Causal Inference under Interference

Ye Wang
University of North Carolina at Chapel Hill

Why do we care about causality?

- Most social science theories take the form of causal relationships.

Why do we care about causality?

- Most social science theories take the form of causal relationships.
- What would happen to Y if Z changes?

Why do we care about causality?

- Most social science theories take the form of causal relationships.
- What would happen to Y if Z changes?
- We call Y the outcome and Z the treatment.

Why do we care about causality?

- Most social science theories take the form of causal relationships.
- What would happen to Y if Z changes?
- We call Y the outcome and Z the treatment.
- The better we understand causal relationships, the better we can design policy interventions.

How do we define causality?

- We possess a sample of N units.

How do we define causality?

- We possess a sample of N units.
- Denote the outcome of interest for unit i as Y_i and the treatment as Z_i.

How do we define causality?

- We possess a sample of N units.
- Denote the outcome of interest for unit i as Y_i and the treatment as Z_i.
- Then, we have the Neyman-Rubin framework:

$$Y_i = \begin{cases} Y_i(0), \ Z_i = 0 \\ Y_i(1), \ Z_i = 1. \end{cases}$$

How do we define causality?

- We possess a sample of N units.
- Denote the outcome of interest for unit i as Y_i and the treatment as Z_i.
- Then, we have the Neyman-Rubin framework:

$$Y_i = \begin{cases} Y_i(0), \ Z_i = 0 \\ Y_i(1), \ Z_i = 1. \end{cases}$$

- $Y_i(z)$ is called the "potential outcome."

How do we define causality?

- We possess a sample of N units.
- Denote the outcome of interest for unit i as Y_i and the treatment as Z_i.
- Then, we have the Neyman-Rubin framework:

$$Y_i = \begin{cases} Y_i(0), \; Z_i = 0 \\ Y_i(1), \; Z_i = 1. \end{cases}$$

- $Y_i(z)$ is called the "potential outcome."
- $\tau_i = Y_i(1) - Y_i(0)$ is the individualistic treatment effect.

How do we define causality?

- We possess a sample of N units.
- Denote the outcome of interest for unit i as Y_i and the treatment as Z_i.
- Then, we have the Neyman-Rubin framework:

$$Y_i = \begin{cases} Y_i(0), \; Z_i = 0 \\ Y_i(1), \; Z_i = 1. \end{cases}$$

- $Y_i(z)$ is called the "potential outcome."
- $\tau_i = Y_i(1) - Y_i(0)$ is the individualistic treatment effect.
- $\tau = \frac{1}{N} \sum_{i=1}^{N} \tau_i$ is known as the average treatment effect (ATE).

Interference

- The Neyman-Rubin framework is built upon the Stable Unit Treatment Value Assumption (SUTVA).

Interference

- The Neyman-Rubin framework is built upon the Stable Unit Treatment Value Assumption (SUTVA).
- Interference occurs when SUTVA is violated: the outcome of one unit can be affected by the treatment of other units.

Interference

- The Neyman-Rubin framework is built upon the Stable Unit Treatment Value Assumption (SUTVA).
- Interference occurs when SUTVA is violated: the outcome of one unit can be affected by the treatment of other units.
- It is also known as “spillover effect,” “diffusion effect,” or “peer effect” in the literature.

Interference

- The Neyman-Rubin framework is built upon the Stable Unit Treatment Value Assumption (SUTVA).
- Interference occurs when SUTVA is violated: the outcome of one unit can be affected by the treatment of other units.
- It is also known as "spillover effect," "diffusion effect," or "peer effect" in the literature.
- Interference differs from contagion.

Interference

- The Neyman-Rubin framework is built upon the Stable Unit Treatment Value Assumption (SUTVA).
- Interference occurs when SUTVA is violated: the outcome of one unit can be affected by the treatment of other units.
- It is also known as "spillover effect," "diffusion effect," or "peer effect" in the literature.
- Interference differs from contagion.
- The phenomenon is prevalent in the real world:
  - Wind turbines erected in one district may change public opinion in nearby areas over a long period (Stokes 2016).

Interference

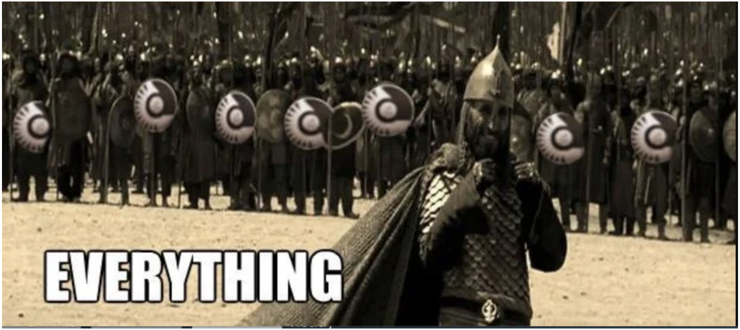
- The Neyman-Rubin framework is built upon the Stable Unit Treatment Value Assumption (SUTVA).
- Interference occurs when SUTVA is violated: the outcome of one unit can be affected by the treatment of other units.
- It is also known as “spillover effect,” “diffusion effect,” or “peer effect” in the literature.
- Interference differs from contagion.
- The phenomenon is prevalent in the real world:
  - Wind turbines erected in one district may change public opinion in nearby areas over a long period (Stokes 2016).
  - Protests in fixed locations can alter the political choice of bystanders over subsequent elections (Wang and Wong 2021).

Interference

- What does interference change in the Neyman-Rubin framework?

