Session II
Examples and Paradigms

Thomas J. Leeper

Government Department
London School of Economics and Political Science

Analysis

See website for

- Data from yesterday's activity
- R code for basic experimental analysis
- Stata code for basic experimental analysis

Aside: Complex Designs

- An experiment can have any number of conditions
  - Up to the limits of sample size
  - More than 8–10 conditions is typically unwieldy
- Still analyze complex designs using regression, but focus on pairwise comparisons to estimates SATEs
  - Treatment–treatment, or treatment-control
  - Without control group, we don't know which treatment(s) affected the outcome

What makes a good research question?

KKV's two criteria:

1. Politically important
2. Contribute to scientific literature

What makes a good research question?

KKV's two criteria:

1. Politically important
2. Contribute to scientific literature
3. Personally interesting
4. Unresolved

What makes a good research question?

KKV's two criteria:

1. Politically important
2. Contribute to scientific literature
3. Personally interesting
4. Unresolved
5. For experiments, *forward* in nature

What kinds of questions can we answer with (survey) experiments?

What kinds of questions can we answer with (survey) experiments?

- Forward causal questions
  - Can X cause Y?
  - What effects does X have?

What kinds of questions can we answer with (survey) experiments?

- Forward causal questions
  - Can X cause Y?
  - What effects does X have?
- Backward causal questions
  - What causes Y?
  - How much of Y is attributable to X?

What kinds of questions can we answer with (survey) experiments?

- Forward causal questions
  - Can X cause Y?
  - What effects does X have?
- Backward causal questions
  - What causes Y?
  - How much of Y is attributable to X?
- Even though answering "forward" causal question, we start with an outcome concept

Hypothesis Testing

- From theory, we derive testable hypotheses
  - Hypotheses are expectations about differences in outcomes across levels of a putatively causal variable
  - Hypothesis must be testable by an SATE
- Manipulations are developed to create variation in that causal variable

Example: News Framing

- Theory: Presentation of news affects opinion
- Hypotheses:
  - News emphasizing free speech increases support for a hate group rally
  - News emphasizing public safety decreases support for a hate group rally
- Manipulation:
  - Control group: no information
  - Free speech group: article emphasizing rights
  - Public safety group: article emphasizing safety

Example: Partisan Identity

- Theory: Strength of partisan identity affects tendency to accept party position
- Hypotheses:
  - Strong partisans are more likely to accept their party's position on an issue
- Manipulation:
  - Control group: no manipulation
  - "Univalent" condition
  - "Ambivalent" condition

Univalent

These days, Democrats and Republicans differ from one another considerably. The two groups seem to be growing further and further apart, not only in terms of their opinions but also their lifestyles. Earlier in the survey, you said you tend to identify as a *Democrat/ Republican*. Please take a few minutes to think about what you like about *Democrats/ Republicans* compared to the *Republicans/ Democrats*. Think of 2 to 3 things you especially like best about **your party**. Then think of 2 to 3 things you especially dislike about **the other party**. Now please write those thoughts in the space below.

Ambivalent

These days, Democrats and Republicans differ from one another considerably. The two groups seem to be growing further and further apart, not only in terms of their opinions but also their lifestyles. Earlier in the survey, you said you tend to identify as a *Democrat/ Republican*. Please take a few minutes to think about what you like about *Democrats/ Republicans* compared to the *Republicans/ Democrats*. Think of 2 to 3 things you especially like best about **the other party**. Then think of 2 to 3 things you especially dislike about **your party**. Now please write those thoughts in the space below.

Hypothesis Testing

- Derive experimental design from hypotheses
- Experimental "factors" are expressions of hypotheses as randomized groups
- What intervention each group receives depends on hypotheses
  - presence/absence
  - levels/doses
  - qualitative variations

Ex.: Presence/Absence

- Theory: Negative campaigning reduces support for the party described negatively.
- Hypothesis: Exposure to a negative advertisement criticizing a party reduces support for that party.
- Manipulation:
  - Control group receives no advertisement.
  - Treatment group watches a video containing a negative ad describing a party.

Ex.: Levels/doses

- Theory: Negative campaigning reduces support for the party described negatively.
- Hypothesis: Exposure to higher levels of negative advertising criticizing a party reduces support for that party.
- Manipulation:
  - Control group receives no advertisement.
  - Treatment group 1 watches a video containing 1 negative ad describing a party.
  - Treatment group 2 watches a video containing 2 negative ads describing a party.
  - Treatment group 3 watches a video containing 3 negative ads describing a party.

Ex.: Qualitative variation

- Theory: Negative campaigning reduces support for the party described negatively.
- Hypothesis: Exposure to a negative advertisement criticizing a party reduces support for that party, while a positive advertisement has no effect.
- Manipulation:
  - Control group receives no advertisement.
  - Negative treatment group watches a video containing a negative ad describing a party.
  - Positive treatment group watches a video containing a positive ad describing a party.

Questions?

Share your TESS Examples

In groups of 4–5, share your TESS examples

- What was the researcher's question?
- How did they test it experimentally?
- What was interesting or surprising about the designs?

Take about 7–8 minutes.

Protocol

Protocol is the complete planning document for how to design, implement, and analyze an experiment.[1]

[1] Thomas J. Leeper. 2011. "The Use of Protocol in the Design and Reporting of Experiments." *The Experimental Political Scientist*.

Protocol

1. Theory/hypotheses
2. Instrumentation
    - Manipulation(s)
    - Outcome(s)
    - Covariate(s)
    - Manipulation check(s)
3. Sampling
4. Implementation
5. Analysis

Why bother?

Why bother?

- Be clear to yourself what you're trying to do before you do it

Why bother?

- Be clear to yourself what you're trying to do before you do it
- Assess the literature for best practices

Why bother?

- Be clear to yourself what you're trying to do before you do it
- Assess the literature for best practices
- Highlight areas in need of pilot testing

Why bother?

- Be clear to yourself what you're trying to do before you do it
- Assess the literature for best practices
- Highlight areas in need of pilot testing
- Economize questionnaire development

Why bother?

- Be clear to yourself what you're trying to do before you do it
- Assess the literature for best practices
- Highlight areas in need of pilot testing
- Economize questionnaire development
- Study preregistration

Questions?

Activity!

- How do we know if an experiment is any good?
- Talk with a partner for about 3 minutes
- Try to develop some criteria that allow you to evaluate "what makes for a good experiment?"

Some possible criteria

- Significant results
- Face validity
- Coherent for respondents
- Non-obvious to respondents
- Simple
- Indirect/unobtrusive
- Validated by prior work
- Innovative/creative
- ...

The best criterion for evaluating the quality of an experiment is whether it manipulated the intended independent variable and controlled everything else by design.

The best criterion for evaluating the quality of an experiment is whether it manipulated the intended independent variable and controlled everything else by design.

–Thomas J. Leeper (5 July 2016)

How do we know we manipulated what we think we manipulated?

How do we know we manipulated what we think we manipulated?

- Outcomes are affected consistent with theory

How do we know we manipulated what we think we manipulated?

- Outcomes are affected consistent with theory
- Before the study using *pilot testing* (or *pretesting*)

How do we know we manipulated what we think we manipulated?

- Outcomes are affected consistent with theory
- Before the study using *pilot testing* (or *pretesting*)
- During the study, using *manipulation checks*

How do we know we manipulated what we think we manipulated?

- Outcomes are affected consistent with theory
- Before the study using *pilot testing* (or *pretesting*)
- During the study, using *manipulation checks*
- During the study, using *placebos*

How do we know we manipulated what we think we manipulated?

- Outcomes are affected consistent with theory
- Before the study using *pilot testing* (or *pretesting*)
- During the study, using *manipulation checks*
- During the study, using *placebos*
- During the study, using *non-equivalent outcomes*

I. Outcomes Affected

- Follows a circular logic!
- Doesn't tell us anything if we hypothesize null effects

II. Pilot Testing

- Goal: establish construct validity of manipulation
- Assess whether a set of possible manipulations affect a measure of the *independent* variable

II. Pilot Testing

- Goal: establish construct validity of manipulation
- Assess whether a set of possible manipulations affect a measure of the *independent* variable
- Example:
  - Goal: Manipulate the "strength" of an argument
  - Write several arguments
  - Ask pilot test respondents to report how strong each one was

III. Manipulation Checks

- Manipulation checks are items added post-treatment, post-outcome that assess whether the *independent* variable was affected by treatment
- We typically talk about manipulations as directly setting the value of X, but in practice we are typically manipulating something *that we think* strongly modifies X

III. Manipulation Checks

- Manipulation checks are items added post-treatment, post-outcome that assess whether the *independent* variable was affected by treatment
- We typically talk about manipulations as directly setting the value of X, but in practice we are typically manipulating something *that we think* strongly modifies X
- Example: information manipulations aim to modify knowledge or beliefs, but are necessarily imperfect at doing so

Manipulation check example[2]

1. Treatment 1: Supply Information
2. Manipulation check 1: measure beliefs
3. Treatment 2: Prime a set of considerations
4. Outcome: Measure opinion
5. Manipulation check 2: measure dimension salience

[2] Leeper & Slothuus. n.d. "Can Citizens Be Framed?" Available from: `http://thomasleeper.com/research.html`.

Some Best Practices

Some Best Practices

- Manipulation checks should be innocuous
  - Shouldn't modify independent variable
  - Shouldn't modify outcome variable

Some Best Practices

- Manipulation checks should be innocuous
  - Shouldn't modify independent variable
  - Shouldn't modify outcome variable
- Generally, measure post-outcome

Some Best Practices

- Manipulation checks should be innocuous
  - Shouldn't modify independent variable
  - Shouldn't modify outcome variable
- Generally, measure post-outcome
- Measure both what you wanted to manipulate *and* what you didn't want to manipulate
  - Most treatments are *compound*!

