others just for info)
  • 17 for trial runs, 33 in experiment, 2 in both
  • Median size: *The Progressive*, 50,000 subscribers
Setup

- Signup 48 small media outlets (& > 12 others just for info)
  - 17 for trial runs, 33 in experiment, 2 in both
  - Median size: *The Progressive*, 50,000 subscribers
  - Examples:

---

Research Design
Setup

- Signup 48 small media outlets (& > 12 others just for info)
  - 17 for trial runs, 33 in experiment, 2 in both
  - Median size: *The Progressive*, 50,000 subscribers
  - Examples:
    - Establish 11 broad *policy areas*
Setup

• Signup 48 small media outlets (& > 12 others just for info)
  • 17 for trial runs, 33 in experiment, 2 in both
  • Median size: *The Progressive*, 50,000 subscribers
  • Examples:

  - Establish 11 broad *policy areas*
    • Rules: (a) major national importance; (b) interest to outlets
Setup

- Signup 48 small media outlets (& > 12 others just for info)
  - 17 for trial runs, 33 in experiment, 2 in both
  - Median size: *The Progressive*, 50,000 subscribers
  - Examples:
    - Establish 11 broad *policy areas*
      - Rules: (a) major national importance; (b) interest to outlets
      - race, immigration, jobs, abortion, climate, food policy, water, education policy, refugees, domestic energy production, and reproductive rights
Setup

• Signup 48 small media outlets (& > 12 others just for info)
  • 17 for trial runs, 33 in experiment, 2 in both
  • Median size: *The Progressive*, 50,000 subscribers
  • Examples:

- Establish 11 broad *policy areas*
  • Rules: (a) major national importance; (b) interest to outlets
  • race, immigration, jobs, abortion, climate, food policy, water, education policy, refugees, domestic energy production, and reproductive rights
  • Using 11 rather than 1: more representative; larger $n$ needed
Treatment

• We choose a policy area
• Outlets volunteer for a pack of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
• The pack chooses subject for articles
• We approve: If rejected outlets can publish outside experiment
• Requirement: No breaking news (stories may be held for weeks)
• Options: large investigations, interview-based journalism, opinion pieces, or others normally published by pack members
• Example. Policy area: technology policy. Subject: what Uber drivers think about driverless cars, or how a trade agreement affects hiring in Philadelphia
• Outlets Publish Simultaneously: (following usual procedures)
• One article on subject per pack member
• Distribute via website, print, video, podcast, etc.
• Promote via Google adwords, social media, email lists, SEO…
• Co- and cross-promote with outlets in same pack

Research Design
Treatment

• **We choose a policy area** (1 of 11)
Treatment

- We choose a policy area (1 of 11)
- Outlets volunteer for a pack of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
Treatment

- We choose a *policy area* (1 of 11)
- Outlets volunteer for a *pack* of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
- The pack chooses *subject* for articles
Treatment

- We choose a *policy area* (1 of 11)
- **Outlets volunteer for a pack** of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
- The pack chooses *subject* for articles

- Outlets Publish Simultaneously: (following usual procedures)
Treatment

- **We choose a policy area** (1 of 11)
- **Outlets volunteer for a pack** of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
- **The pack chooses subject** for articles
  - **We approve:** If rejected outlets can publish outside experiment

- **Outlets Publish Simultaneously:** (following usual procedures)
Treatment

• **We choose a policy area** (1 of 11)

• **Outlets volunteer for a pack** of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)

• **The pack chooses subject** for articles
  - **We approve:** If rejected outlets can publish outside experiment
  - **Requirement:** No breaking news (stories may be held for weeks)

• **Outlets Publish Simultaneously:** (following usual procedures)
Treatment

• **We choose a policy area** (1 of 11)
• **Outlets volunteer for a pack** of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
• **The pack chooses subject** for articles
  • **We approve:** If rejected outlets can publish outside experiment
  • **Requirement:** No breaking news (stories may be held for weeks)
  • **Options:** large investigations, interview-based journalism, opinion pieces, or others normally published by pack members

• **Outlets Publish Simultaneously:** (following usual procedures)
Treatment

- We choose a *policy area* (1 of 11)
- Outlets volunteer for a *pack* of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
- The pack chooses *subject* for articles
  - We approve: If rejected outlets can publish outside experiment
  - Requirement: No breaking news (stories may be held for weeks)
  - Options: large investigations, interview-based journalism, opinion pieces, or others normally published by pack members
  - Example.

- Outlets Publish Simultaneously: (following usual procedures)
Treatment

- We choose a *policy area* (1 of 11)
- Outlets volunteer for a *pack* of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
- The pack chooses *subject* for articles
  - *We approve:* If rejected outlets can publish outside experiment
  - *Requirement:* No breaking news (stories may be held for weeks)
  - *Options:* large investigations, interview-based journalism, opinion pieces, or others normally published by pack members
  - *Example.* *Policy area:* technology policy.

- Outlets Publish Simultaneously: (following usual procedures)
Treatment

- **We choose a policy area** (1 of 11)
- **Outlets volunteer for a pack** of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
- **The pack chooses subject** for articles
  - **We approve:** If rejected outlets can publish outside experiment
  - **Requirement:** No breaking news (stories may be held for weeks)
  - **Options:** large investigations, interview-based journalism, opinion pieces, or others normally published by pack members
  - **Example. Policy area:** technology policy.
    - **Subject:** what Uber drivers think about driverless cars,

- **Outlets Publish Simultaneously:** (following usual procedures)
Treatment

• We choose a policy area (1 of 11)
• Outlets volunteer for a pack of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
• The pack chooses subject for articles
  • We approve: If rejected outlets can publish outside experiment
  • Requirement: No breaking news (stories may be held for weeks)
  • Options: large investigations, interview-based journalism, opinion pieces, or others normally published by pack members
  • Example. Policy area: technology policy.
    Subject: what Uber drivers think about driverless cars, or how a trade agreement affects hiring in Philadelphia
• Outlets Publish Simultaneously: (following usual procedures)
Treatment

• **We choose a policy area** (1 of 11)

• **Outlets volunteer for a pack** of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)

• **The pack chooses subject** for articles
  - **We approve**: If rejected outlets can publish outside experiment
  - **Requirement**: No breaking news (stories may be held for weeks)
  - **Options**: large investigations, interview-based journalism, opinion pieces, or others normally published by pack members
  - **Example. Policy area**: technology policy.
    - **Subject**: what Uber drivers think about driverless cars,
    - *or* how a trade agreement affects hiring in Philadelphia

• **Outlets Publish Simultaneously**: (following usual procedures)
  - One article on subject per pack member
Treatment

- We choose a policy area (1 of 11)
- Outlets volunteer for a pack of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
- The pack chooses subject for articles
  - We approve: If rejected outlets can publish outside experiment
  - Requirement: No breaking news (stories may be held for weeks)
  - Options: large investigations, interview-based journalism, opinion pieces, or others normally published by pack members
  - Example. Policy area: technology policy.
    Subject: what Uber drivers think about driverless cars, or how a trade agreement affects hiring in Philadelphia
- Outlets Publish Simultaneously: (following usual procedures)
  - One article on subject per pack member
  - Distribute via website, print, video, podcast, etc.
Treatment