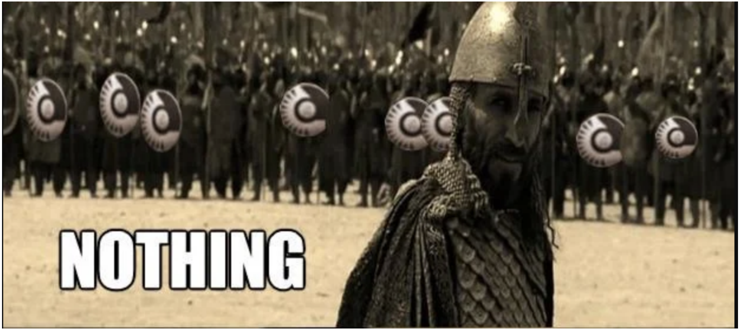
Interference

- What does interference change in the Neyman-Rubin framework?



Interference

- What does interference change in the Neyman-Rubin framework?



Interference in the Neyman-Rubin framework

- Under interference, the potential outcome of unit i can be written as

$$Y_i = Y_i\left(\mathbf{Z}\right), \text{ where } \mathbf{Z} = (Z_1, Z_2, \ldots, Z_N)'.$$

Interference in the Neyman-Rubin framework

- Under interference, the potential outcome of unit i can be written as

$$Y_i = Y_i(\mathbf{Z}), \text{ where } \mathbf{Z} = (Z_1, Z_2, \ldots, Z_N)'.$$

- For each i, there can be 2^N potential outcomes.

Interference in the Neyman-Rubin framework

- Under interference, the potential outcome of unit i can be written as

$$Y_i = Y_i\left(\mathbf{Z}\right), \text{ where } \mathbf{Z} = (Z_1, Z_2, \ldots, Z_N)'.$$

- For each i, there can be 2^N potential outcomes.
- There are multiple challenges:
  - What causal estimands can be defined, and what identification assumptions do we need?

Interference in the Neyman-Rubin framework

- Under interference, the potential outcome of unit i can be written as

$$Y_i = Y_i(\mathbf{Z}), \text{ where } \mathbf{Z} = (Z_1, Z_2, \ldots, Z_N)'.$$

- For each i, there can be 2^N potential outcomes.
- There are multiple challenges:
  - What causal estimands can be defined, and what identification assumptions do we need?
  - How do we isolate the direct effect from the spillover effect driven by interference?

Interference in the Neyman-Rubin framework

- Under interference, the potential outcome of unit i can be written as

$$Y_i = Y_i(\mathbf{Z}), \text{ where } \mathbf{Z} = (Z_1, Z_2, \ldots, Z_N)'.$$

- For each i, there can be 2^N potential outcomes.
- There are multiple challenges:
  - What causal estimands can be defined, and what identification assumptions do we need?
  - How do we isolate the direct effect from the spillover effect driven by interference?
  - How can we conduct statistical inference?

Define causal effects under interference

- Consider a simple experiment with two subjects and Bernoulli assignment

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	7	5
(1, 0)	0.25	7	5
(0, 1)	0.25	2	3
(0, 0)	0.25	2	3

Define causal effects under interference

- Consider a simple experiment with two subjects and Bernoulli assignment

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	7	5
(1, 0)	0.25	7	5
(0, 1)	0.25	2	3
(0, 0)	0.25	2	3

- $Y_{Ye}(1) = 7$, $Y_{Ye}(0) = 2$, and $\tau_{Ye} = 5$.

Define causal effects under interference

- Consider a simple experiment with two subjects and Bernoulli assignment

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	7	5
(1, 0)	0.25	7	5
(0, 1)	0.25	2	3
(0, 0)	0.25	2	3

- $Y_{Ye}(1) = 7,\ Y_{Ye}(0) = 2$, and $\tau_{Ye} = 5$.
- $Y_{Jiawei}(1) = 5,\ Y_{Jiawei}(0) = 3$, and $\tau_{Jiawei} = 2$.

Define causal effects under interference

- Consider a simple experiment with two subjects and Bernoulli assignment

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	7	5
(1, 0)	0.25	7	5
(0, 1)	0.25	2	3
(0, 0)	0.25	2	3

- $Y_{Ye}(1) = 7,\ Y_{Ye}(0) = 2$, and $\tau_{Ye} = 5$.
- $Y_{Jiawei}(1) = 5,\ Y_{Jiawei}(0) = 3$, and $\tau_{Jiawei} = 2$.
- $\tau = \frac{1}{2}(5 + 2) = 3.5$.

Define causal effects under interference

- Consider the same experiment under interference

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

Define causal effects under interference

- Consider the same experiment under interference

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

- It is obvious that Y_{Ye} depends on the treatment status of both subjects: $Y_{Ye} = Y_{Ye}(Z_{Ye}, Z_{Jiawei})$.

Define causal effects under interference

- Consider the same experiment under interference

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

- It is obvious that Y_{Ye} depends on the treatment status of both subjects: $Y_{Ye} = Y_{Ye}\left(Z_{Ye}, Z_{Jiawei}\right)$.
- $\tau_{Ye}(Z_{Jiawei}) = Y_{Ye}\left(1, Z_{Jiawei}\right) - Y_{Ye}\left(0, Z_{Jiawei}\right)$ is now a random variable.

Define causal effects under interference

- Consider the same experiment under interference

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

- It is obvious that Y_{Ye} depends on the treatment status of both subjects: $Y_{Ye} = Y_{Ye}(Z_{Ye}, Z_{Jiawei})$.
- $\tau_{Ye}(Z_{Jiawei}) = Y_{Ye}(1, Z_{Jiawei}) - Y_{Ye}(0, Z_{Jiawei})$ is now a random variable.
- So is the ATE.