IV. Placebos

- Include an experimental condition that *does not* manipulate the variable of interest (but might affect the outcome)

IV. Placebos

- Include an experimental condition that *does not* manipulate the variable of interest (but might affect the outcome)
- Example:
  - Study whether risk-related arguments about climate change increase support for a climate change policy
  - Placebo condition: control article with risk-related arguments about non-environmental issue (e.g., terrorism)

V. Non-equivalent outcomes

- Measures an outcome that *should not* be affected by independent variable

V. Non-equivalent outcomes

- Measures an outcome that *should not* be affected by independent variable
- Example:
  - Assess effect of some treatment on attitudes toward group A
  - Focal outcome: attitudes toward group A
  - Non-equivalent outcome: attitudes toward group B

Aside: Demand Characteristics

- "Demand characteristics" are features of experiments that (unintentionally) imply the purpose of the study and thereby change respondents' behavior (to be consistent with theory)

[3] But, consider the ethics of not doing so (more Friday)

Aside: Demand Characteristics

- "Demand characteristics" are features of experiments that (unintentionally) imply the purpose of the study and thereby change respondents' behavior (to be consistent with theory)
- Implications:
  - Design experimental treatments that are non-obvious
  - Do not disclose the purpose of the study up front[3]

[3] But, consider the ethics of not doing so (more Friday)

Question Wording Designs

- Kahneman and Tversky is a "question wording" experiment
- Hypothesized difference in outcomes according to the decision being faced
  - Risky or not risky
  - Gains or losses
- Manipulation operationalizes this by asking two different questions
- Outcome is the answer to the question

Activity!

- With a partner, identify each of the protocol elements of the experiment described by Kahneman and Tversky
  1. Theory/hypotheses
  2. Instrumentation
  3. Sampling
  4. Implementation
  5. Analysis
- Report back in about 4 minutes

"Framing" or "Priming" Experiments

Example: Schuldt et al. "'Global Warming' or 'Climate Change'? Whether the Planet is Warming Depends on Question Wording."

What's this study about?

You may have heard about the idea that the world's temperature may have been **going up** over the past 100 years, a phenomenon sometimes called **global warming**. What is your personal opinion regarding whether or not this has been happening?

- Definitely has not been happening
- Probably has not been happening
- Unsure, but leaning toward it has not been happening
- Not sure either way
- Unsure, but leaning toward it has been happening
- Probably has been happening
- Definitely has been happening

You may have heard about the idea that the world's temperature may have been **changing** over the past 100 years, a phenomenon sometimes called **climate change**. What is your personal opinion regarding whether or not this has been happening?

- Definitely has not been happening
- Probably has not been happening
- Unsure, but leaning toward it has not been happening
- Not sure either way
- Unsure, but leaning toward it has been happening
- Probably has been happening
- Definitely has been happening

Another framing example[4]

Today, tests are being developed that make it possible to detect serious genetic defects **before a baby is born**. But so far, it is impossible either to treat or to correct most of them. If (you/your partner) were pregnant, would you want (her) to have a test to find out if the **baby** has any serious genetic defects? (Yes/No)

Suppose a test shows the **baby** has a serious genetic defect. Would you, yourself, want (your partner) to have an abortion if a test shows the **baby** has a serious genetic defect? (Yes/No)

---

[4]Singer & Couper. 2014. "The Effect of Question Wording on Attitudes toward Prenatal Testing and Abortion." *Public Opinion Quarterly* 78(3): 751–760.

Another framing example[4]

Today, tests are being developed that make it possible to detect serious genetic defects **in the fetus during pregnancy**. But so far, it is impossible either to treat or to correct most of them. If (you/your partner) were pregnant, would you want (her) to have a test to find out if the **fetus** has any serious genetic defects? (Yes/No)

Suppose a test shows the **fetus** has a serious genetic defect. Would you, yourself, want (your partner) to have an abortion if a test shows the **fetus** has a serious genetic defect? (Yes/No)

---

[4] Singer & Couper. 2014. "The Effect of Question Wording on Attitudes toward Prenatal Testing and Abortion." *Public Opinion Quarterly* 78(3): 751–760.

Another framing example[5]

Do you favor or oppose the death penalty for persons convicted of murder?

[5] Bobo & Johnson. 2004. "A Taste for Punishment: Black and White Americans' Views on the Death Penalty and the War on Drugs." Du Bois Review 1(1): 151–180.

Another framing example[5]

Blacks are about 12% of the U.S. population, but they were half of the homicide offenders last year. Do you favor or oppose the death penalty for persons convicted of murder?

[5] Bobo & Johnson. 2004. "A Taste for Punishment: Black and White Americans' Views on the Death Penalty and the War on Drugs." Du Bois Review 1(1): 151–180.

Another framing example[6]

Concealed handgun laws have recently received national attention. Some people have argued that law-abiding citizens have the right to protect themselves. What do you think about concealed handgun laws?

---

[6]Haider-Markel & Joslyn. 2001. "Gun Policy, Opinion, Tragedy, and Blame Attribution: The Conditional Influence of Issue Frames." *Journal of Politics* 63(2): 520–543.

Another framing example[6]

Concealed handgun laws have recently received national attention. Some people have argued that laws allowing citizens to carry concealed handguns threaten public safety because they would allow almost anyone to carry a gun almost anywhere, even onto school grounds. What do you think about concealed handgun laws?

[6] Haider-Markel & Joslyn. 2001. "Gun Policy, Opinion, Tragedy, and Blame Attribution: The Conditional Influence of Issue Frames." *Journal of Politics* 63(2): 520–543.

Question testing

Use question wording designs to select which survey measures we want to use

- Select possible question wordings
- Select some criterion(-ia) for assessing which is better
- Pilot test and then use the item that performs better

Aside: Experimentation vs. Other Pretesting Methods

Aside: Experimentation vs. Other Pretesting Methods

- Experiments are complementary to other pretesting methods

Aside: Experimentation vs. Other Pretesting Methods

- Experiments are complementary to other pretesting methods
- Specific value added of an experiment: optimize questions or other survey features against a specific criterion, e.g.:
  - (Non-)Response or drop-off rates
  - "Don't know" rates
  - Item characteristics
  - Reading times or response latencies

Aside: Experimentation vs. Other Pretesting Methods

- Experiments are complementary to other pretesting methods
- Specific value added of an experiment: optimize questions or other survey features against a specific criterion, e.g.:
  - (Non-)Response or drop-off rates
  - "Don't know" rates
  - Item characteristics
  - Reading times or response latencies
- But! Power considerations...

Classic question testing experiment[7]

Some people feel that The 1975 Public Affairs Act should be repealed-do you agree or disagree with this idea?

[7] Bishop, G.F., Tuchfarber, A. & Oldendick, R.W. 1986. "Opinions on Fictitious Issues: The Pressure to Answer Survey Questions." *Public Opinion Quarterly* 50(2): 240–250.

Classic question testing experiment[7]

Some people feel that The 1975 Public Affairs Act should be repealed-do you agree or disagree with this idea, or haven't you thought much about this issue?

[7] Bishop, G.F., Tuchfarber, A. & Oldendick, R.W. 1986. "Opinions on Fictitious Issues: The Pressure to Answer Survey Questions." *Public Opinion Quarterly* 50(2): 240–250.

An example[8]

In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. How about you–did you vote in the elections this November?

[8] Holbrook & Krosnick. 2013. "A New Question Sequence to Measure Voter Turnout in Telephone Surveys: Results of an Experiment in the 2006 ANES Pilot Study." *Public Opinion Quarterly* 77: 106–123.

An example[8]

In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time. Which of the following statements best describes you?

- One, I did not vote in the November 3 election
- two, I thought about voting this time but didn't
- three, I usually vote but didn't this time
- four, I am sure I voted

[8] Holbrook & Krosnick. 2013. "A New Question Sequence to Measure Voter Turnout in Telephone Surveys: Results of an Experiment in the 2006 ANES Pilot Study." *Public Opinion Quarterly* 77: 106–123.

An Instructional Manipulation[9]

For the next few questions, I am going to read out some statements, and for each one, please tell me if it is true or false. If you don't know, just say so and we will skip to the next one.

1. Britain's electoral system is based on proportional representation.
2. MPs from different parties are on parliamentary committees.
3. The Conservatives are opposed to the ratification of a constitution for the European Union.

---

[9] Sturgis, Allum & Smith. 2008. "An Experiment on the Measurement of Political Knowledge in Surveys." *Public Opinion Quarterly* 72(1): 90–102.

An Instructional Manipulation[9]

For the next few questions, I am going to read out some statements, and for each one, please tell me if it is true or false. If you don't know, please just give me your best guess.

1. Britain's electoral system is based on proportional representation.
2. MPs from different parties are on parliamentary committees.
3. The Conservatives are opposed to the ratification of a constitution for the European Union.

---

[9] Sturgis, Allum & Smith. 2008. "An Experiment on the Measurement of Political Knowledge in Surveys." *Public Opinion Quarterly* 72(1): 90–102.

An Instructional Manipulation + [10]

In the next part of this study, you will be asked 14 questions about politics, public policy, and economics. Many people don't know the answers to these questions, but it is helpful for us if you answer, even if you're not sure what the correct answer is. We encourage you to take a guess on every question. At the end of this study, you will see a summary of how many questions you answered correctly.

---

[10] Prior & Lupia. 2008. "Money, Time, and Political Knowledge: Distinguishing Quick Recall and Political Learning Skills." *American journal of Political Science* 52(1): 169–183.

An Instructional Manipulation +[10]

We will pay you for answering questions correctly. You will earn $1 for every correct answer you give. So, if you answer 3 of the 14 questions correctly, you will earn $3. If you answer 7 of the 14 questions correctly, you will earn $7. The more questions you answer correctly, the more you will earn.