- **We choose a policy area** (1 of 11)
- **Outlets volunteer for a pack** of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
- **The pack chooses subject** for articles
  - **We approve:** If rejected outlets can publish outside experiment
  - **Requirement:** No breaking news (stories may be held for weeks)
  - **Options:** large investigations, interview-based journalism, opinion pieces, or others normally published by pack members
  - **Example.** Policy area: technology policy.
    Subject: what Uber drivers think about driverless cars, or how a trade agreement affects hiring in Philadelphia
- **Outlets Publish Simultaneously:** (following usual procedures)
  - One article on subject per pack member
  - Distribute via website, print, video, podcast, etc.
  - Promote via Google adwords, social media, email lists, SEO...
Treatment

• **We choose a policy area** (1 of 11)
• **Outlets volunteer for a pack** of 2–5 (with our approval), following “project manager” protocol (e.g., Panama Papers)
• **The pack chooses subject** for articles
  • **We approve:** If rejected outlets can publish outside experiment
  • **Requirement:** No breaking news (stories may be held for weeks)
  • **Options:** large investigations, interview-based journalism, opinion pieces, or others normally published by pack members
  • **Example. Policy area:** technology policy.
    • **Subject:** what Uber drivers think about driverless cars, or how a trade agreement affects hiring in Philadelphia
• **Outlets Publish Simultaneously:** (following usual procedures)
  • One article on subject per pack member
  • Distribute via website, print, video, podcast, etc.
  • Promote via Google adwords, social media, email lists, SEO...
  • Co- and cross-promote with outlets in same pack
Randomization

Matched Pair Randomization

- Select pair of weeks: matched on similarity of predicted news
- One coin flip: which week is treatment and which control
- Treatment week: publish & promote articles (usually Tuesday)
- Control week: no compensation or special actions

(Ex post: Predictions accurate; flips, news shocks uncorrelated)
Randomization

Matched Pair Randomization

• Select pair of weeks: matched on similarity of predicted news
• One coin flip: which week is treatment and which control
• Treatment week: publish & promote articles (usually Tuesday)
• Control week: no compensation or special actions

(Ex post: Predictions accurate; flips, news shocks uncorrelated)
Randomization

Matched Pair Randomization

- Select pair of weeks: matched on similarity of *predicted* news

[September 2015 Calendar]
Randomization

Matched Pair Randomization

- **Select pair of weeks:** matched on similarity of *predicted* news
- **One coin flip:** which week is treatment and which control

![Calendar Image]
Randomization

Matched Pair Randomization

- **Select pair of weeks:** matched on similarity of *predicted* news
- **One coin flip:** which week is treatment and which control
  - **Treatment week:** publish & promote articles (usually Tuesday)

Research Design

10/23
Randomization

Matched Pair Randomization

- **Select pair of weeks:** matched on similarity of *predicted* news
- **One coin flip:** which week is treatment and which control
  - **Treatment week:** publish & promote articles (usually Tuesday)
  - **Control week:** no compensation or special actions

![Calendar for SEPTEMBER 2015]

- **Treatment Week:**
  - Monday 7
  - Tuesday 8
  - Wednesday 9
  - Thursday 10
  - Friday 11
  - Saturday 12
  - Sunday 13

- **Control Week:**
  - Monday 14
  - Tuesday 15
  - Wednesday 16
  - Thursday 17
  - Friday 18
  - Saturday 19
  - Sunday 20
Randomization

Matched Pair Randomization

- **Select pair of weeks:** matched on similarity of *predicted* news
- **One coin flip:** which week is treatment and which control
  - **Treatment week:** publish & promote articles (usually Tuesday)
  - **Control week:** no compensation or special actions
Randomization

Matched Pair Randomization

- **Select pair of weeks**: matched on similarity of *predicted* news
- **One coin flip**: which week is treatment and which control
  - **Treatment week**: publish & promote articles (usually Tuesday)
  - **Control week**: no compensation or special actions

![Calendar](calendar.png)
Randomization

Matched Pair Randomization

- **Select pair of weeks**: matched on similarity of *predicted* news
- **One coin flip**: which week is treatment and which control
  - **Treatment week**: publish & promote articles (usually Tuesday)
  - **Control week**: no compensation or special actions
- **(Ex post)**: Predictions accurate; flips, news shocks uncorrelated
Randomization

Matched Pair Randomization

- **Select pair of weeks:** matched on similarity of *predicted* news
- **One coin flip:** which week is treatment and which control
  - **Treatment week:** publish & promote articles (usually Tuesday)
  - **Control week:** no compensation or special actions
- **(Ex post:** Predictions accurate; flips, news shocks uncorrelated)**

Reasoning
Randomization

Matched Pair Randomization

- **Select pair of weeks**: matched on similarity of *predicted* news
- **One coin flip**: which week is treatment and which control
  - **Treatment week**: publish & promote articles (usually Tuesday)
  - **Control week**: no compensation or special actions
- **(Ex post)**: Predictions accurate; flips, news shocks uncorrelated

Reasoning

- **Cf. complete randomization:**
Randomization

Matched Pair Randomization

- **Select pair of weeks:** matched on similarity of *predicted* news
- **One coin flip:** which week is treatment and which control
  - **Treatment week:** publish & promote articles (usually Tuesday)
  - **Control week:** no compensation or special actions
- **(Ex post:** Predictions accurate; flips, news shocks uncorrelated)

Reasoning

- **Cf. complete randomization:** more power, efficiency, & “political” robustness;
Randomization

Matched Pair Randomization

- **Select pair of weeks:** matched on similarity of *predicted* news
- **One coin flip:** which week is treatment and which control
  - **Treatment week:** publish & promote articles (usually Tuesday)
  - **Control week:** no compensation or special actions
- **(Ex post:** Predictions accurate; flips, news shocks uncorrelated)

Reasoning

- **Cf. complete randomization:** more power, efficiency, & “political” robustness; less bias, model dependence, & research costs;
Randomization

Matched Pair Randomization

- **Select pair of weeks**: matched on similarity of *predicted* news
- **One coin flip**: which week is treatment and which control
  - **Treatment week**: publish & promote articles (usually Tuesday)
  - **Control week**: no compensation or special actions
- **(Ex post)**: Predictions accurate; flips, news shocks uncorrelated

Reasoning

- **Cf. complete randomization**: more power, efficiency, & “political” robustness; less bias, model dependence, & research costs; SEs as much as 600% smaller (Imai, King, Nall 2008)
Randomization

Matched Pair Randomization

- **Select pair of weeks:** matched on similarity of *predicted* news
- **One coin flip:** which week is treatment and which control
  - **Treatment week:** publish & promote articles (usually Tuesday)
  - **Control week:** no compensation or special actions
- (Ex post: Predictions accurate; flips, news shocks uncorrelated)

Reasoning

- Cf. complete randomization: more power, efficiency, & “political” robustness; less bias, model dependence, & research costs; SEs as much as 600% smaller (Imai, King, Nall 2008)
- Few experiments/outlet: Less interference; more heterogeneity
Randomization

Matched Pair Randomization

- Select pair of weeks: matched on similarity of predicted news
- *One* coin flip: which week is treatment and which control
  - Treatment week: publish & promote articles (usually Tuesday)
  - Control week: no compensation or special actions
- (Ex post: Predictions accurate; flips, news shocks uncorrelated)

Reasoning

- Cf. complete randomization: more power, efficiency, & “political” robustness; less bias, model dependence, & research costs; SEs as much as 600% smaller (Imai, King, Nall 2008)
- Few experiments/outlet: Less interference; more heterogeneity
- Nation as unit of treatment: no spillover, more cost
Randomization