Define causal effects under interference

- Consider the same experiment under interference

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

- It is obvious that Y_{Ye} depends on the treatment status of both subjects: $Y_{Ye} = Y_{Ye}(Z_{Ye}, Z_{Jiawei})$.
- $\tau_{Ye}(Z_{Jiawei}) = Y_{Ye}(1, Z_{Jiawei}) - Y_{Ye}(0, Z_{Jiawei})$ is now a random variable.
- So is the ATE.
- To obtain meaningful estimands, we marginalize $\tau_{Ye}(Z_{Jiawei})$ over Z_{Jiawei}.

Define causal effects under interference

- Consider the same experiment under interference

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

Define causal effects under interference

- Consider the same experiment under interference

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

$$\tau_{Ye} = 0.5 * 4 + 0.5 * 5 = 4.5$$

$$\tau_{Jiawei} = 0.5 * 1 + 0.5 * 2 = 1.5$$

Define causal effects under interference

- Consider the same experiment under interference

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

$$\tau_{Ye} = 0.5 * 4 + 0.5 * 5 = 4.5$$
$$\tau_{Jiawei} = 0.5 * 1 + 0.5 * 2 = 1.5$$

- Then,
$$\tau = \frac{1}{2}(4.5 + 1.5) = 3.$$

Define causal effects under interference

- Consider the general case with N units and an arbitrary treatment assignment mechanism.

Define causal effects under interference

- Consider the general case with N units and an arbitrary treatment assignment mechanism.
- For each unit, we have

$$\tau_i = E_{\mathbf{Z}_{-i}}\left[Y_i(1, \mathbf{Z}_{-i}) - Y_i(0, \mathbf{Z}_{-i})\right]$$

Define causal effects under interference

- Consider the general case with N units and an arbitrary treatment assignment mechanism.
- For each unit, we have

$$\tau_i = E_{\mathbf{Z}_{-i}}\left[Y_i(1, \mathbf{Z}_{-i}) - Y_i(0, \mathbf{Z}_{-i})\right]$$

- The expected average treatment effect (EATE) is defined as

$$\begin{aligned}\tau &= \frac{1}{N}\sum_{i=1}^{N}\tau_i \\ &= \frac{1}{N}\sum_{i=1}^{N}E_{\mathbf{Z}_{-i}}\left[Y_i(1, \mathbf{Z}_{-i})\right] - \frac{1}{N}\sum_{i=1}^{N}E_{\mathbf{Z}_{-i}}\left[Y_i(0, \mathbf{Z}_{-i})\right].\end{aligned}$$

Define causal effects under interference

- Consider the general case with N units and an arbitrary treatment assignment mechanism.
- For each unit, we have

$$\tau_i = E_{\mathbf{Z}_{-i}}\left[Y_i(1, \mathbf{Z}_{-i}) - Y_i(0, \mathbf{Z}_{-i})\right]$$

- The expected average treatment effect (EATE) is defined as

$$\begin{aligned}\tau &= \frac{1}{N}\sum_{i=1}^{N}\tau_i \\ &= \frac{1}{N}\sum_{i=1}^{N}E_{\mathbf{Z}_{-i}}\left[Y_i(1, \mathbf{Z}_{-i})\right] - \frac{1}{N}\sum_{i=1}^{N}E_{\mathbf{Z}_{-i}}\left[Y_i(0, \mathbf{Z}_{-i})\right].\end{aligned}$$

- The average effect from switching one unit's treatment status from 0 to 1 on its own outcome under the current treatment assignment mechanism.

Estimate the EATE

- Sävje, Aronow, and Hudgens (2021): the EATE is identifiable and can be estimated by the difference-in-means estimator.

Estimate the EATE

- Sävje, Aronow, and Hudgens (2021): the EATE is identifiable and can be estimated by the difference-in-means estimator.
- The estimate will be unbiased and consistent under most experimental designs.

Estimate the EATE

- Sävje, Aronow, and Hudgens (2021): the EATE is identifiable and can be estimated by the difference-in-means estimator.
- The estimate will be unbiased and consistent under most experimental designs.
- In other words, nothing but the interpretation has changed.

Estimate the EATE

- Sävje, Aronow, and Hudgens (2021): the EATE is identifiable and can be estimated by the difference-in-means estimator.
- The estimate will be unbiased and consistent under most experimental designs.
- In other words, nothing but the interpretation has changed.
- The requirement is that treatment assignment is relatively independent across units.

Estimate the EATE

- Sävje, Aronow, and Hudgens (2021): the EATE is identifiable and can be estimated by the difference-in-means estimator.
- The estimate will be unbiased and consistent under most experimental designs.
- In other words, nothing but the interpretation has changed.
- The requirement is that treatment assignment is relatively independent across units.
- Consistency holds for complete randomization but not for paired randomization.

Estimate the EATE

- Sävje, Aronow, and Hudgens (2021): the EATE is identifiable and can be estimated by the difference-in-means estimator.
- The estimate will be unbiased and consistent under most experimental designs.
- In other words, nothing but the interpretation has changed.
- The requirement is that treatment assignment is relatively independent across units.
- Consistency holds for complete randomization but not for paired randomization.
- How about the indirect effect?

Estimate the EATE

- Sävje, Aronow, and Hudgens (2021): the EATE is identifiable and can be estimated by the difference-in-means estimator.
- The estimate will be unbiased and consistent under most experimental designs.
- In other words, nothing but the interpretation has changed.
- The requirement is that treatment assignment is relatively independent across units.
- Consistency holds for complete randomization but not for paired randomization.
- How about the indirect effect?
- Classical methods assume that we know the interference structure: how one's outcome is affected by others' treatments.