---

[10] Prior & Lupia. 2008. "Money, Time, and Political Knowledge: Distinguishing Quick Recall and Political Learning Skills." *American journal of Political Science* 52(1): 169–183.

Question Order Designs

- Manipulation of pre-outcome questionnaire

Question Order Designs

- Manipulation of pre-outcome questionnaire
- Example:
  - Goal: assess influence of value salience on support for a policy
  - Manipulate by asking different questions:
    - Battery of 5 "rights" questions, or
    - Battery of 5 "life" questions
  - Measure support for legalized abortion

Question Order Designs

- Manipulation of pre-outcome questionnaire
- Example:
  - Goal: assess influence of value salience on support for a policy
  - Manipulate by asking different questions:
    - Battery of 5 "rights" questions, or
    - Battery of 5 "life" questions
  - Measure support for legalized abortion
- If answers to manipulated questions matter, can measure rest post-outcome

Ex. Question-as-treatment[11]

- How close do you feel to your ethnic or racial group?
- Some people have said that taxes need to be raised to take care of pressing national needs. How willing would you be to have your taxes raised to improve education in public schools?

---

[11] Transue. 2007. "Identity Salience, Identity Acceptance, and Racial Policy Attitudes: American National Identity as a Uniting Force." *American Journal of Political Science* 51(1): 78–91.

Ex. Question-as-treatment[11]

- How close do you feel to other Americans?
- Some people have said that taxes need to be raised to take care of pressing national needs. How willing would you be to have your taxes raised to improve education in public schools?

---

[11] Transue. 2007. "Identity Salience, Identity Acceptance, and Racial Policy Attitudes: American National Identity as a Uniting Force." *American Journal of Political Science* 51(1): 78–91.

Ex. Question-as-treatment[11]

- How close do you feel to your ethnic or racial group?
- Some people have said that taxes need to be raised to take care of pressing national needs. How willing would you be to have your taxes raised to improve educational opportunities for minorities?

---

[11] Transue. 2007. "Identity Salience, Identity Acceptance, and Racial Policy Attitudes: American National Identity as a Uniting Force." *American Journal of Political Science* 51(1): 78–91.

Ex. Question-as-treatment[11]

- How close do you feel to other Americans?
- Some people have said that taxes need to be raised to take care of pressing national needs. How willing would you be to have your taxes raised to improve educational opportunities for minorities?

---

[11] Transue. 2007. "Identity Salience, Identity Acceptance, and Racial Policy Attitudes: American National Identity as a Uniting Force." *American Journal of Political Science* 51(1): 78–91.

Ex.: Knowledge and Political Interest

1. Do you happen to remember anything special that your U.S. Representative has done for your district or for the people in your district while he has been in Congress?
2. Is there any legislative bill that has come up in the House of Representatives, on which you remember how your congressman has voted in the last couple of years?
3. Now, some people seem to follow what's going on in government and public affairs most of the time, whether there's an election going on or not. Others aren't that interested. Would you say that you follow what's going on in government and public affairs most of the time, some of the time, only now and then, or hardly at all?

Ex.: Knowledge and Political Interest

1. Now, some people seem to follow what's going on in government and public affairs most of the time, whether there's an election going on or not. Others aren't that interested. Would you say that you follow what's going on in government and public affairs most of the time, some of the time, only now and then, or hardly at all?
2. Do you happen to remember anything special that your U.S. Representative has done for your district or for the people in your district while he has been in Congress?
3. Is there any legislative bill that has come up in the House of Representatives, on which you remember how your congressman has voted in the last couple of years?

Vignettes

- A "vignette" is a short paragraph of text describing a situation
- Vignettes are probably the most common survey experimental paradigm, after question wording designs
- Take many forms and increasingly encompass non-textual stimuli
- Basically limited to web-based mode

A classic vignette[12]

Now think about a **(black/white)** woman in her early thirties. She is a high school **(graduate/drop out)** with a ten-year-old child, and she has been on welfare for the past year.

- How likely is it that she will have more children in order to get a bigger welfare check? (1 = Very likely, . . . , 7 = Not at all likely)
- How likely do you think it is that she will really try hard to find a job in the next year? (1 = Very likely, . . . , 7 = Not at all likely)

---

[12] Gilens, M. 1996. "'Race coding' and white opposition to welfare. *American Political Science Review* 90(3): 593–604.

Newer vignette[13]

Imagine that you were living in a village in another district in Uttar Pradesh and that you were voting for candidates in **(village/state/national)** election. Here are the two candidates who are running against each other: The first candidate is named **(caste name)** and is running as the **(BJP/SP/BSP)** party candidate. **(Corrupt/criminality allegation)**. His opponent is named **(caste name)** and is running as the **(BJP/SP/BSP)** party candidate. **(Opposite corrupt/criminality allegation)**. From this information, please indicate which candidate you would vote for in the **(village/state/national)** election.

[13] Banerjee et al. 2012. "Are Poor Voters Indifferent to Whether Elected Leaders are Criminal or Corrupt? A Vignette Experiment in Rural India." Working paper.

Longer texts[14]

We are testing materials for use in a study **of the structure of sentences people use when writing news editorials**. Along these lines, we would like you to read a series of paragraphs, taken from recent major newspaper editorials.

[14] Druckman & Leeper. 2012. "Learning More from Political Communication Experiments: Pretreatment and Its Effects." *American Journal of Political Science* 56(4): 875–896.

Longer texts[14]

We are testing materials for use in a study **that is related to the kinds of opinions people form about public policies**. Along these lines, we would like you to read a series of paragraphs, taken from recent major newspaper editorials.

[14] Druckman & Leeper. 2012. "Learning More from Political Communication Experiments: Pretreatment and Its Effects." *American Journal of Political Science* 56(4): 875–896.

Please read the following paragraphs and, for each, rate **how *dynamic* you think it is. A paragraph is more "dynamic" when it uses more vivid action words. For example, a statement like, "He sped up and raced through the light before crashing into the swerving truck," seems more dynamic than, "He went faster to get through the light before having an accident." The action words in the first sentence (which we have highlighted in bold) seem more dynamic or vivid than those contained in the second sentence.** There are no right or wrong opinions and your responses to all questions are completely confidential.

Please read the following paragraphs and, for each, rate **the extent to which it decreases or increases your support for the Patriot Act. In subsequent surveys we will ask you for your overall opinion about the state-run casino (i.e., the extent to which you oppose or support the state-run casino).** There are no right or wrong opinions and your responses to all questions are completely confidential.

Please read the paragraphs carefully and, after each one, rate **the extent to which you think it is *dynamic***.

With the passage of the Patriot Act in 2001, the FBI can now enter your home, search around, and doesn't ever have to tell you it was there. You could be perfectly innocent, yet federal agents can go through your most personal effects. When considering new laws, a test of the impact on liberty should be required. On that test, the Patriot Act fails. At a massive 342 pages, it potentially violates at least six of the ten original amendments known as the Bill of Rights — the First, Fourth, Fifth, Sixth, Seventh and Eighth Amendments — and possibly the Thirteenth and Fourteenth as well.

Please read the paragraphs carefully and, after each one, rate **the extent to which it decreases or increases your support for the Patriot Act**.

With the passage of the Patriot Act in 2001, the FBI can now enter your home, search around, and doesn't ever have to tell you it was there. You could be perfectly innocent, yet federal agents can go through your most personal effects. When considering new laws, a test of the impact on liberty should be required. On that test, the Patriot Act fails. At a massive 342 pages, it potentially violates at least six of the ten original amendments known as the Bill of Rights — the First, Fourth, Fifth, Sixth, Seventh and Eighth Amendments — and possibly the Thirteenth and Fourteenth as well.

Example[15]

Fears of Future Terror Attacks Warranted

By Andrew Tardaca
Published: January 17, 2009

U.S. citizens are bracing for another 9/11 type terrorist attack, according to a variety of reports. A recent Gallup poll finds that 87% of the American public is highly concerned about the possibility of a terrorist attack at home. According to new information from several international sources, these fears are well supported.

A raid on a London terrorist hideout on November 9, 2008 resulted in the capture of computer files that identified numerous U.S. financial districts, cultural centers, and transportation systems on a list of future Al Qaeda targets. According to a recent overseas intelligence report, "al Qaeda already has several cells operating in the U.S. that may be on the verge of mounting a large-scale terrorist attack."

On September 11, 2001, Al Qaeda's attacks killed nearly 3,000 men, women, and children, and injured over 6,000 more. Since September 11th, Al Qaeda and groups affiliated with Al Qaeda have waged attacks in countries such as Egypt, Indonesia, Kenya, Morocco, Saudi Arabia, Spain, Turkey, the United Kingdom, and most recently India. U.S. security officials are warning that current terrorist plots include plans for attacks on U.S. soil at least twice the magnitude of 9/11. An anonymous source reported that recent intelligence documents contain "sobering information" concerning the magnitude of future terrorist attacks.

Warnings issued by extremist groups such as Al Qaeda to "attack U.S. interests and allies on its soil" are even more alarming given the state of preparedness for future incidents. Experts have issued warnings about

---

[15] Merolla & Zechmeister. 2013. "Evaluating Political Leaders in Times of Terror and Economic Threat: The Conditioning Influence of Politician Partisanship." *Journal of Politics* 75(3): 599–712.

Example[15]

Economic Recession Projected to Deepen

By Andrew Tardaca
Published: January 17, 2009

U.S. citizens are bracing for a drastic deepening of the current economic recession. A recent Gallup poll finds that 87% of the American public is highly concerned about economic conditions in the country. The report further states "The economic mood is grimmer than it has been since 1992."