Matched Pair Randomization

- **Select pair of weeks:** matched on similarity of *predicted* news
- **One coin flip:** which week is treatment and which control
  - **Treatment week:** publish & promote articles (usually Tuesday)
  - **Control week:** no compensation or special actions
- *(Ex post: Predictions accurate; flips, news shocks uncorrelated)*

Reasoning

- **Cf. complete randomization:** more power, efficiency, & “political” robustness; less bias, model dependence, & research costs; SEs as much as 600% smaller (Imai, King, Nall 2008)
- **Few experiments/outlet:** Less interference; more heterogeneity
- **Nation as unit of treatment:** no spillover, more cost
- *(Ex post: Automated text analysis & qualitative evidence: indistinguishable from normal publications & practices; no outlet received a single complaint)*
Quantities of Interest (& observable implications)
Quantities of Interest (& observable implications)

Random Treatment → Articles Published → Pageviews → Posts on Subject → Posts in Policy Area
Quantities of Interest (& observable implications)

- Random Treatment
- Articles Published
- Pageviews
- Posts on Subject
- Posts in Policy Area

• Intervention

Methods: readme, 2010; readme2, 2019

Social media: King, Pan, Roberts (2017)

Downloads from outlets

Special access to Google Analytics
Quantities of Interest (& observable implications)

- Random Treatment
- Articles Published
- Pageviews
- Posts on Subject
- Posts in Policy Area

• Intervention

• Social media: Crimson Hexagon, Inc.; Methods: readme, 2010; readme2, 2019
Quantities of Interest (& observable implications)

Random Treatment → Articles Published → Pageviews → Posts on Subject → Posts in Policy Area

- Intervention
- Downloads from outlets

- Social media: Crimson Hexagon, Inc.; Methods: readme, 2010; readme2, 2019
Quantities of Interest (& observable implications)

- Intervention
- Downloads from outlets
- Special access to Google Analytics

- Social media: Crimson Hexagon, Inc.; Methods: readme, 2010; readme2, 2019
Quantities of Interest (& observable implications)

- Intervention
- Downloads from outlets
- Special access to Google Analytics
- Social media: King, Pan, Roberts (2017)
- Social media: Crimson Hexagon, Inc.; Methods: readme, 2010; readme2, 2019
Determining $n$ via Sequential Hypothesis Testing

- Most analysts: fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    $\Rightarrow$ neither is acceptable with such massive logistical costs
- Power calculations: require knowing QOI!
- Better: fix $p$-value, run experiment sequentially, discover $n$
  - Collect only as much data as you need
    (Why should you be in grad school longer than necessary?)
  - Valid statistically under likelihood or Bayes
    (Careful of misinformation in some applied literatures)
- Need to check sensitivity to priors and models
- We introduce new methods to:
  - Evaluate robustness under frequentist theory
  - Remove parametric assumptions

Research Design
Determining $n$ via Sequential Hypothesis Testing

- Most analysts: fix $n$, run experiment, discover $p$-value
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts:** fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources

Research Design
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts**: fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts**: fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    ~~ neither is acceptable with such massive logistical costs

---

Research Design
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts**: fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    $\leadsto$ neither is acceptable with such massive logistical costs
- **Power calculations**: require knowing QOI!
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts**: fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    $\implies$ neither is acceptable with such massive logistical costs
- **Power calculations**: require knowing QOI!
- **Better**: fix $p$-value, run experiment sequentially, discover $n$
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts**: fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    - $\Rightarrow$ neither is acceptable with such massive logistical costs
  - **Power calculations**: require knowing QOI!
- **Better**: fix $p$-value, run experiment sequentially, discover $n$
  - Collect only as much data as you need
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts:** fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    $\leadsto$ neither is acceptable with such massive logistical costs
- **Power calculations:** require knowing QOI!
- **Better:** fix $p$-value, run experiment sequentially, discover $n$
  - Collect only as much data as you need
    (Why should you be in grad school longer than necessary?)
Determining \( n \) via Sequential Hypothesis Testing

- **Most analysts:** fix \( n \), run experiment, discover \( p \)-value
  - If \( n \) is too large: waste time & resources
  - If \( n \) is too small: waste the entire experiment
    \( \Rightarrow \) neither is acceptable with such massive logistical costs
- **Power calculations:** require knowing QOI!
- **Better:** fix \( p \)-value, run experiment sequentially, discover \( n \)
  - Collect only as much data as you need
    (Why should you be in grad school longer than necessary?)
  - Valid statistically under likelihood or Bayes
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts**: fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    $\leadsto$ neither is acceptable with such massive logistical costs
- **Power calculations**: require knowing QOI!
- **Better**: fix $p$-value, run experiment sequentially, discover $n$
  - Collect only as much data as you need
    (Why should you be in grad school *longer* than necessary?)
  - **Valid statistically** under likelihood or Bayes
    (Careful of misinformation in some applied literatures)
Determining $n$ via Sequential Hypothesis Testing

• **Most analysts:** fix $n$, run experiment, discover $p$-value
  • If $n$ is too large: waste time & resources
  • If $n$ is too small: waste the entire experiment
    $\rightsquigarrow$ neither is acceptable with such massive logistical costs
  • **Power calculations:** require knowing QOI!

• **Better:** fix $p$-value, run experiment sequentially, discover $n$
  • Collect only as much data as you need
    (Why should you be in grad school longer than necessary?)
  • **Valid statistically** under likelihood or Bayes
    (Careful of misinformation in some applied literatures)
  • Need to check sensitivity to priors and models
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts**: fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    $\Rightarrow$ neither is acceptable with such massive logistical costs
- **Power calculations**: require knowing QOI!
- **Better**: fix $p$-value, run experiment sequentially, discover $n$
  - Collect only as much data as you need
    (Why should you be in grad school longer than necessary?)
  - Valid statistically under likelihood or Bayes
    (Careful of misinformation in some applied literatures)
  - Need to check sensitivity to priors and models
  - We introduce new methods to:
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts:** fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    $\Rightarrow$ neither is acceptable with such massive logistical costs
  - Power calculations: require knowing QOI!

- **Better:** fix $p$-value, run experiment sequentially, discover $n$
  - Collect only as much data as you need
    (Why should you be in grad school *longer* than necessary?)
  - Valid statistically under likelihood or Bayes
    (Careful of misinformation in some applied literatures)
  - Need to check sensitivity to priors and models
  - We introduce new methods to:
    - Evaluate robustness under frequentist theory
Determining $n$ via Sequential Hypothesis Testing

- **Most analysts**: fix $n$, run experiment, discover $p$-value
  - If $n$ is too large: waste time & resources
  - If $n$ is too small: waste the entire experiment
    $\leadsto$ neither is acceptable with such massive logistical costs
- **Power calculations**: require knowing QOI!
- **Better**: fix $p$-value, run experiment sequentially, discover $n$
  - Collect only as much data as you need
    (Why should you be in grad school longer than necessary?)
  - Valid statistically under likelihood or Bayes
    (Careful of misinformation in some applied literatures)
  - Need to check sensitivity to priors and models
  - We introduce new methods to:
    - Evaluate robustness under frequentist theory
    - Remove parametric assumptions
Results from Sequential Hypothesis Tests

- **Our Stopping Rule:**
  \[ p \leq 0.05 \]
  - joint test: day 1, 2, 3, policy, subject;
  - for \( n, n-1, \& n-2 \)

- Recognizing more data is better and logistics are complicated (they might stop us!)