Application I

- In Duflo and Saez (2003), the authors try to promote a retirement plan among staff members in a university.

Application I

- In Duflo and Saez (2003), the authors try to promote a retirement plan among staff members in a university.
- 220 out of 330 departments were assigned into the treatment group.

Application I

- In Duflo and Saez (2003), the authors try to promote a retirement plan among staff members in a university.
- 220 out of 330 departments were assigned into the treatment group.
- In each treated department, 50% of staff members who did not enroll in the plan were treated.

Application I

- In Duflo and Saez (2003), the authors try to promote a retirement plan among staff members in a university.
- 220 out of 330 departments were assigned into the treatment group.
- In each treated department, 50% of staff members who did not enroll in the plan were treated.
- Treated staff members received an invite to an information fair on the plan.

Application I

- This two-stage randomization is called a “split-plot” design.

Application I

- This two-stage randomization is called a "split-plot" design.
- There were 2,039 treated staff members and 2,129 untreated ones from departments in the treatment group.

Application I

- This two-stage randomization is called a “split-plot” design.
- There were 2,039 treated staff members and 2,129 untreated ones from departments in the treatment group.
- There were 2,043 staff members from departments in the control group.

Application I

- This two-stage randomization is called a "split-plot" design.
- There were 2,039 treated staff members and 2,129 untreated ones from departments in the treatment group.
- There were 2,043 staff members from departments in the control group.
- The outcome is whether staff member attended the fair and whether they enrolled in the plan.

Application I

- This two-stage randomization is called a “split-plot” design.
- There were 2,039 treated staff members and 2,129 untreated ones from departments in the treatment group.
- There were 2,043 staff members from departments in the control group.
- The outcome is whether staff member attended the fair and whether they enrolled in the plan.
- Staff members may share the information with each other.

Application I

- This two-stage randomization is called a "split-plot" design.
- There were 2,039 treated staff members and 2,129 untreated ones from departments in the treatment group.
- There were 2,043 staff members from departments in the control group.
- The outcome is whether staff member attended the fair and whether they enrolled in the plan.
- Staff members may share the information with each other.
- But we can still estimate the direct effect or the EATE.

Application I

- This two-stage randomization is called a “split-plot” design.
- There were 2,039 treated staff members and 2,129 untreated ones from departments in the treatment group.
- There were 2,043 staff members from departments in the control group.
- The outcome is whether staff member attended the fair and whether they enrolled in the plan.
- Staff members may share the information with each other.
- But we can still estimate the direct effect or the EATE.
- How do we estimate the indirect or spillover effects?

Partial interference

- The authors assume that there exists no interference across departments.

Partial interference

- The authors assume that there exists no interference across departments.
- A staff member's outcome may only be affected by the treatment status of those from her department.

Partial interference

- The authors assume that there exists no interference across departments.
- A staff member's outcome may only be affected by the treatment status of those from her department.
- This assumption is called "partial interference" by Hudgens and Halloran (2008).

Partial interference

- The authors assume that there exists no interference across departments.
- A staff member's outcome may only be affected by the treatment status of those from her department.
- This assumption is called "partial interference" by Hudgens and Halloran (2008).
- We can summarize the effect generated by the other people's treatment status with the department level probability of being treated.

Partial interference

- The authors assume that there exists no interference across departments.
- A staff member's outcome may only be affected by the treatment status of those from her department.
- This assumption is called "partial interference" by Hudgens and Halloran (2008).
- We can summarize the effect generated by the other people's treatment status with the department level probability of being treated.
- In this example, it equals to 50% or 0.

Partial interference

- Hudgens and Halloran (2008) show that we can estimate the indirect effect with the difference in the average outcome across untreated staff members from the treated and untreated departments.

Partial interference

- Hudgens and Halloran (2008) show that we can estimate the indirect effect with the difference in the average outcome across untreated staff members from the treated and untreated departments.
- In Duflo and Saez (2003), the average fair attendance rate is 0.280 among treated staff members from treated departments, 0.151 among untreated staff members from treated departments, and 0.049 among untreated staff members from untreated departments.

Partial interference

- Hudgens and Halloran (2008) show that we can estimate the indirect effect with the difference in the average outcome across untreated staff members from the treated and untreated departments.
- In Duflo and Saez (2003), the average fair attendance rate is 0.280 among treated staff members from treated departments, 0.151 among untreated staff members from treated departments, and 0.049 among untreated staff members from untreated departments.
- The direct effect estimate is $0.280 - 0.151 = 0.129$.

Partial interference

- Hudgens and Halloran (2008) show that we can estimate the indirect effect with the difference in the average outcome across untreated staff members from the treated and untreated departments.
- In Duflo and Saez (2003), the average fair attendance rate is 0.280 among treated staff members from treated departments, 0.151 among untreated staff members from treated departments, and 0.049 among untreated staff members from untreated departments.
- The direct effect estimate is $0.280 - 0.151 = 0.129$.
- The estimated indirect effect on untreated staff members is $0.151 - 0.049 = 0.102$.

Partial interference

- Hudgens and Halloran (2008) show that we can estimate the indirect effect with the difference in the average outcome across untreated staff members from the treated and untreated departments.
- In Duflo and Saez (2003), the average fair attendance rate is 0.280 among treated staff members from treated departments, 0.151 among untreated staff members from treated departments, and 0.049 among untreated staff members from untreated departments.
- The direct effect estimate is $0.280 - 0.151 = 0.129$.
- The estimated indirect effect on untreated staff members is $0.151 - 0.049 = 0.102$.
- We need a large number of departments to estimate the indirect effect precisely.