On September 16, failures of large financial institutions in the United States, such as Lehman Brothers and AIG, rapidly evolved into a global crisis resulting in bank failures across the U.S. and Europe. In the United States alone, 15 banks failed in 2008, while several others were rescued through government intervention or acquisitions by other banks. These events led to sharp reductions in the value of stocks and commodities worldwide. Over the past year, the Dow Jones Industrial Average lost 33.8%, the third worst loss in our nation's history. On October 11, 2008, the head of the International Monetary Fund (IMF) warned that the world financial system is teetering on the "brink of systemic meltdown".

The bank failures and subsequent market collapse were tied to sub-prime loans and credit default swaps. Increasing interest rates on loans hit the housing market particularly hard, as individuals were unable to keep up with mortgage payments. 2008 witnessed a record number of foreclosures, leading to the worst housing crisis, banking failure, and market collapse since the Great Depression.

Future projections are looking even grimmer. Experts predict that the housing market will not recover for at least a decade, especially now that banks are hesitant to make loans. The downturn in the economy has led to

---

[15] Merolla & Zechmeister. 2013. "Evaluating Political Leaders in Times of Terror and Economic Threat: The Conditioning Influence of Politician Partisanship." *Journal of Politics* 75(3): 599–712.

Some vignette considerations

Some vignette considerations

- Comparability across conditions
  - Length
  - Readability

Some vignette considerations

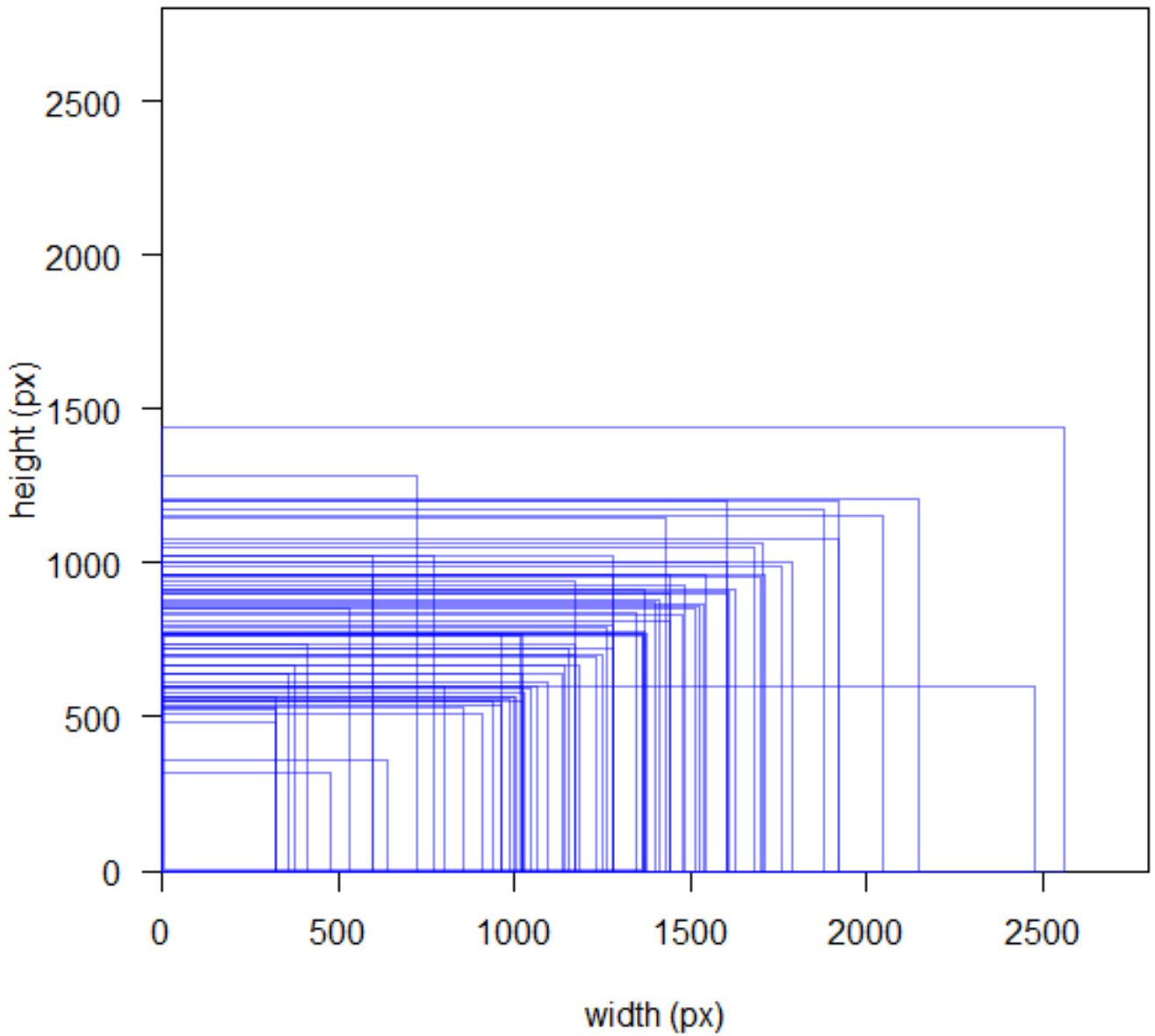
- Comparability across conditions
  - Length
  - Readability
- Language proficiency

Some vignette considerations

- Comparability across conditions
  - Length
  - Readability
- Language proficiency
- Length
  - Timers
  - Forced exposure
  - Mouse trackers

Some vignette considerations

- Comparability across conditions
  - Length
  - Readability
- Language proficiency
- Length
  - Timers
  - Forced exposure
  - Mouse trackers
- Devices
  - Browser-specificity
  - Device sizes (e.g., mobile)


height (px)
2500
2000
1500
1000
500
0
0
500
1000
1500
2000
2500
width (px)

Aside: Unique features of online studies

Aside: Unique features of online studies

- Capacity for audio-visual treatments and measurements

Aside: Unique features of online studies

- Capacity for audio-visual treatments and measurements
- Paradata collection
  - Implicit outcomes like response times, answer switching, mouse click behavior, browser focus, eye tracking, etc.

Aside: Unique features of online studies

- Capacity for audio-visual treatments and measurements
- Paradata collection
  - Implicit outcomes like response times, answer switching, mouse click behavior, browser focus, eye tracking, etc.
- Complex randomization (Thursday)

Aside: Unique features of online studies

- Capacity for audio-visual treatments and measurements
- Paradata collection
  - Implicit outcomes like response times, answer switching, mouse click behavior, browser focus, eye tracking, etc.
- Complex randomization (Thursday)
- Panel data
- Synchronous, multi-person designs

Non-textual Manipulations

- Images can work well
- Standalone or embedded in a text or question

[16] "Cueing Patriotism, Prejudice, and Partisanship in the Age of Obama: Experimental Tests of U.S. Flag Imagery Effects in Presidential Elections." *Political Psychology*: in press.

Non-textual Manipulations

- Images can work well
- Standalone or embedded in a text or question
- Examples
  - Kalmoe & Gross[16] measure impact of patriotic cues on candidate support by showing images of candidates with and without flags

[16] "Cueing Patriotism, Prejudice, and Partisanship in the Age of Obama: Experimental Tests of U.S. Flag Imagery Effects in Presidential Elections." *Political Psychology*: in press.

Non-textual Manipulations

- Images can work well
- Standalone or embedded in a text or question
- Examples
  - Kalmoe & Gross[16] measure impact of patriotic cues on candidate support by showing images of candidates with and without flags
  - Subliminal primes possible, depending on software

[16] "Cueing Patriotism, Prejudice, and Partisanship in the Age of Obama: Experimental Tests of U.S. Flag Imagery Effects in Presidential Elections." *Political Psychology*: in press.

Non-textual Manipulations

- Images can work well
- Standalone or embedded in a text or question
- Examples
  - Kalmoe & Gross[16] measure impact of patriotic cues on candidate support by showing images of candidates with and without flags
  - Subliminal primes possible, depending on software
  - Lots of recent examples of facial manipulation

---

[16] "Cueing Patriotism, Prejudice, and Partisanship in the Age of Obama: Experimental Tests of U.S. Flag Imagery Effects in Presidential Elections." *Political Psychology*: in press.

Example[17]

Light Complexion

Original

Dark Complexion

[17] Iyengar et al. 2010. "Do Explicit Racial Cues Influence Candidate Preference? The Case of Skin Complexion in the 2008 Campaign." Working paper.

Example[18]



[18] Laustsen & Petersen. 2016. "Winning Faces vary by Ideology." *Political Communication* 33(2): 188–211.

Example[19]

[19] Bailenson et al. 2006. "Transformed Facial Similarity as a Political Cue: A Preliminary Investigation." *Political Psychology* 27(3): 373–385.

Audio & Video manipulations

- Problematic for same reasons as long texts

[20] Vavreck. 2007 "The Exaggerated Effects of Advertising on Turnout: The Dangers of Self-Reports." *Quarterly Journal of Political Science* 2: 325–343.

[21] Mutz. 2007. "Effects of 'In-Your-Face' Television Discourse on Perceptions of a Legitimate Opposition." *American Political Science Review* 101(4): 621–635.

Audio & Video manipulations

- Problematic for same reasons as long texts
- Best practices
  - Keep it short
  - Have the video play automatically
  - Disallow survey progression
  - Control and validate

---

[20] Vavreck. 2007 "The Exaggerated Effects of Advertising on Turnout: The Dangers of Self-Reports." *Quarterly Journal of Political Science* 2: 325–343.

[21] Mutz. 2007. "Effects of 'In-Your-Face' Television Discourse on Perceptions of a Legitimate Opposition." *American Political Science Review* 101(4): 621–635.

Audio & Video manipulations

- Problematic for same reasons as long texts
- Best practices
  - Keep it short
  - Have the video play automatically
  - Disallow survey progression
  - Control and validate
- Examples
  - Television Advertisements[20]
  - News Programs[21]

---

[20] Vavreck. 2007 "The Exaggerated Effects of Advertising on Turnout: The Dangers of Self-Reports." *Quarterly Journal of Political Science* 2: 325–343.