- **Empirical result:**
  \[ n = 70 \] (35 experiments)

- **Frequentist validation:**
  extensive [non]parametric tests
Results from Sequential Hypothesis Tests

- Our Stopping Rule:
  - $p \leq 0.05$
  - Joint test: day 1, 2, 3, policy, subject; for $n$, $n-1$, & $n-2$
  - Recognizing more data is better
  - And logistics are complicated (they might stop us!)

Empirical result: $n = 70$ (35 experiments)

Frequentist validation: extensive [non]parametric tests
Results from Sequential Hypothesis Tests

• Our Stopping Rule:
  • \( p \leq 0.05, \)
Results from Sequential Hypothesis Tests

- **Our Stopping Rule:**
  - \( p \leq 0.05 \), joint test: day 1, 2, 3, policy, subject;
Results from Sequential Hypothesis Tests

- **Our Stopping Rule:**
  - \( p \leq 0.05 \), joint test: day 1, 2, 3, policy, subject; for \( n, n - 1, \) & \( n - 2 \)
Results from Sequential Hypothesis Tests

• **Our Stopping Rule:**
  • \( p \leq 0.05 \), joint test: day 1,2,3, policy, subject; for \( n, n-1, \& n-2 \)
  • recognizing more data is better
Results from Sequential Hypothesis Tests

- **Our Stopping Rule:**
  - $p \leq 0.05$, joint test: day 1, 2, 3, policy, subject; for $n$, $n-1$, & $n-2$
  - recognizing more data is better
  - and logistics are complicated (they might stop us!)
Results from Sequential Hypothesis Tests

- **Our Stopping Rule:**
  - \( p \leq 0.05 \), joint test: day 1,2,3, policy, subject; for \( n, n - 1, \) \& \( n - 2 \)
  - recognizing more data is better
  - and logistics are complicated (they might stop us!)

- **Empirical result:**

\( n = 70 \) (35 experiments)
Results from Sequential Hypothesis Tests

• **Our Stopping Rule:**
  • $p \leq 0.05$, joint test: day 1,2,3, policy, subject; for $n$, $n-1$, & $n-2$
  • recognizing more data is better
  • and logistics are complicated (they might stop us!)

• **Empirical result:** $n = 70$ (35 experiments)
Results from Sequential Hypothesis Tests

- **Our Stopping Rule:**
  - $p \leq 0.05$, joint test: day 1, 2, 3, policy, subject; for $n$, $n-1$, & $n-2$
  - recognizing more data is better
  - and logistics are complicated (they might stop us!)

- **Empirical result:** $n = 70$ (35 experiments)
Results from Sequential Hypothesis Tests

- **Our Stopping Rule:**
  - \( p \leq 0.05 \), joint test: day 1, 2, 3, policy, subject; for \( n, n - 1, \) & \( n - 2 \)
  - recognizing more data is better
  - and logistics are complicated (they might stop us!)

- **Empirical result:** \( n = 70 \) (35 experiments)

- **Frequentist validation:** extensive [non]parametric tests
Main Causal Effect: Public Expression in Policy Areas

- **Red Dots:** model-based estimate (assumes linearity over days)
- **Open circles:** model-free estimate (no model, higher variance)
- **Causal effects:**
  - 1st day: 19.4% increase,
  - Total: 62.7% increase

- **Context:** 3 small media outlets have huge effect on the national conversation

---

Results
Main Causal Effect: Public Expression in Policy Areas

- **Red Dots:** model-based estimate (assumes linearity over days)
- **Open circles:** model-free estimate (no model, higher variance)

Causal effects:
- 1st day: 19.4% increase,
- Total: 62.7% increase

Context:
- 3 small media outlets have huge effect on the national conversation

Results
Main Causal Effect: Public Expression in Policy Areas

- Red Dots: model-based estimate (assumes linearity over days)
Main Causal Effect: Public Expression in Policy Areas

- **Red Dots**: model-based estimate (assumes linearity over days)
- **Open circles**: model-free estimate (no model, higher variance)

**Results:**
Main Causal Effect: Public Expression in Policy Areas

- **Red Dots**: model-based estimate (assumes linearity over days)
- **Open circles**: model-free estimate (no model, higher variance)
- **Causal effects:**

Results
Main Causal Effect: Public Expression in Policy Areas

- **Red Dots**: model-based estimate (assumes linearity over days)
- **Open circles**: model-free estimate (no model, higher variance)
- **Causal effects**: 1st day: 19.4% increase,
Main Causal Effect: Public Expression in Policy Areas

• **Red Dots:** model-based estimate (assumes linearity over days)
• **Open circles:** model-free estimate (no model, higher variance)
• **Causal effects:** 1st day: 19.4% increase, Total: 62.7% increase
Main Causal Effect: Public Expression in Policy Areas

- Red Dots: model-based estimate (assumes linearity over days)
- Open circles: model-free estimate (no model, higher variance)
- Causal effects: 1st day: 19.4% increase, Total: 62.7% increase
- Context: 3 small media outlets have huge effect on the national conversation
Causal Effect: Indistinguishable Across Subgroups
Causal Effect: Indistinguishable Across Subgroups

Region

% Change in Posts

Day 1 Day 2 Day 3 Day 4 Day 5 Day 6

Northeastern

Midwestern

Western

Southern

Results
Causal Effect: Indistinguishable Across Subgroups

Effect on the national conversation in major policy areas is national
Effect on the national conversation in major policy areas is national
Causal Effect: Indistinguishable Across Subgroups

Effect on the national conversation in major policy areas is national
Causal Effect: Indistinguishable Across Subgroups

Influence

% Change in Posts

Day 1
Day 2
Day 3
Day 4
Day 5
Day 6

High

Low

Effect on the national conversation in major policy areas is national
Causal Heterogeneity: Leave-One-Outlet-Out

• Red Dots: Original (model-based) estimates
• Open circles: same, with one outlet dropped from any packs
• Results: no dominant outlet; high heterogeneity
Causal Heterogeneity: Leave-One-Outlet-Out

Jackknife Estimation on Policy Area Effects

Red Dots: Original (model-based) estimates
Open circles: same, with one outlet dropped from any packs
Results: no dominant outlet; high heterogeneity
Causal Heterogeneity: Leave-One-Outlet-Out

Jackknife Estimation on Policy Area Effects

- Red Dots: Original (model-based) estimates
Causal Heterogeneity: Leave-One-Outlet-Out

Jackknife Estimation on Policy Area Effects

- **Red Dots**: Original (model-based) estimates
- **Open circles**: same, with one outlet dropped from any packs
Causal Heterogeneity: Leave-One-Outlet-Out

Jackknife Estimation on Policy Area Effects

- **Red Dots**: Original (model-based) estimates
- **Open circles**: same, with one outlet dropped from any packs
- **Results**: no dominant outlet; high heterogeneity
Supporting Analyses
High Experimental Compliance

- # Articles published by pack in policy area
- What's the goal?
- Average # media outlets per pack: 3.1
- Causal effect on # articles: 2.94
  \[\Rightarrow\] high compliance
- Pageviews (on subject of articles, relative to a day's volume)
- Causal effect on # pageviews: 969.6% (52,223 views) increase
  \[\Rightarrow\] high compliance

Supporting Analyses
High Experimental Compliance

- # Articles published by pack in policy area
High Experimental Compliance

- # Articles published by pack in policy area
  - What’s the goal? Average # media outlets per pack:

  - Average # media outlets per pack: 3.1
  - Causal effect on # articles: 2.94
    $$\Rightarrow$$ high compliance
  - Pageviews (on subject of articles, relative to a day’s volume)
    - Causal effect on # pageviews: 969.6% (52,223 views) increase
    $$\Rightarrow$$ high compliance
High Experimental Compliance