Application II

- Paluck, Shepherd, and Aronow (2016) examine the consequences of an anti-conflict campaign across 56 schools in New Jersey.

Application II

- Paluck, Shepherd, and Aronow (2016) examine the consequences of an anti-conflict campaign across 56 schools in New Jersey.
- 28 schools were assigned into the treatment group.

Application II

- Paluck, Shepherd, and Aronow (2016) examine the consequences of an anti-conflict campaign across 56 schools in New Jersey.
- 28 schools were assigned into the treatment group.
- In each treated school, treatment was then assigned at the individual level.

Application II

- Paluck, Shepherd, and Aronow (2016) examine the consequences of an anti-conflict campaign across 56 schools in New Jersey.
- 28 schools were assigned into the treatment group.
- In each treated school, treatment was then assigned at the individual level.
- Treated students were invited to participate in a bi-weekly meeting to discuss the consequences of conflicts and behavioral strategies.

Application II

- The intervention lasted for an academic year.

Application II

- The intervention lasted for an academic year.
- The authors measure the outcome in various ways (e.g., whether the student wears a wristband signaling commitment to anti-conflict norms.)

Application II

- The intervention lasted for an academic year.
- The authors measure the outcome in various ways (e.g., whether the student wears a wristband signaling commitment to anti-conflict norms.)
- Untreated students may learn about the content of the meetings from their friends.

Application II

- The intervention lasted for an academic year.
- The authors measure the outcome in various ways (e.g., whether the student wears a wristband signaling commitment to anti-conflict norms.)
- Untreated students may learn about the content of the meetings from their friends.
- Treated students may reinforce each other's commitment to anti-conflict norms.

Application II

- We may still assume partial interference and estimate the indirect effect as before.

Application II

- We may still assume partial interference and estimate the indirect effect as before.
- But a school can be big and the indirect effect estimate may not be meaningful.

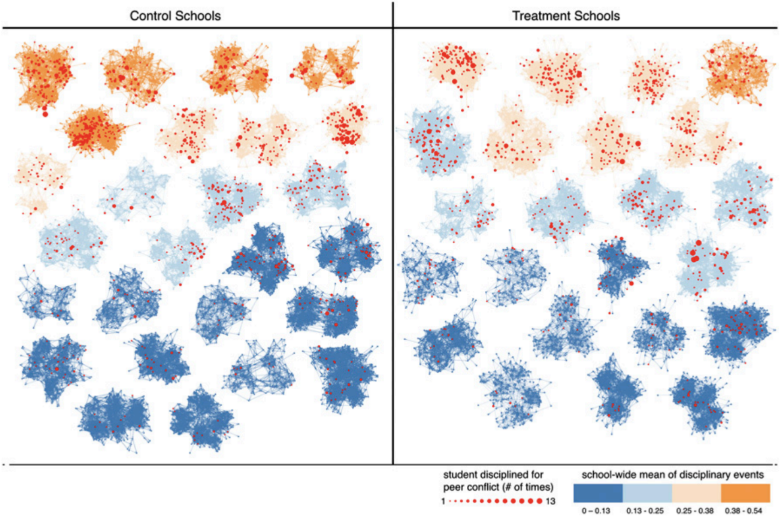
Application II

- We may still assume partial interference and estimate the indirect effect as before.
- But a school can be big and the indirect effect estimate may not be meaningful.
- Instead, the authors collect information on the social network among the students.

Application II

- We may still assume partial interference and estimate the indirect effect as before.
- But a school can be big and the indirect effect estimate may not be meaningful.
- Instead, the authors collect information on the social network among the students.
- They ask the students to list all their friends in the school before the treatment.

Application II



Exposure mapping

- This information allows the authors to adopt the approach of "exposure mapping" developed by Aronow and Samii (2017).

Exposure mapping

- This information allows the authors to adopt the approach of "exposure mapping" developed by Aronow and Samii (2017).
- We assume that there exists a mapping from treatment assignment to the actual treatment exposure.

Exposure mapping

- This information allows the authors to adopt the approach of "exposure mapping" developed by Aronow and Samii (2017).
- We assume that there exists a mapping from treatment assignment to the actual treatment exposure.
- If only one's outcome is only affected by her direct friends in the social network, we can measure the level of exposure with the proportion of treated friends.

Exposure mapping

- We estimate the effect generated by exposure.

Exposure mapping

- We estimate the effect generated by exposure.
- It is a standard causal inference problem with a transformed treatment variable.

Exposure mapping

- We estimate the effect generated by exposure.
- It is a standard causal inference problem with a transformed treatment variable.
- This method requires the correct specification of the exposure mapping.

Exposure mapping

- Aronow and Samii (2017) consider a simpler transformation.

Exposure mapping

- Aronow and Samii (2017) consider a simpler transformation.
- There are five different levels of exposure: $(1, 1, 1)$ (direct $+$ indirect exposure), $(1, 0, 1)$ (isolated indirect exposure), $(0, 1, 1)$ (indirect exposure), $(0, 0, 1)$ (school exposure), and $(0, 0, 0)$ (no exposure).

Exposure mapping

- Aronow and Samii (2017) consider a simpler transformation.
- There are five different levels of exposure: $(1, 1, 1)$ (direct + indirect exposure), $(1, 0, 1)$ (isolated indirect exposure), $(0, 1, 1)$ (indirect exposure), $(0, 0, 1)$ (school exposure), and $(0, 0, 0)$ (no exposure).
- We use no exposure as the benchmark.