[21] Mutz. 2007. "Effects of 'In-Your-Face' Television Discourse on Perceptions of a Legitimate Opposition." *American Political Science Review* 101(4): 621–635.

"Task" Designs

- Task designs ask respondents to perform a task
- Often developed for laboratory settings

"Task" Designs

- Task designs ask respondents to perform a task
- Often developed for laboratory settings
- Most common example: writing something

"Task" Designs

- Task designs ask respondents to perform a task
- Often developed for laboratory settings
- Most common example: writing something
- Can be problematic:
  - Time-intensive
  - Invites drop-off
  - Compliance problems

Univalent

These days, Democrats and Republicans differ from one another considerably. The two groups seem to be growing further and further apart, not only in terms of their opinions but also their lifestyles. Earlier in the survey, you said you tend to identify as a *Democrat/ Republican*. Please take a few minutes to think about what you like about *Democrats/ Republicans* compared to the *Republicans/ Democrats*. Think of 2 to 3 things you especially like best about **your party**. Then think of 2 to 3 things you especially dislike about **the other party**. Now please write those thoughts in the space below.

Ambivalent

These days, Democrats and Republicans differ from one another considerably. The two groups seem to be growing further and further apart, not only in terms of their opinions but also their lifestyles. Earlier in the survey, you said you tend to identify as a *Democrat/ Republican*. Please take a few minutes to think about what you like about *Democrats/ Republicans* compared to the *Republicans/ Democrats*. Think of 2 to 3 things you especially like best about **the other party**. Then think of 2 to 3 things you especially dislike about **your party**. Now please write those thoughts in the space below.

Questions?

Sensitive Item Designs

- Experiments can also be used to measure something
- Goal here is not necessarily causal inference, though the causal insight of the experiment provides *descriptively* useful information
- Paradigms
  - List experiments
  - Endorsement experiments

List Experiments [22]

Now I'm going to read you three things that sometimes make people angry or upset. After I read all three, just tell me *how many* of them upset you. I don't want to know which ones. just *how many*.

1. the federal government increasing the tax on gasoline
2. professional athletes getting million-dollar salaries
3. large corporations polluting the environment

[22] Kuklinski et al. 1997. "Racial Prejudice and Attitudes Toward Affirmative Action." *American Journal of Political Science* 41(2): 402–419.

List Experiments [22]

Now I'm going to read you three things that sometimes make people angry or upset. After I read all three, just tell me *how many* of them upset you. I don't want to know which ones. just *how many*.

1. the federal government increasing the tax on gasoline
2. professional athletes getting million-dollar salaries
3. large corporations polluting the environment
4. **a black family moving in next door**

---

[22] Kuklinski et al. 1997. "Racial Prejudice and Attitudes Toward Affirmative Action." *American Journal of Political Science* 41(2): 402–419.

Endorsement experiments[23]

A recent proposal calls for the sweeping reform of the Afghan prison system, including the construction of new prisons in every district to help alleviate overcrowding in existing facilities. Though expensive, new programs for inmates would also be offered, and new judges and prosecutors would be trained. How do you feel about this proposal?

---

[23] Lyall, Blair, & Imai. 2013. "Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan." *American Political Science Review* 107(4): 679–705.

Endorsement experiments[23]

A recent proposal **by the Taliban** calls for the sweeping reform of the Afghan prison system, including the construction of new prisons in every district to help alleviate overcrowding in existing facilities. Though expensive, new programs for inmates would also be offered, and new judges and prosecutors would be trained. How do you feel about this proposal?

---

[23] Lyall, Blair, & Imai. 2013. "Explaining Support for Combatants during Wartime: A Survey Experiment in Afghanistan." *American Political Science Review* 107(4): 679–705.

Questions?

Beyond One-shot Designs

- Surveys can be used as a measurement instrument for a field treatment or a manipulation applied in a different survey panel wave
  1. Measure effect duration in two-wave panel
  2. Solicit pre-treatment outcome measures in a two-wave panel
  3. Measure effects of field treatment in post-test only design
  4. Randomly encourage field treatment in pre-test and measure effects in post-test

Beyond One-shot Designs

- Surveys can be used as a measurement instrument for a field treatment or a manipulation applied in a different survey panel wave
  1. Measure effect duration in two-wave panel
  2. Solicit pre-treatment outcome measures in a two-wave panel
  3. Measure effects of field treatment in post-test only design
  4. Randomly encourage field treatment in pre-test and measure effects in post-test
- Problems? Compliance & nonresponse

I. Effect Duration

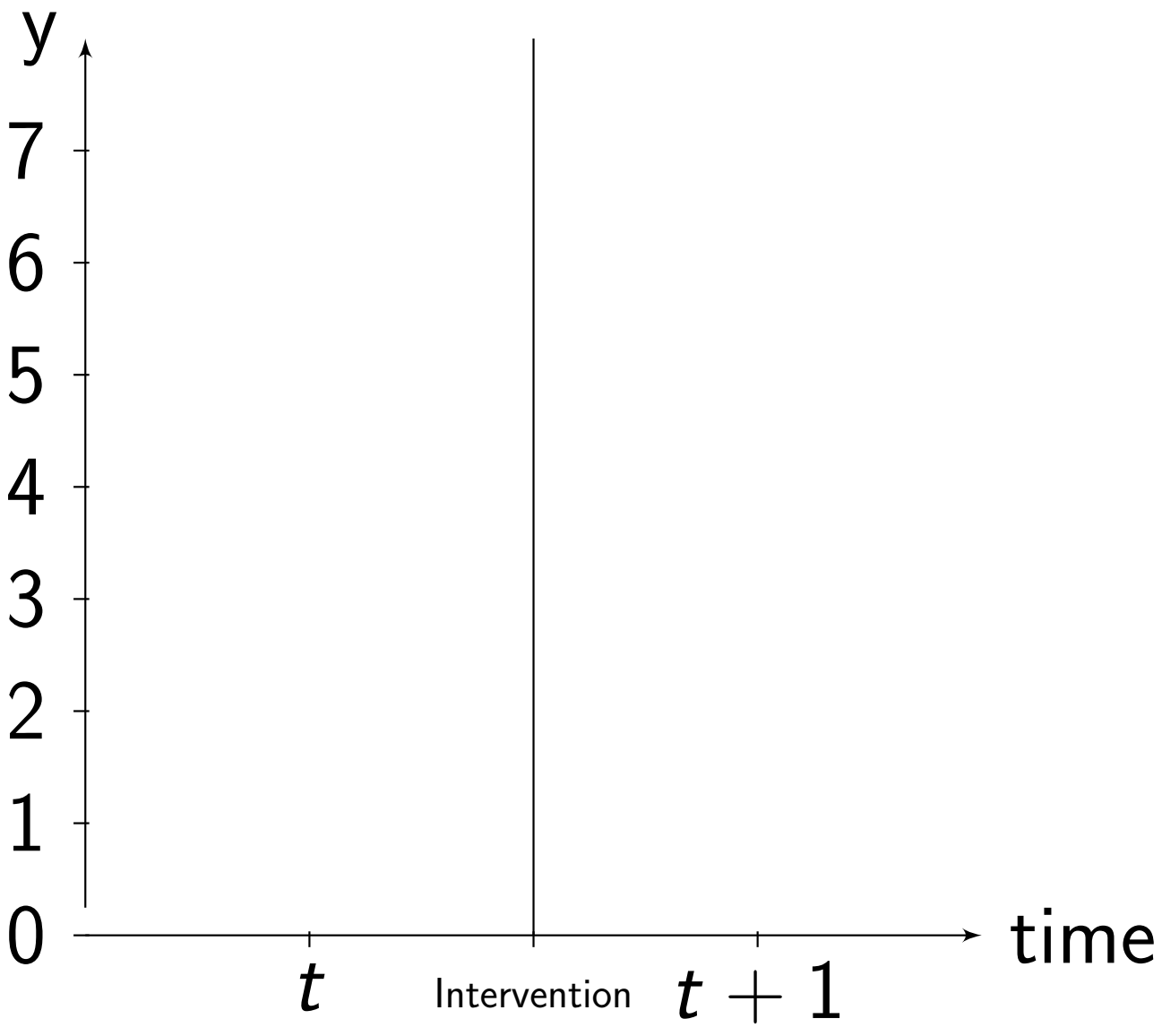
- Use a two- (or more-) wave panel to measure duration of effects
  - T1: Treatment and outcome measurement
  - T2+: Outcome measurement
- Two main concerns (non unique to survey experiments)
  - Attrition
  - Panel conditioning