- # Articles published by pack in policy area
  - What’s the goal? Average # media outlets per pack: 3.1
High Experimental Compliance

- # Articles published by pack in policy area
  - What’s the goal? Average # media outlets per pack: 3.1
  - Causal effect on # articles:
High Experimental Compliance

- # Articles published by pack in policy area
  - What’s the goal? Average # media outlets per pack: 3.1
  - Causal effect on # articles: 2.94

Pageviews (on subject of articles, relative to a day’s volume)
- Causal effect on # pageviews: 969.6% (52,223 views) increase
  - High compliance
High Experimental Compliance

- # Articles published by pack in policy area
  - What’s the goal? Average # media outlets per pack: 3.1
  - Causal effect on # articles: 2.94
  - ⟹ high compliance
High Experimental Compliance

- # Articles published by pack in policy area
  - What’s the goal? Average # media outlets per pack: 3.1
  - Causal effect on # articles: 2.94
  - ⟹ high compliance

- Pageviews (on subject of articles, relative to a day’s volume)
High Experimental Compliance

• # Articles published by pack in policy area
  • What’s the goal? Average # media outlets per pack: 3.1
  • Causal effect on # articles: 2.94
  • $\Rightarrow$ high compliance

• Pageviews (on subject of articles, relative to a day’s volume)
  • Causal effect on # pageviews: 969.6% (52,223 views) increase
High Experimental Compliance

- **# Articles published by pack in policy area**
  - What’s the goal? Average # media outlets per pack: 3.1
  - Causal effect on # articles: 2.94
  - ⟹ high compliance

- **Pageviews** (on subject of articles, relative to a day’s volume)
  - Causal effect on # pageviews: 969.6% (52,223 views) increase
  - ⟹ high compliance

Supporting Analyses
Causal Effect on Subject of Articles

- Red Dots: model-based estimate (assumes linearity over days)
- Open circles: model-free estimate (no model, higher variance)
- Causal effects:
  - 1st day: 454% increase,
  - Total: 1,666% increase
Causal Effect on Subject of Articles

- Red Dots: model-based estimate (assumes linearity over days)
- Open circles: model-free estimate (no model, higher variance)

Causal effects:
- 1st day: 454% increase,
- Total: 1,666% increase
Causal Effect on Subject of Articles

- **Red Dots**: model-based estimate (assumes linearity over days)

**Supporting Analyses**
Causal Effect on Subject of Articles

- **Red Dots**: model-based estimate (assumes linearity over days)
- **Open circles**: model-free estimate (no model, higher variance)
Causal Effect on Subject of Articles

- **Red Dots**: model-based estimate (assumes linearity over days)
- **Open circles**: model-free estimate (no model, higher variance)
- **Causal effects:**
Causal Effect on Subject of Articles

- **Red Dots**: model-based estimate (assumes linearity over days)
- **Open circles**: model-free estimate (no model, higher variance)
- **Causal effects**: 1st day: 454% increase,
Causal Effect on Subject of Articles

- **Red Dots**: model-based estimate (assumes linearity over days)
- **Open circles**: model-free estimate (no model, higher variance)
- **Causal effects**: 1st day: 454% increase, Total: 1,666% increase

Supporting Analyses
Other Supporting Analyses

- More Results
- Opinion change: 2.3% change in direction of article opinion
- Large news media outlets: Observational evidence, >15x effect
- Robustness checks
  - # of unique authors: little change from effect on posts
  - Removing bots, retweets: No real change
  - Week 1 to 2 spillover, noncompliance: No evidence
  - Treatment articles: representative of all on complexity, type
Other Supporting Analyses

- More Results
Other Supporting Analyses

- **More Results**
  - **Opinion change:** 2.3% change in direction of article opinion

- **Large news media outlets:** Observational evidence, >15x effect

- **Robustness checks**
  - # of unique authors: little change from effect on posts
  - Removing bots, retweets: No real change
  - Week 1 to 2 spillover, noncompliance: No evidence

- **Treatment articles:** representative of all on complexity, type
Other Supporting Analyses

• More Results
  • Opinion change: 2.3% change in direction of article opinion
  • Large news media outlets: Observational evidence, >15x effect
Other Supporting Analyses

- More Results
  - Opinion change: 2.3% change in direction of article opinion
  - Large news media outlets: Observational evidence, >15x effect
- Robustness checks
Other Supporting Analyses

• **More Results**
  - Opinion change: 2.3% change in direction of article opinion
  - Large news media outlets: Observational evidence, >15x effect

• **Robustness checks**
  - # of unique authors: little change from effect on posts
Other Supporting Analyses

• **More Results**
  • Opinion change: 2.3% change in direction of article opinion
  • Large news media outlets: Observational evidence, >15x effect

• **Robustness checks**
  • # of unique authors: little change from effect on posts
  • Removing bots, retweets: No real change
Other Supporting Analyses

- **More Results**
  - Opinion change: 2.3% change in direction of article opinion
  - Large news media outlets: Observational evidence, >15x effect

- **Robustness checks**
  - # of unique authors: little change from effect on posts
  - Removing bots, retweets: No real change
  - Week 1 to 2 spillover, noncompliance: No evidence
Other Supporting Analyses

- **More Results**
  - Opinion change: 2.3% change in direction of article opinion
  - Large news media outlets: Observational evidence, >15x effect

- **Robustness checks**
  - # of unique authors: little change from effect on posts
  - Removing bots, retweets: No real change
  - Week 1 to 2 spillover, noncompliance: No evidence
  - Treatment articles: representative of all on complexity, type
Introduction

Research Design

Results

Supporting Analyses

Implications
Summary and Implications

• Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change

• Larger outlets: even bigger average effects

• Heterogeneous effects: large, highly variable viral patterns

• Implications: for individual journalists
  • Remarkable power; serious responsibility; not just another job

• Implications: for ecosystem of media outlets
  • Control over editorial boards and mastheads
  • Balance and diversity of outlet opinion
  • Effects of fake news
    • Impact on agendas, elections, public policy, discourse

• Journalism jobs: 25% drop in last decade

• What should be next?
  • We wrote a paper, built a platform, & showed how others can

• What experiment would you (or should we) run next?

For more information:
GaryKing.org/media
Summary and Implications

• Summary

• Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change

• Larger outlets: even bigger average effects

• Heterogeneous effects: large, highly variable viral patterns

• Implications: for individual journalists
  • Remarkable power; serious responsibility; not just another job

• Implications: for ecosystem of media outlets
  • Control over editorial boards and mastheads
  • Balance and diversity of outlet opinion
  • Effects of fake news
  • Impact on agendas, elections, public policy, discourse

• Journalism jobs: 25% drop in last decade

• What should be next?
  • We wrote a paper, built a platform, & showed how others can

For more information: GaryKing.org/media
Summary and Implications

• Summary

• Implications: for individual journalists
Summary and Implications

- Summary

- Implications: for individual journalists

- Implications: for ecosystem of media outlets
Summary and Implications

• Summary

• Implications: for individual journalists

• Implications: for ecosystem of media outlets

• What should be next?
Summary and Implications

• Summary
  • Small outlets: very large average effects

• Implications: for individual journalists

• Implications: for ecosystem of media outlets

• What should be next?