Exposure mapping

Estimator	**Estimand**	**Estimate**	**S.E.**	**95% CI**
HT	$\tau(d_{001}, d_{000})$	0.057	0.062	$(-0.065, 0.179)$
	$\tau(d_{011}, d_{000})$	0.154	0.029	$(0.097, 0.211)$
	$\tau(d_{101}, d_{000})$	0.305	0.141	$(0.029, 0.581)$
	$\tau(d_{111}, d_{000})$	0.299	0.020	$(0.260, 0.338)$
Hajek	$\tau(d_{001}, d_{000})$	0.058	0.064	$(-0.067, 0.183)$
	$\tau(d_{011}, d_{000})$	0.154	0.037	$(0.081, 0.227)$
	$\tau(d_{101}, d_{000})$	0.292	0.123	$(0.051, 0.533)$
	$\tau(d_{111}, d_{000})$	0.307	0.049	$(0.211, 0.403)$
WLS	$\tau(d_{001}, d_{000})$	0.056	0.066	$(-0.072, 0.186)$
	$\tau(d_{011}, d_{000})$	0.156	0.037	$(0.083, 0.229)$
	$\tau(d_{101}, d_{000})$	0.295	0.124	$(0.050, 0.536)$
	$\tau(d_{111}, d_{000})$	0.306	0.049	$(0.212, 0.404)$

HT = Horvitz–Thompson estimator with conservative variance estimator.
Hajek = Hajek estimator with linearized variance estimator.
WLS = Least squares weighted by exposure probabilities with covariate adjustment for network degree and linearized variance estimator.
S.E. = Estimated standard error; CI = Normal approximation confidence interval.

Application III

- Stokes (2016): political consequences of building wind turbines in Ontario, Canada.

Application III

- Stokes (2016): political consequences of building wind turbines in Ontario, Canada.
- 6186 precincts over three elections: 2003, 2007, and 2011.

Application III

- Stokes (2016): political consequences of building wind turbines in Ontario, Canada.
- 6186 precincts over three elections: 2003, 2007, and 2011.
- Wind turbine constructions started after 2003.

Application III

- Stokes (2016): political consequences of building wind turbines in Ontario, Canada.
- 6186 precincts over three elections: 2003, 2007, and 2011.
- Wind turbine constructions started after 2003.
- Outcomes: turnout rate and the Liberal Party's vote share.

Application III

- Stokes (2016): political consequences of building wind turbines in Ontario, Canada.
- 6186 precincts over three elections: 2003, 2007, and 2011.
- Wind turbine constructions started after 2003.
- Outcomes: turnout rate and the Liberal Party's vote share.
- How to estimate the spillover effects?

Application III

- Stokes (2016): political consequences of building wind turbines in Ontario, Canada.
- 6186 precincts over three elections: 2003, 2007, and 2011.
- Wind turbine constructions started after 2003.
- Outcomes: turnout rate and the Liberal Party's vote share.
- How to estimate the spillover effects?
- Can we assume partial interference?

Application III

- Stokes (2016): political consequences of building wind turbines in Ontario, Canada.
- 6186 precincts over three elections: 2003, 2007, and 2011.
- Wind turbine constructions started after 2003.
- Outcomes: turnout rate and the Liberal Party's vote share.
- How to estimate the spillover effects?
- Can we assume partial interference?
- Does any exposure mapping exist?

Application III

- Stokes (2016): political consequences of building wind turbines in Ontario, Canada.
- 6186 precincts over three elections: 2003, 2007, and 2011.
- Wind turbine constructions started after 2003.
- Outcomes: turnout rate and the Liberal Party's vote share.
- How to estimate the spillover effects?
- Can we assume partial interference?
- Does any exposure mapping exist?
- We need new methods for this scenario.

Define the indirect effect

- Use the previous example

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

Define the indirect effect

- Use the previous example

Treatment status	Prob	Ye	Jiawei
(1, 1)	0.25	8	6
(1, 0)	0.25	7	5
(0, 1)	0.25	4	5
(0, 0)	0.25	2	3

$$\tau_{Ye;Jiawei} = 0.5 * 1 + 0.5 * 2 = 1.5$$
$$\tau_{Jiawei;Ye} = 0.5 * 1 + 0.5 * 2 = 1.5$$

Define the indirect effect

- We can similarly construct the indirect effect generated by unit j on any other unit i.

Define the indirect effect

- We can similarly construct the indirect effect generated by unit j on any other unit i.
- The problem is that we do not want to simply take an average.

Define the indirect effect

- We can similarly construct the indirect effect generated by unit j on any other unit i.
- The problem is that we do not want to simply take an average.
- It is hard to identify such an estimand.

Define the indirect effect

- We can similarly construct the indirect effect generated by unit j on any other unit i.
- The problem is that we do not want to simply take an average.
- It is hard to identify such an estimand.
- Instead, we can aggregate $\tau_{i;j}$ over a specific set of neighbors of j's, $\Omega_j(d)$, where d is a proximity metric.

Define the indirect effect

- We can similarly construct the indirect effect generated by unit j on any other unit i.
- The problem is that we do not want to simply take an average.
- It is hard to identify such an estimand.
- Instead, we can aggregate $\tau_{i;j}$ over a specific set of neighbors of j's, $\Omega_j(d)$, where d is a proximity metric.
- This is a linear mapping from $\mathcal{R}^N$ to $\mathcal{R}$ that captures the influence of a unit on the others.