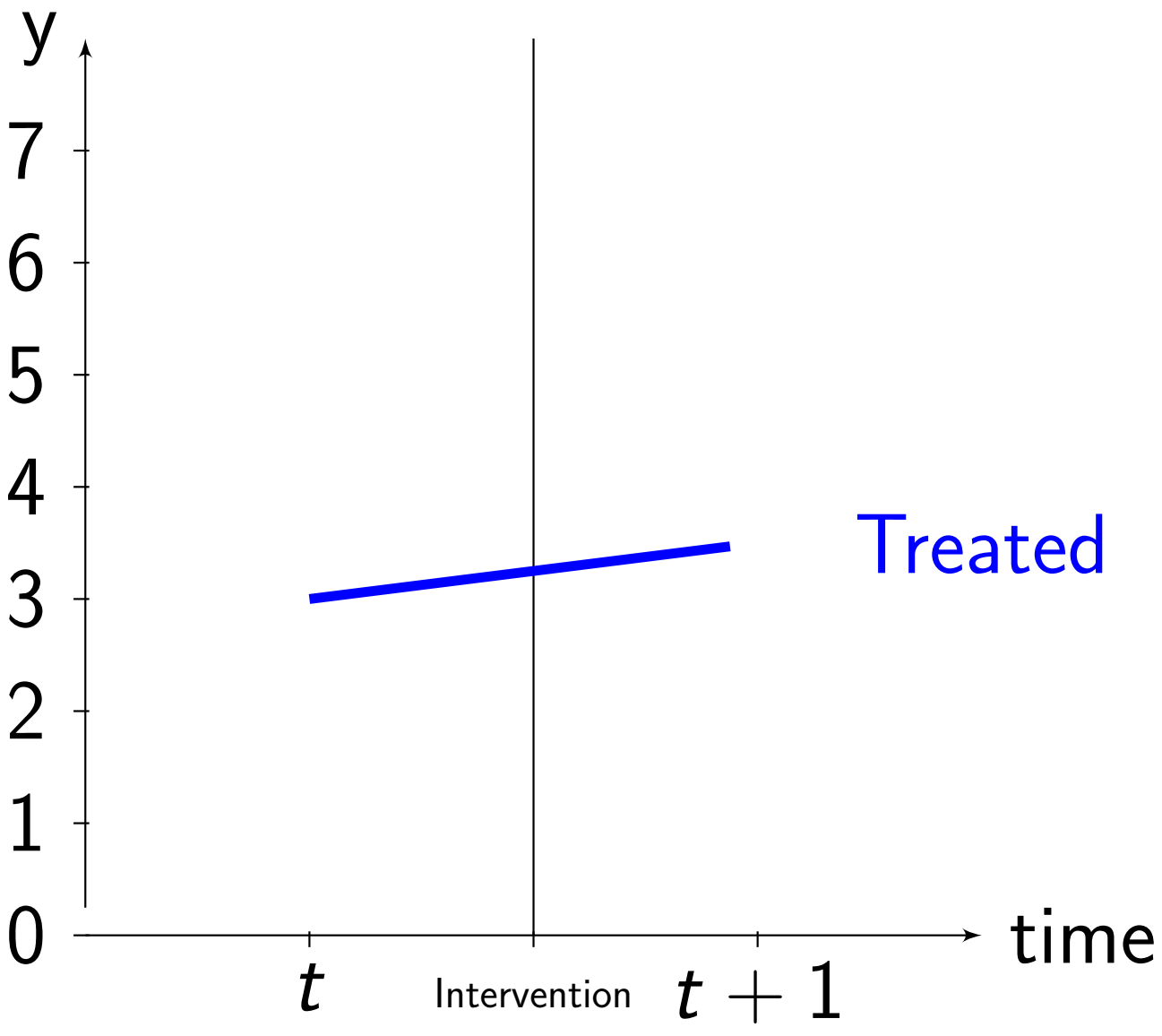
II. Within-Subjects Designs

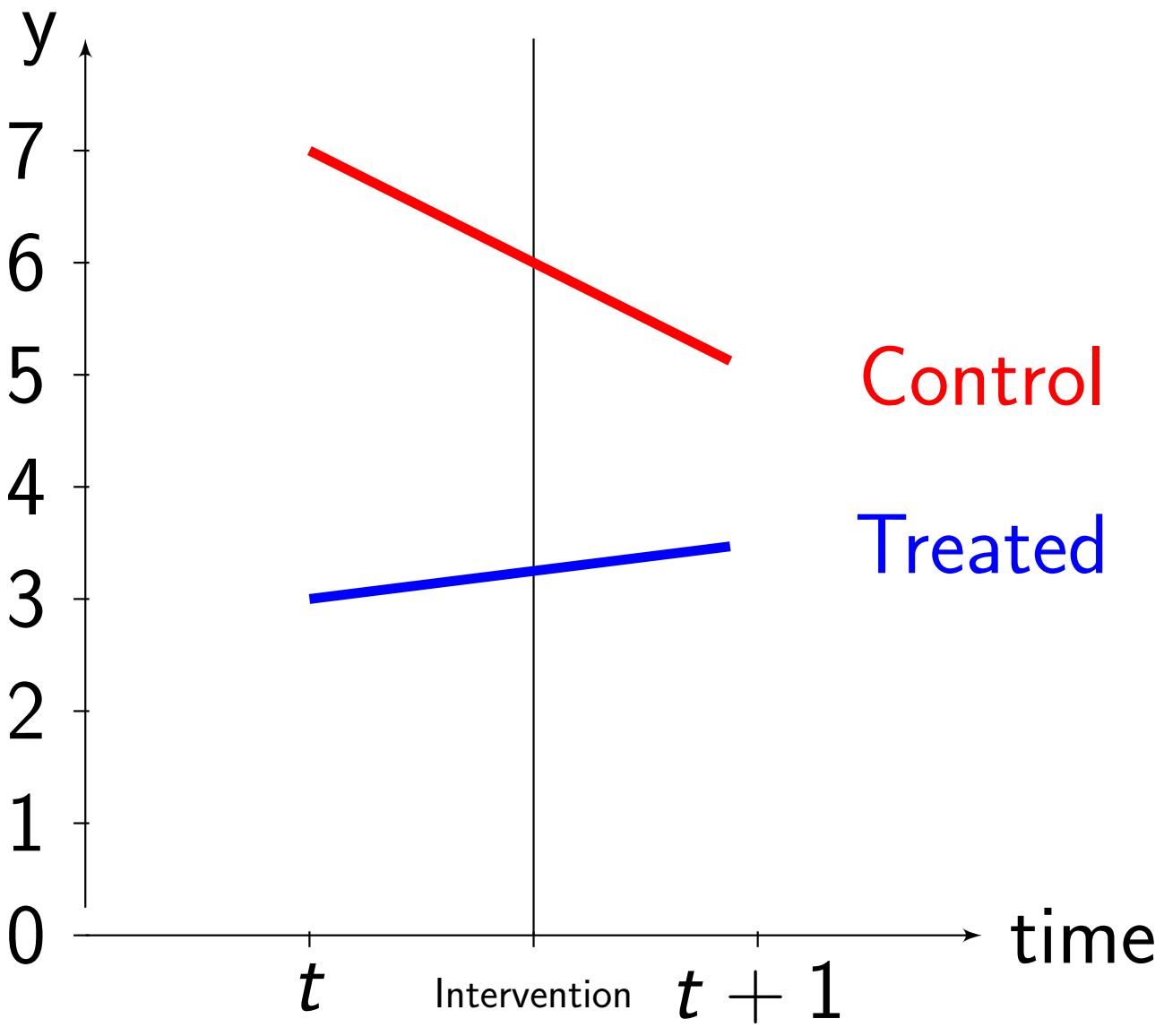
- Estimate treatment effects as a difference-in-differences
- Instead of using the post-treatment mean-difference in Y to estimate the causal effect, use the difference in pre-post differences for the two groups:

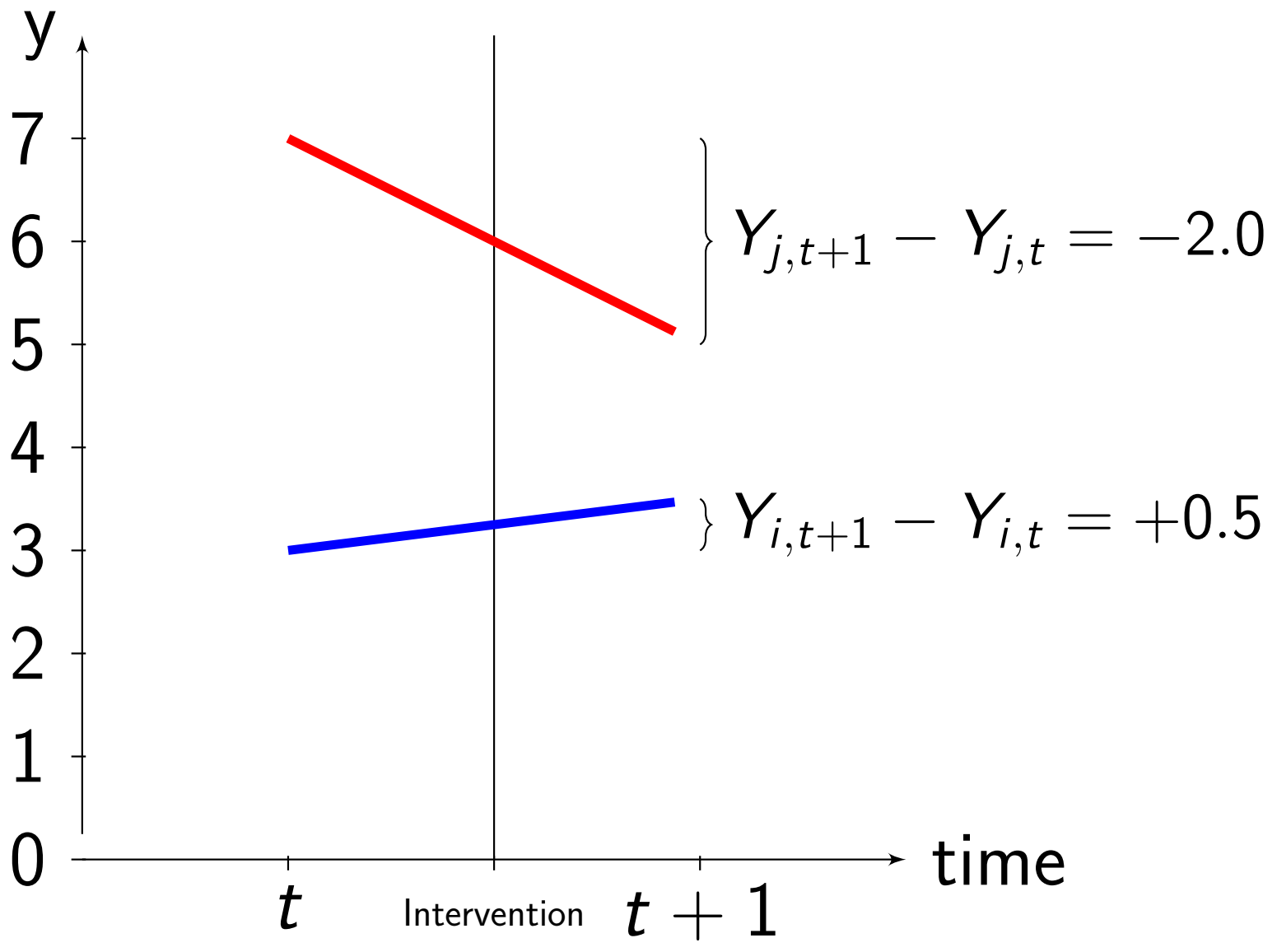
$$(\hat{Y}_{0,t+1} - \hat{Y}_{0,t}) - (\hat{Y}_{j,t+1} - \hat{Y}_{j,t})$$