For more information: GaryKing.org/media
Summary and Implications

• Summary
  • Small outlets: very large average effects

• Implications: for individual journalists

• Implications: for ecosystem of media outlets

• What should be next?
Summary and Implications

• Summary
  • Small outlets: very large average effects on pageviews,

• Implications: for individual journalists

• Implications: for ecosystem of media outlets

• What should be next?
Summary and Implications

- **Summary**
  - **Small outlets:** very large average effects on pageviews, agenda (subject & policy),

- **Implications: for individual journalists**

- **Implications: for ecosystem of media outlets**

- **What should be next?**

For more information: GaryKing.org/media
Summary and Implications

• Summary
  • Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change

• Implications: for individual journalists

• Implications: for ecosystem of media outlets

• What should be next?
Summary and Implications

• **Summary**
  • Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change
  • Larger outlets: even bigger average effects

• Implications: for individual journalists

• Implications: for ecosystem of media outlets

• What should be next?

For more information: GaryKing.org/media
Summary and Implications

• Summary
  • Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change
  • Larger outlets: even bigger average effects
  • Heterogeneous effects: large, highly variable viral patterns

• Implications: for individual journalists

• Implications: for ecosystem of media outlets

• What should be next?
Summary and Implications

- **Summary**
  - **Small outlets:** very large average effects on pageviews, agenda (subject & policy), opinion change
  - **Larger outlets:** even bigger average effects
  - **Heterogeneous effects:** large, highly variable viral patterns

- **Implications: for individual journalists**
  - Remarkable power; serious responsibility; not just another job

- **Implications: for ecosystem of media outlets**
  - Control over editorial boards and mastheads
  - Balance and diversity of outlet opinion
  - Effects of fake news
  - Impact on agendas, elections, public policy, discourse
  - Journalism jobs: 25% drop in last decade

- **What should be next?**

- For more information: GaryKing.org/media
Summary and Implications

- **Summary**
  - **Small outlets:** very large average effects on pageviews, agenda (subject & policy), opinion change
  - **Larger outlets:** even bigger average effects
  - **Heterogeneous effects:** large, highly variable viral patterns

- **Implications: for individual journalists**
  - Remarkable power; serious responsibility; not just another job

- **Implications: for ecosystem of media outlets**
  - Control over editorial boards and mastheads

- **What should be next?**
Summary and Implications

• **Summary**
  • Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change
  • Larger outlets: even bigger average effects
  • Heterogeneous effects: large, highly variable viral patterns

• **Implications: for individual journalists**
  • Remarkable power; serious responsibility; not just another job

• **Implications: for ecosystem of media outlets**
  • Control over editorial boards and mastheads
  • Balance and diversity of outlet opinion

• What should be next?
Summary and Implications

• **Summary**
  - **Small outlets:** very large average effects on pageviews, agenda (subject & policy), opinion change
  - **Larger outlets:** even bigger average effects
  - **Heterogeneous effects:** large, highly variable viral patterns

• **Implications: for individual journalists**
  - Remarkable power; serious responsibility; not just another job

• **Implications: for ecosystem of media outlets**
  - Control over editorial boards and mastheads
  - Balance and diversity of outlet opinion
  - Effects of fake news

• What should be next?
Summary and Implications

• Summary
  • Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change
  • Larger outlets: even bigger average effects
  • Heterogeneous effects: large, highly variable viral patterns

• Implications: for individual journalists
  • Remarkable power; serious responsibility; not just another job

• Implications: for ecosystem of media outlets
  • Control over editorial boards and mastheads
  • Balance and diversity of outlet opinion
  • Effects of fake news
  • Impact on agendas, elections, public policy, discourse

• What should be next?
Summary and Implications

• **Summary**
  • Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change
  • Larger outlets: even bigger average effects
  • Heterogeneous effects: large, highly variable viral patterns

• **Implications: for individual journalists**
  • Remarkable power; serious responsibility; not just another job

• **Implications: for ecosystem of media outlets**
  • Control over editorial boards and mastheads
  • Balance and diversity of outlet opinion
  • Effects of fake news
  • Impact on agendas, elections, public policy, discourse
  • Journalism jobs: 25% drop in last decade

• **What should be next?**
Summary and Implications

• **Summary**
  • Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change
  • Larger outlets: even bigger average effects
  • Heterogeneous effects: large, highly variable viral patterns

• **Implications: for individual journalists**
  • Remarkable power; serious responsibility; not just another job

• **Implications: for ecosystem of media outlets**
  • Control over editorial boards and mastheads
  • Balance and diversity of outlet opinion
  • Effects of fake news
  • Impact on agendas, elections, public policy, discourse
  • Journalism jobs: 25% drop in last decade

• **What should be next?**
  • We wrote a paper,
Summary and Implications

• **Summary**
  - Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change
  - Larger outlets: even bigger average effects
  - Heterogeneous effects: large, highly variable viral patterns

• **Implications: for individual journalists**
  - Remarkable power; serious responsibility; not just another job

• **Implications: for ecosystem of media outlets**
  - Control over editorial boards and mastheads
  - Balance and diversity of outlet opinion
  - Effects of fake news
  - Impact on agendas, elections, public policy, discourse
  - Journalism jobs: 25% drop in last decade

• **What should be next?**
  - We wrote a paper, built a platform,
Summary and Implications

• **Summary**
  • **Small outlets:** very large average effects on pageviews, agenda (subject & policy), opinion change
  • **Larger outlets:** even bigger average effects
  • **Heterogeneous effects:** large, highly variable viral patterns

• **Implications: for individual journalists**
  • Remarkable power; serious responsibility; not just another job

• **Implications: for ecosystem of media outlets**
  • Control over editorial boards and mastheads
  • Balance and diversity of outlet opinion
  • Effects of fake news
  • Impact on agendas, elections, public policy, discourse
  • Journalism jobs: 25% drop in last decade

• **What should be next?**
  • We wrote a paper, built a platform, & showed how others can
Summary and Implications

• **Summary**
  - **Small outlets:** very large average effects on pageviews, agenda (subject & policy), opinion change
  - **Larger outlets:** even bigger average effects
  - **Heterogeneous effects:** large, highly variable viral patterns

• **Implications: for individual journalists**
  - Remarkable power; serious responsibility; not just another job

• **Implications: for ecosystem of media outlets**
  - Control over editorial boards and mastheads
  - Balance and diversity of outlet opinion
  - Effects of fake news
  - Impact on agendas, elections, public policy, discourse
  - Journalism jobs: 25% drop in last decade

• **What should be next?**
  - We wrote a **paper**, built a platform, & showed how others can
  - What experiment would you (or should we) run next?

Implications
Summary and Implications

• Summary
  • Small outlets: very large average effects on pageviews, agenda (subject & policy), opinion change
  • Larger outlets: even bigger average effects
  • Heterogeneous effects: large, highly variable viral patterns

• Implications: for individual journalists
  • Remarkable power; serious responsibility; not just another job

• Implications: for ecosystem of media outlets
  • Control over editorial boards and mastheads
  • Balance and diversity of outlet opinion
  • Effects of fake news
  • Impact on agendas, elections, public policy, discourse
  • Journalism jobs: 25% drop in last decade

• What should be next?
  • We wrote a paper, built a platform, & showed how others can
  • What experiment would you (or should we) run next?