Define the indirect effect

- We can similarly construct the indirect effect generated by unit j on any other unit i.
- The problem is that we do not want to simply take an average.
- It is hard to identify such an estimand.
- Instead, we can aggregate $\tau_{i;j}$ over a specific set of neighbors of j's, $\Omega_j(d)$, where d is a proximity metric.
- This is a linear mapping from $\mathcal{R}^N$ to $\mathcal{R}$ that captures the influence of a unit on the others.
- The form of $\Omega_j(d)$ is decided by the purpose of the study.

Define the indirect effect

- We can similarly construct the indirect effect generated by unit j on any other unit i.
- The problem is that we do not want to simply take an average.
- It is hard to identify such an estimand.
- Instead, we can aggregate $\tau_{i;j}$ over a specific set of neighbors of j's, $\Omega_j(d)$, where d is a proximity metric.
- This is a linear mapping from $\mathcal{R}^N$ to $\mathcal{R}$ that captures the influence of a unit on the others.
- The form of $\Omega_j(d)$ is decided by the purpose of the study.
- Wang et al. (2025) propose the "circle average:" $\Omega_j(d) = \{i : d_{ij} = d\}$.

Define the indirect effect

- We can similarly construct the indirect effect generated by unit j on any other unit i.
- The problem is that we do not want to simply take an average.
- It is hard to identify such an estimand.
- Instead, we can aggregate $\tau_{i;j}$ over a specific set of neighbors of j's, $\Omega_j(d)$, where d is a proximity metric.
- This is a linear mapping from $\mathcal{R}^N$ to $\mathcal{R}$ that captures the influence of a unit on the others.
- The form of $\Omega_j(d)$ is decided by the purpose of the study.
- Wang et al. (2025) propose the "circle average:" $\Omega_j(d) = \{i : d_{ij} = d\}$.
- Then, we can define the average marginalized effect generated by unit j on its neighbors in $\Omega_j(d)$:

$$\tau_j(d) = \frac{\sum_{i=1}^{N} \mathbf{1}\{i \in \Omega_j(d)\}\tau_{i;j}}{\sum_{i=1}^{N} \mathbf{1}\{i \in \Omega_j(d)\}}.$$

Estimate the indirect effect

- Finally, our estimand is obtained by averaging over $\tau_j(d)$:

$$\tau(d) = \frac{1}{N} \sum_{j=1}^{N} \tau_j(d).$$

Estimate the indirect effect

- Finally, our estimand is obtained by averaging over $\tau_j(d)$:

$$\tau(d) = \frac{1}{N} \sum_{j=1}^{N} \tau_j(d).$$

- We can identify this quantity by constructing the observed outcome's circle average:

$$\mu_j(d) = \frac{\sum_{i=1}^{N} \mathbf{1}\{i \in \Omega_j(d)\} Y_i}{\sum_{i=1}^{N} \mathbf{1}\{i \in \Omega_j(d)\}}.$$

Estimate the indirect effect

- Finally, our estimand is obtained by averaging over $\tau_j(d)$:

$$\tau(d) = \frac{1}{N} \sum_{j=1}^{N} \tau_j(d).$$

- We can identify this quantity by constructing the observed outcome's circle average:

$$\mu_j(d) = \frac{\sum_{i=1}^{N} \mathbf{1}\{i \in \Omega_j(d)\} Y_i}{\sum_{i=1}^{N} \mathbf{1}\{i \in \Omega_j(d)\}}.$$

- Then, we can show that

$$\tau(d) = \mathrm{E}\left[\frac{1}{N} \sum_{j=1}^{N} \frac{Z_j \mu_j(d)}{p_j} - \frac{1}{N} \sum_{j=1}^{N} \frac{(1 - Z_j)\mu_j(d)}{1 - p_j}\right].$$

Estimate the indirect effect

- We can obtain the estimate $\hat{\tau}(d)$ by regressing $\mu_j(d)$ on Z_j, with the weight
$$w_j = \frac{Z_j}{p_j} + \frac{1 - Z_j}{1 - p_j}.$$

Estimate the indirect effect

- We can obtain the estimate $\hat{\tau}(d)$ by regressing $\mu_j(d)$ on Z_j, with the weight
$$w_j = \frac{Z_j}{p_j} + \frac{1 - Z_j}{1 - p_j}.$$
- If the degree of dependence across units isn't growing too fast with the sample size, $\sqrt{N}(\hat{\tau}(d) - \tau(d))$ converges to a normal distribution centered around 0.

Estimate the indirect effect

- We can obtain the estimate $\hat{\tau}(d)$ by regressing $\mu_j(d)$ on Z_j, with the weight
$$w_j = \frac{Z_j}{p_j} + \frac{1 - Z_j}{1 - p_j}.$$
- If the degree of dependence across units isn't growing too fast with the sample size, $\sqrt{N}(\hat{\tau}(d) - \tau(d))$ converges to a normal distribution centered around 0.
- By examining $\hat{\tau}(d)$ at different values of d, we can see how spillover effects vary with proximity.

Application III

- We assume that the construction of wind turbines is independent to the potential outcomes conditional on covariates and lagged outcomes.

Application III

- We assume that the construction of wind turbines is independent to the potential outcomes conditional on covariates and lagged outcomes.
- This is known as “sequential ignorability” in biostatistics.

Application III

- We assume that the construction of wind turbines is independent to the potential outcomes conditional on covariates and lagged outcomes.
- This is known as "sequential ignorability" in biostatistics.
- We focus on the effect of two treatment histories, $(1, 1)$ and $(0, 1)$, relative to $(0, 0)$, on the outcome in 2011.