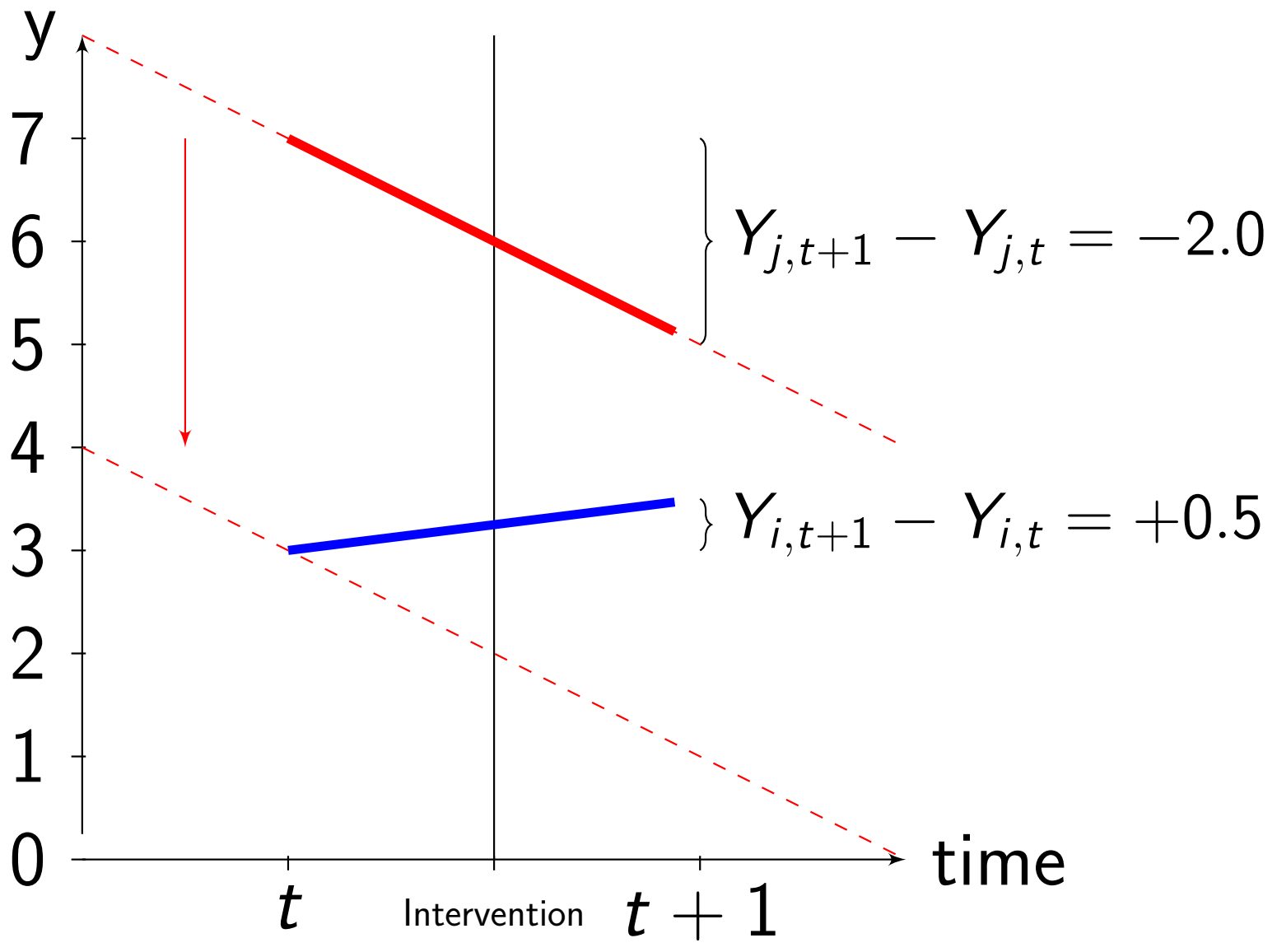
- Advantageous because variance for paired samples decreases as correlation between t_0 and t_1 observations increases