For more information: GaryKing.org/media
Notation and Quantities of Interest

• Outcome Variable: $y_{ped}$, # social media posts in policy area $p$ ($p = 1, … , 11$)
• Experiment $e$ ($e = 1, … , E_p$)
• Day of and after intervention ($d = 1, … , 6$)

• Treatment Variable: $T_{ped}$, instruction to pack (of 2-5 outlets) to write, publish, and promote articles, like a project manager

• Treated weeks: $T_{ped1} = \cdots = T_{ped6} = 1$
• Control weeks: $T_{ped1} = \cdots = T_{ped6} = 0$

• Quantities of Interest

• Absolute Increase: $\lambda_d = \frac{\text{mean}_{p,e}[Y_{ped}(1)] - \text{mean}_{p,e}[Y_{ped}(0)]}{23}$

• Proportionate Increase: $\phi_d = \frac{\lambda_d}{\text{mean}_{p,e}[Y_{ped}(0)]}$
Notation and Quantities of Interest

- **Outcome Variable**: $y_{ped}$, # social media posts in

- **Treatment Variable**: $T_{ped}$, instruction to pack (of 2-5 outlets) to write, publish, and promote articles, like a project manager

- **Treated weeks**: $T_{pe1} = \ldots = T_{pe6} = 1$

- **Control weeks**: $T_{pe1} = \ldots = T_{pe6} = 0$

- **Quantities of Interest**
  - Absolute Increase: $\lambda_d = \text{mean}_{p,e}[Y_{ped}(1)] - \text{mean}_{p,e}[Y_{ped}(0)]$
  - Proportionate Increase: $\phi_d = \frac{\lambda_d}{\text{mean}_{p,e}[Y_{ped}(0)]}$
Notation and Quantities of Interest

- **Outcome Variable:** $y_{ped}$, # social media posts in policy area $p$ ($p = 1, \ldots, 11$)
Notation and Quantities of Interest

- **Outcome Variable**: \( y_{ped} \), # social media posts in
  - policy area \( p \) (\( p = 1, \ldots, 11 \))
  - experiment \( e \) (\( e = 1, \ldots, E_p \))
Notation and Quantities of Interest

- **Outcome Variable**: $y_{ped}$, # social media posts in
  - policy area $p$ ($p = 1, \ldots, 11$)
  - experiment $e$ ($e = 1, \ldots, E_p$)
  - day $d$ of and after intervention ($d = 1, \ldots, 6$)
Notation and Quantities of Interest

• **Outcome Variable:** $y_{ped}$, # social media posts in
  - policy area $p$ ($p = 1, \ldots, 11$)
  - experiment $e$ ($e = 1, \ldots, E_p$)
  - day $d$ of and after intervention ($d = 1, \ldots, 6$)

• **Treatment Variable:** $T_{ped}$, instruction to pack (of 2-5 outlets) to write, publish, and promote articles, like a project manager
Notation and Quantities of Interest

- **Outcome Variable:** $y_{ped}$, # social media posts in
  - policy area $p$ ($p = 1, \ldots, 11$)
  - experiment $e$ ($e = 1, \ldots, E_p$)
  - day $d$ of and after intervention ($d = 1, \ldots, 6$)

- **Treatment Variable:** $T_{ped}$, instruction to pack (of 2-5 outlets) to write, publish, and promote articles, like a project manager
  - Treated weeks: $T_{pe1} = \cdots = T_{pe6} = 1$
Notation and Quantities of Interest

- **Outcome Variable**: $y_{ped}$, # social media posts in
  - policy area $p$ ($p = 1, \ldots, 11$)
  - experiment $e$ ($e = 1, \ldots, E_p$)
  - day $d$ of and after intervention ($d = 1, \ldots, 6$)

- **Treatment Variable**: $T_{ped}$, instruction to pack (of 2-5 outlets) to write, publish, and promote articles, like a project manager
  - Treated weeks: $T_{pe1} = \cdots = T_{pe6} = 1$
  - Control weeks: $T_{pe1} = \cdots = T_{pe6} = 0$
Notation and Quantities of Interest

- **Outcome Variable**: $y_{ped}$, # social media posts in
  - policy area $p$ ($p = 1, \ldots, 11$)
  - experiment $e$ ($e = 1, \ldots, E_p$)
  - day $d$ of and after intervention ($d = 1, \ldots, 6$)

- **Treatment Variable**: $T_{ped}$, instruction to pack (of 2-5 outlets) to write, publish, and promote articles, like a project manager
  - Treated weeks: $T_{pe1} = \cdots = T_{pe6} = 1$
  - Control weeks: $T_{pe1} = \cdots = T_{pe6} = 0$

- **Quantities of Interest**
Notation and Quantities of Interest

- **Outcome Variable:** \( y_{ped} \), # social media posts in
  - policy area \( p \) (\( p = 1, \ldots, 11 \))
  - experiment \( e \) (\( e = 1, \ldots, E_p \))
  - day \( d \) of and after intervention (\( d = 1, \ldots, 6 \))

- **Treatment Variable:** \( T_{ped} \), instruction to pack (of 2-5 outlets) to write, publish, and promote articles, like a project manager
  - Treated weeks: \( T_{pe1} = \ldots = T_{pe6} = 1 \)
  - Control weeks: \( T_{pe1} = \ldots = T_{pe6} = 0 \)

- **Quantities of Interest**
  - Absolute Increase: \( \lambda_d = \text{mean}_{p,e}[Y_{ped}(1)] - \text{mean}_{p,e}[Y_{ped}(0)] \)
Notation and Quantities of Interest

• Outcome Variable: \( y_{ped} \), # social media posts in
  - policy area \( p \) (\( p = 1, \ldots, 11 \))
  - experiment \( e \) (\( e = 1, \ldots, E_p \))
  - day \( d \) of and after intervention (\( d = 1, \ldots, 6 \))

• Treatment Variable: \( T_{ped} \), instruction to pack (of 2-5 outlets) to write, publish, and promote articles, like a project manager
  - Treated weeks: \( T_{pe1} = \cdots = T_{pe6} = 1 \)
  - Control weeks: \( T_{pe1} = \cdots = T_{pe6} = 0 \)

• Quantities of Interest
  - Absolute Increase: \( \lambda_d = \text{mean}_{p,e}[Y_{ped}(1)] - \text{mean}_{p,e}[Y_{ped}(0)] \)
  - Proportionate Increase: \( \phi_d = \frac{\lambda_d}{\text{mean}_{p,e}[Y_{ped}(0)]} \)
Estimation Approaches

Model-Based Approach

Transform outcome variable for normality & homoskedasticity:
\[ z_{ped} = \ln(y_{ped} + 0.5) \]

The Model:
\[ E(z_{ped} | T_{ped}) = \beta_0 + \beta_p T_{ped} + \eta_d + \gamma_d T_{ped} \]

- \[ \beta_0 \]: constant term
- \[ \beta_p \]: fixed effects for the 11 policy areas

Assume linearity over days:
\[ \eta_d = \eta_0 + \eta_1 d \]
\[ \gamma_d = \gamma_0 + \gamma_1 d \]

Assume conditional independence over \( p, e, d \)

Model-Free Approach:

Drop linearity & conditional independence assumptions

Regress \( z_{ped} \) on \( T_{ped} \) separately for each \( d \)

Equivalent to difference in means for each day

(potentially with policy fixed effects)
Estimation Approaches

• Model-Based Approach

\[ z_{ped} = \ln (y_{ped} + 0.5) \]

The Model:

\[ E(z_{ped} | T_{ped}) = \beta_0 + \beta_p T_{ped} + \eta_d + \gamma_d T_{ped} \]

- \( \beta_0 \): constant term
- \( \beta_p \): fixed effects for the 11 policy areas
- Assume linearity over days:
  \[ \eta_d = \eta_0 + \eta_1 d \] and \[ \gamma_d = \gamma_0 + \gamma_1 d \]