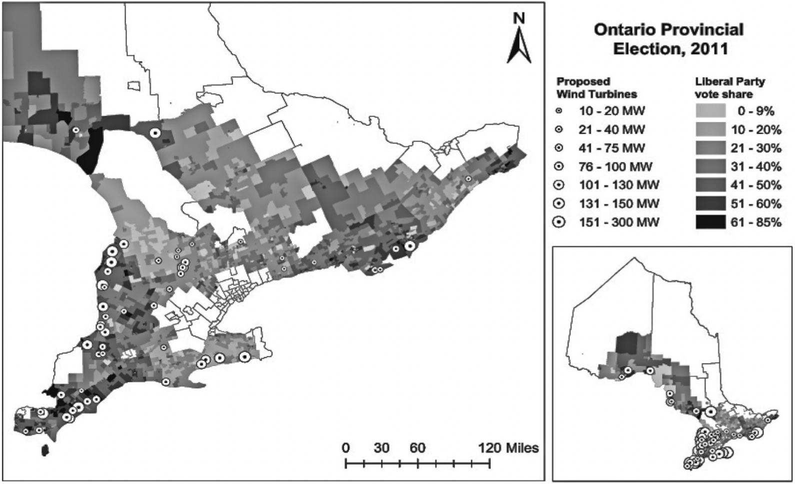
Application III

- We assume that the construction of wind turbines is independent to the potential outcomes conditional on covariates and lagged outcomes.
- This is known as "sequential ignorability" in biostatistics.
- We focus on the effect of two treatment histories, $(1, 1)$ and $(0, 1)$, relative to $(0, 0)$, on the outcome in 2011.
- Propensity scores are estimated via logistic regression with the lagged outcomes and a quadratic function of the geographic coordinates.

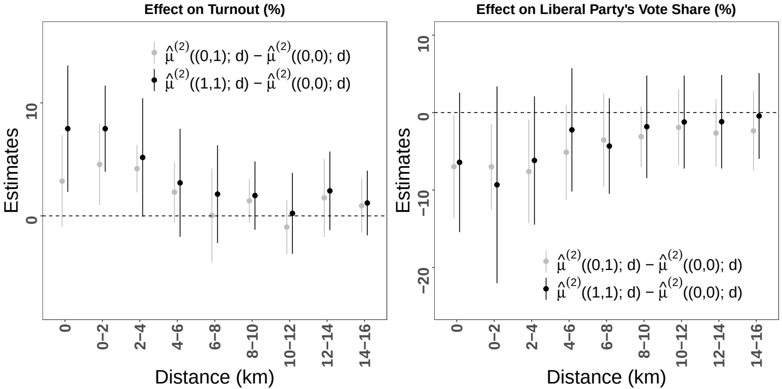
Application III

- We assume that the construction of wind turbines is independent to the potential outcomes conditional on covariates and lagged outcomes.
- This is known as "sequential ignorability" in biostatistics.
- We focus on the effect of two treatment histories, $(1, 1)$ and $(0, 1)$, relative to $(0, 0)$, on the outcome in 2011.
- Propensity scores are estimated via logistic regression with the lagged outcomes and a quadratic function of the geographic coordinates.
- We construct "donuts" with the radius of 2 km around each precinct.

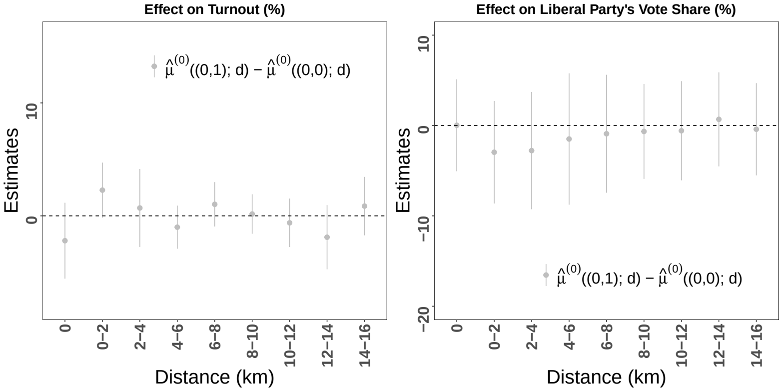
Application III



Application III



Application III



References I

Aronow, Peter M, and Cyrus Samii. 2017. "Estimating Average Causal Effects Under General Interference, with Application to a Social Network Experiment." *The Annals of Applied Statistics* 11 (4): 1912–47.

Duflo, Esther, and Emmanuel Saez. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment." *The Quarterly Journal of Economics* 118 (3): 815–42.

Hudgens, Michael, and M. Elizabeth Halloran. 2008. "Toward Causal Inference with Interference." *Journal of the American Statistical Association* 103 (482): 832–42.

Paluck, Elizabeth Levy, Hana Shepherd, and Peter M. Aronow. 2016. "Changing Climates of Conflict: A Social Network Experiment in 56 Schools." *Proceedings of the National Academy of Science* 113 (3): 566–71.

References II

Sävje, Fredrik, Peter M Aronow, and Michael G Hudgens. 2021. "Average Treatment Effects in the Presence of Unknown Interference." *The Annals of Statistics* 49 (2): 673–701.

Stokes, Leah C. 2016. "Electoral Backlash Against Climate Policy: A Natural Experiment on Retrospective Voting and Local Resistance to Public Policy." *American Journal of Political Science* 60 (4): 958–74.

Wang, Ye, Cyrus Samii, Haoge Chang, and PM Aronow. 2025. "Design-Based Inference for Spatial Experiments Under Unknown Interference." *The Annals of Applied Statistics* 19 (1): 744–68.

Wang, Ye, and Stan Hok-Wui Wong. 2021. "Electoral Impacts of a Failed Uprising: Evidence from Hong Kong's Umbrella Movement." *Electoral Studies* 71: 102336.