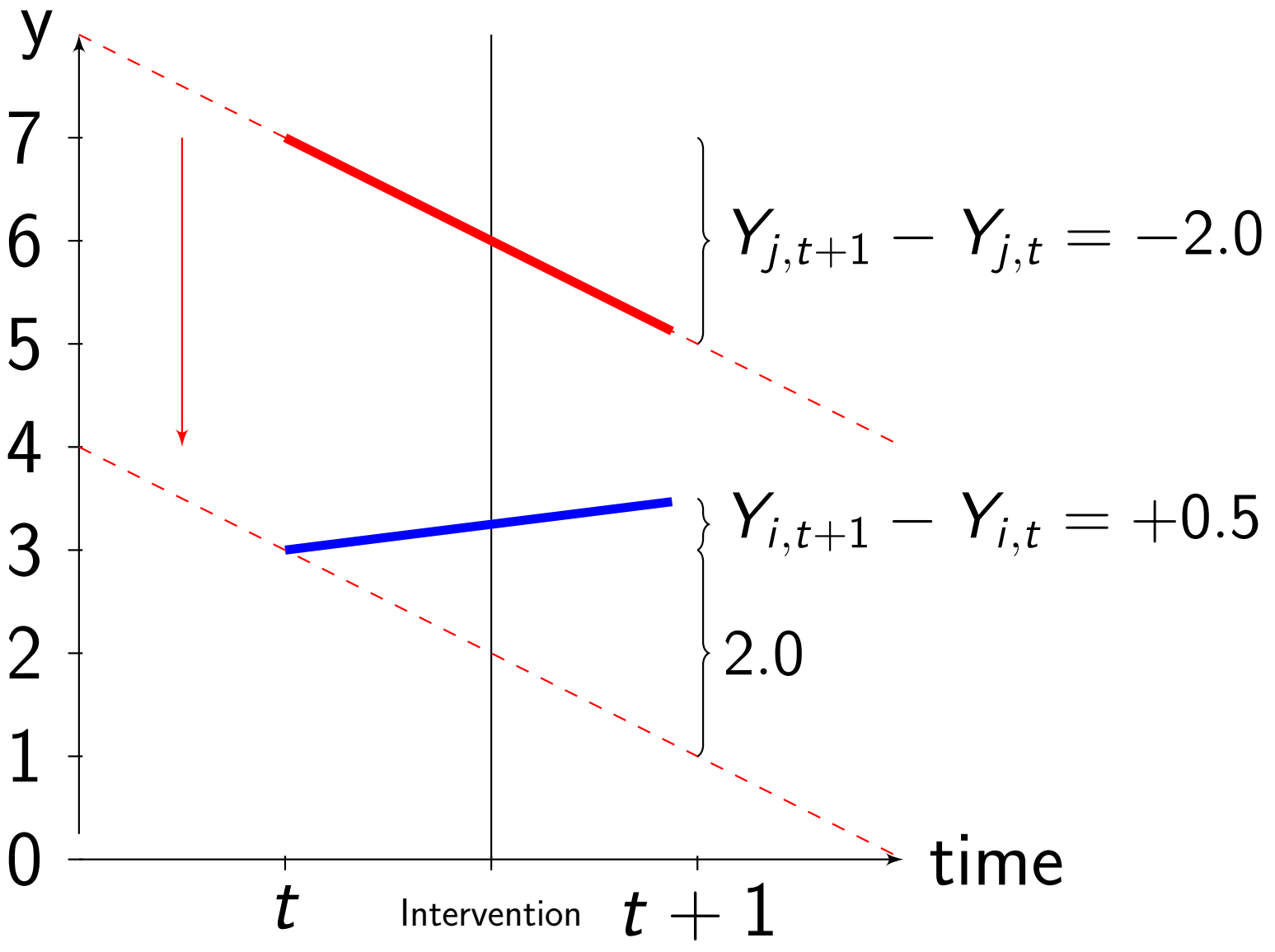

y
7
6
5
4
3
2
1
0
time
t
Intervention
$t+1$

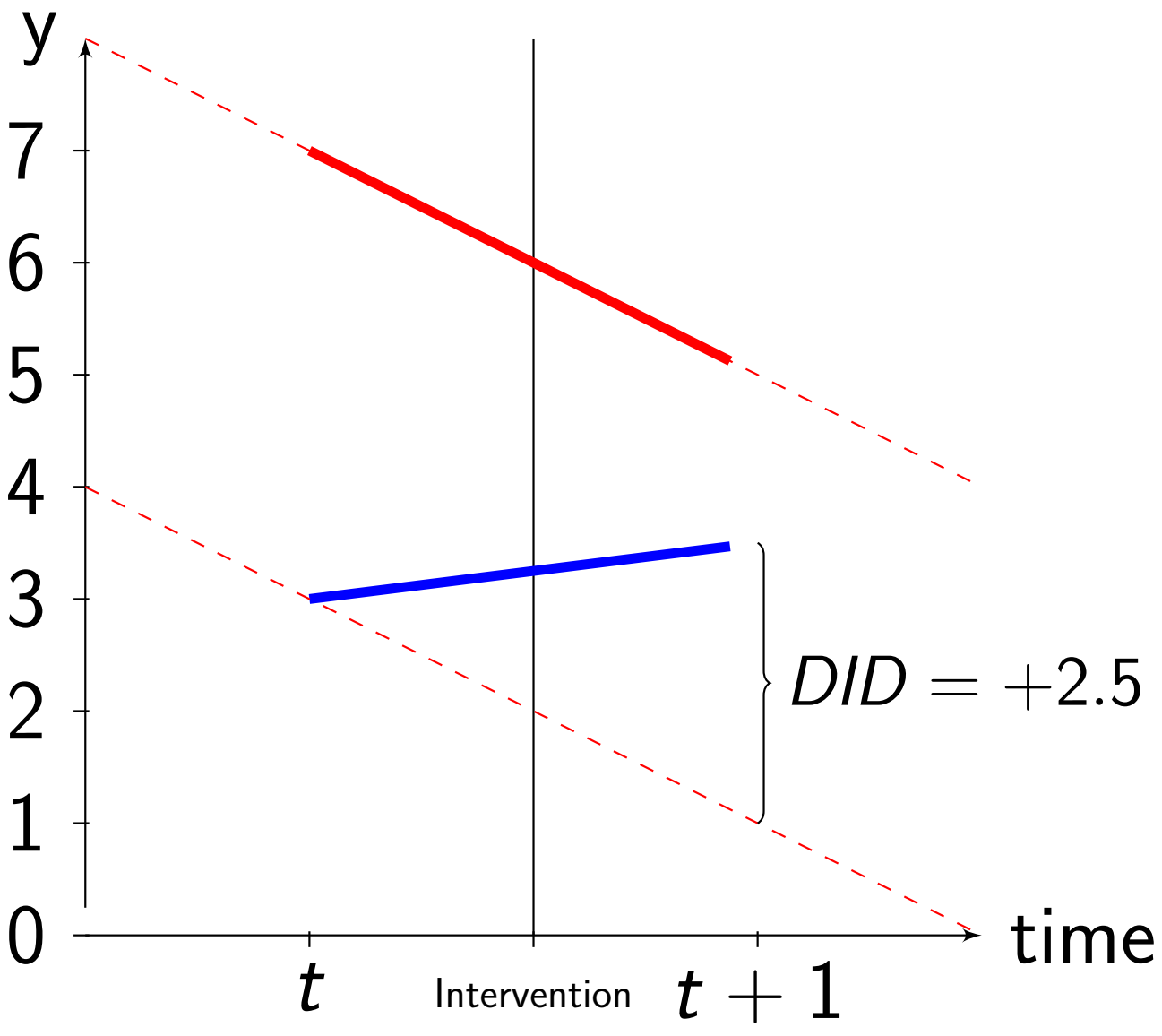

y
7
6
5
4
3
2
1
0
Treated
time
t
Intervention
$t+1$


y
7
6
5
4
3
2
1
0
Control
Treated
time
t
Intervention
$t+1$


y
7
6
5
4
3
2
1
0
$Y_{j,t+1} - Y_{j,t} = -2.0$
$Y_{i,t+1} - Y_{i,t} = +0.5$
time
t
Intervention
$t + 1$


y
7
6
5
4
3
2
1
0
$Y_{j,t+1} - Y_{j,t} = -2.0$
$Y_{i,t+1} - Y_{i,t} = +0.5$
time
t
Intervention
$t+1$


y
7
6
5
4
3
2
1
0
$Y_{j,t+1} - Y_{j,t} = -2.0$
$Y_{i,t+1} - Y_{i,t} = +0.5$
2.0
time
t
Intervention
$t + 1$


y
7
6
5
4
3
2
1
0
DID = +2.5
time
t
Intervention
t + 1

Threats to Validity

As soon as time comes into play, we have to worry about threats to validity.[24]

[24] Shadish, Cook, and Campbell (2002)

Threats to Validity

As soon as time comes into play, we have to worry about threats to validity.[24]

1. History (simultaneous cause)

[24] Shadish, Cook, and Campbell (2002)

Threats to Validity

As soon as time comes into play, we have to worry about threats to validity.[24]

1. History (simultaneous cause)
2. Maturation (time trends)

[24] Shadish, Cook, and Campbell (2002)

Threats to Validity

As soon as time comes into play, we have to worry about threats to validity.[24]

1. History (simultaneous cause)
2. Maturation (time trends)
3. Testing (observation changes respondents)

[24] Shadish, Cook, and Campbell (2002)

Threats to Validity

As soon as time comes into play, we have to worry about threats to validity.[24]

1. History (simultaneous cause)
2. Maturation (time trends)
3. Testing (observation changes respondents)
4. Instrumentation (changing operationalization)

[24] Shadish, Cook, and Campbell (2002)

Threats to Validity

As soon as time comes into play, we have to worry about threats to validity.[24]

1. History (simultaneous cause)
2. Maturation (time trends)
3. Testing (observation changes respondents)
4. Instrumentation (changing operationalization)
5. Instability (measurement error)

[24] Shadish, Cook, and Campbell (2002)

Threats to Validity

As soon as time comes into play, we have to worry about threats to validity.[24]

1. History (simultaneous cause)
2. Maturation (time trends)
3. Testing (observation changes respondents)
4. Instrumentation (changing operationalization)
5. Instability (measurement error)
6. Attrition

[24] Shadish, Cook, and Campbell (2002)

III. Randomized Field Treatment

- Examples:

III. Randomized Field Treatment

- Examples:
  1. Citizens randomly sent a letter by post encouraging them to reduce water usage

III. Randomized Field Treatment

- Examples:
  1. Citizens randomly sent a letter by post encouraging them to reduce water usage
  2. Different local media markets randomly assigned to receive different advertising

III. Randomized Field Treatment

- Examples:
  1. Citizens randomly sent a letter by post encouraging them to reduce water usage
  2. Different local media markets randomly assigned to receive different advertising
- Survey is used to measure outcomes, when treatment assignment is already known

III. Randomized Field Treatment

- Examples:
  1. Citizens randomly sent a letter by post encouraging them to reduce water usage
  2. Different local media markets randomly assigned to receive different advertising
- Survey is used to measure outcomes, when treatment assignment is already known
- Issues

III. Randomized Field Treatment

- Examples:
  1. Citizens randomly sent a letter by post encouraging them to reduce water usage
  2. Different local media markets randomly assigned to receive different advertising
- Survey is used to measure outcomes, when treatment assignment is already known
- Issues
  - Nonresponse

III. Randomized Field Treatment

- Examples:
  1. Citizens randomly sent a letter by post encouraging them to reduce water usage
  2. Different local media markets randomly assigned to receive different advertising
- Survey is used to measure outcomes, when treatment assignment is already known
- Issues
  - Nonresponse
  - Noncompliance

IV. Treatment Encouragement

- Design:
  - T1: Encourage treatment
  - T2: Measure effects
- Examples:
  1. Albertson and Lawrence

IV. Treatment Encouragement

- Design:
  - T1: Encourage treatment
  - T2: Measure effects
- Examples:
  1. Albertson and Lawrence

IV. Treatment Encouragement

- Design:
  - T1: Encourage treatment
  - T2: Measure effects
- Examples:
  1. Albertson and Lawrence

IV. Treatment Encouragement

- Design:
  - T1: Encourage treatment
  - T2: Measure effects
- Examples:
  1. Albertson and Lawrence

IV. Treatment Encouragement

- Design:
  - T1: Encourage treatment
  - T2: Measure effects
- Examples:
  1. Albertson and Lawrence
- Issues

IV. Treatment Encouragement

- Design:
  - T1: Encourage treatment
  - T2: Measure effects
- Examples:
  1. Albertson and Lawrence
- Issues
  - Attrition

IV. Treatment Encouragement

- Design:
  - T1: Encourage treatment
  - T2: Measure effects
- Examples:
  1. Albertson and Lawrence
- Issues
  - Attrition
  - Noncompliance

Treatment Noncompliance

- Definition:

  "when subjects who were assigned to receive the treatment go untreated or when subjects assigned to the control group are treated" [25]

[25] Gerber & Green. 2012. *Field Experiments*, p.132.

Treatment Noncompliance

- Definition:

  "when subjects who were assigned to receive the treatment go untreated or when subjects assigned to the control group are treated" [25]

- Several strategies
  - "As treated" analysis
  - "Intention to treat" analysis
  - Estimate a LATE

[25] Gerber & Green. 2012. *Field Experiments*, p.132.

Analyzing Noncompliance

- If noncompliance only occurs in one group, it is *asymmetric* or *one-sided*
- We can ignore non-compliance and analyze the "intention to treat" effect, which will underestimate our effects because some people were not treated as assigned: $ITT = \overline{Y}_1 - \overline{Y}_0$
- We can use "instrumental variables" to estimate the "local average treatment effect" (LATE) for those that complied with treatment: $LATE = \frac{ITT}{\%Compliant}$

Local Average Treatment Effect

- IV estimate is *local* to the variation in X that is due to variation W (i.e., the LATE)
- This matters if effects are *heterogeneous*
- LATE is effect for those who *comply* with instrument
- Four subpopulations:
  - Compliers: $X = 1$ only if $D = 1$
  - Always-takers: $X = 1$ regardless of D
  - Never-takers: $X = 0$ regardless of D
  - Defiers: $X = 1$ only if $D = 0$
- Exclusion restriction! Monotonicity!

Questions?

Homework!

- Working alone, with a partner, or in a small group, start thinking about a possible experiment
- Start writing a protocol:
  1. Theory/hypotheses
  2. Instrumentation
- Time on Friday to share and get feedback