- Assume conditional independence over \( p, e, d \)

Model-Free Approach:

• Drop linearity & conditional independence assumptions
• Regress \( z_{ped} \) on \( T_{ped} \) separately for each \( d \)
• Equivalent to difference in means for each day (perhaps with policy fixed effects)
Estimation Approaches

• **Model-Based Approach**
  • Transform outcome variable for normality & homoskedasticity:
    \[ z_{\text{ped}} = \ln(y_{\text{ped}} + 0.5) \]
Estimation Approaches

• **Model-Based Approach**
  • Transform outcome variable for normality & homoskedasticity:
    \[ z_{ped} = \ln(y_{ped} + 0.5) \]
  • The Model: \( E(z_{ped}|T_{ped}) = \beta^0 + \beta_p + \eta_d + \gamma_d T_{ped} \)
Estimation Approaches

• Model-Based Approach
  • Transform outcome variable for normality & homoskedasticity:
    \[ z_{\text{ped}} = \ln(y_{\text{ped}} + 0.5) \]
  • The Model: \( E(z_{\text{ped}}|T_{\text{ped}}) = \beta_0^0 + \beta_p + \eta_d + \gamma_d T_{\text{ped}} \)
    • \( \beta_0^0 \): constant term
Estimation Approaches

- **Model-Based Approach**
  - Transform outcome variable for normality & homoskedasticity:
    \[ z_{ped} = \ln(y_{ped} + 0.5) \]
  - The Model: \( E(z_{ped}|T_{ped}) = \beta^0 + \beta_p + \eta_d + \gamma_d T_{ped} \)
    - \( \beta^0 \): constant term
    - \( \beta_p \): fixed effects for the 11 policy areas

- **Model-Free Approach**
  - Drop linearity & conditional independence assumptions
  - Regress \( z_{ped} \) on \( T_{ped} \) separately for each \( d \)
    - Equivalent to difference in means for each day
      (perhaps with policy fixed effects)
Estimation Approaches

• Model-Based Approach
  • Transform outcome variable for normality & homoskedasticity:
    \( z_{ped} = \ln(y_{ped} + 0.5) \)
  • The Model:
    \[ E(z_{ped}|T_{ped}) = \beta^0 + \beta_p + \eta_d + \gamma_d T_{ped} \]
    • \( \beta^0 \): constant term
    • \( \beta_p \): fixed effects for the 11 policy areas
    • Assume linearity over days: \( \eta_d = \eta^0 + \eta^1 d \) and \( \gamma_d = \gamma^0 + \gamma^1 d \)
Estimation Approaches

- **Model-Based Approach**
  - Transform outcome variable for normality & homoskedasticity:
    \[ z_{ped} = \ln(y_{ped} + 0.5) \]
  - The Model:
    \[ E(z_{ped}|T_{ped}) = \beta^0 + \beta_p + \eta_d + \gamma_d T_{ped} \]
    - \( \beta^0 \): constant term
    - \( \beta_p \): fixed effects for the 11 policy areas
    - Assume linearity over days: \( \eta_d = \eta^0 + \eta^1 d \) and \( \gamma_d = \gamma^0 + \gamma^1 d \)
    - Assume *conditional* independence over \( p, e, d \)

- **Model-Free Approach**
  - Drop linearity & conditional independence assumptions
  - Regress \( z_{ped} \) on \( T_{ped} \) separately for each \( d \)
    - Equivalent to difference in means for each day
    - (perhaps with policy fixed effects)
Estimation Approaches

• **Model-Based Approach**
  - Transform outcome variable for normality & homoskedasticity:
    \[ z_{ped} = \ln(y_{ped} + 0.5) \]
  - The Model: \( E(z_{ped}|T_{ped}) = \beta^0 + \beta_p + \eta_d + \gamma_d T_{ped} \)
    - \( \beta^0 \): constant term
    - \( \beta_p \): fixed effects for the 11 policy areas
    - Assume linearity over days: \( \eta_d = \eta^0 + \eta^1 d \) and \( \gamma_d = \gamma^0 + \gamma^1 d \)
    - Assume *conditional* independence over \( p, e, d \)

• **Model-Free Approach:**
Estimation Approaches

• Model-Based Approach
  • Transform outcome variable for normality & homoskedasticity:
    \[ z_{ped} = \ln(y_{ped} + 0.5) \]
  • The Model: \[ E(z_{ped}|T_{ped}) = \beta^0 + \beta_p \eta_d + \gamma_d T_{ped} \]
    • \( \beta^0 \): constant term
    • \( \beta_p \): fixed effects for the 11 policy areas
    • Assume linearity over days: \( \eta_d = \eta^0 + \eta^1 d \) and \( \gamma_d = \gamma^0 + \gamma^1 d \)
    • Assume conditional independence over \( p, e, d \)

• Model-Free Approach:
  • Drop linearity & conditional independence assumptions
Estimation Approaches

- **Model-Based Approach**
  - Transform outcome variable for normality & homoskedasticity:
    \[ z_{ped} = \ln(y_{ped} + 0.5) \]
  - The Model: \( E(z_{ped}|T_{ped}) = \beta^0 + \beta_p + \eta_d + \gamma_d T_{ped} \)
    - \( \beta^0 \): constant term
    - \( \beta_p \): fixed effects for the 11 policy areas
    - Assume linearity over days: \( \eta_d = \eta^0 + \eta^1 d \) and \( \gamma_d = \gamma^0 + \gamma^1 d \)
    - Assume conditional independence over \( p, e, d \)

- **Model-Free Approach:**
  - Drop linearity & conditional independence assumptions
  - Regress \( z_{ped} \) on \( T_{ped} \) separately for each \( d \)
Estimation Approaches

- **Model-Based Approach**
  - Transform outcome variable for normality & homoskedasticity:
    \[
    z_{ped} = \ln(y_{ped} + 0.5)
    \]
  - The Model: \( E(z_{ped}|T_{ped}) = \beta^0 + \beta_p + \eta_d + \gamma_d T_{ped} \)
    - \( \beta^0 \): constant term
    - \( \beta_p \): fixed effects for the 11 policy areas
    - Assume linearity over days: \( \eta_d = \eta^0 + \eta^1 d \) and \( \gamma_d = \gamma^0 + \gamma^1 d \)
    - Assume *conditional* independence over \( p, e, d \)

- **Model-Free Approach:**
  - Drop linearity & conditional independence assumptions
  - Regress \( z_{ped} \) on \( T_{ped} \) separately for each \( d \)
  - Equivalent to difference in means for each day
Estimation Approaches

• **Model-Based Approach**
  - Transform outcome variable for normality & homoskedasticity:
    \[ z_{ped} = \ln(y_{ped} + 0.5) \]
  - The Model: \[ E(z_{ped}|T_{ped}) = \beta^0 + \beta_p + \eta_d + \gamma_d T_{ped} \]
    - \( \beta^0 \): constant term
    - \( \beta_p \): fixed effects for the 11 policy areas
    - Assume linearity over days: \( \eta_d = \eta^0 + \eta^1 d \) and \( \gamma_d = \gamma^0 + \gamma^1 d \)
    - Assume *conditional* independence over \( p, e, d \)

• **Model-Free Approach:**
  - Drop linearity & conditional independence assumptions
  - Regress \( z_{ped} \) on \( T_{ped} \) separately for each \( d \)
  - Equivalent to difference in means for each day
  - (perhaps with policy fixed effects)