Causal Graphs for Conditional Parallel Trends

A Primer

Michael C. Knaus
University of Tübingen

Henri Pfleiderer
University of Tübingen

Purpose of This Primer

Welcome to the primer for Knaus & Pfleiderer (2026)!

This is meant for those who are interested in the results from the main paper but have not been exposed to causal graphs so far.

Instead of the technical details, propositions, … this focuses on the main steps in constructing $\boldsymbol{\Delta}$-SWIGs and path blocking rules.

Outline

Part I — Difference-in-Differences

Part II — Introduction to Causal Graphs

Part III — Our work: Δ-Single-World Intervention Graphs

Part IV — Some results using Δ-Single-World Intervention Graphs

Part V — Multiple Time Periods

Part I: Difference-in-Differences

The Basic 2x2 DiD Setting

We observe units in two time periods: a **before** ($t = 0$) and an **after** ($t = 1$) period. Between the two periods, some units get treated ($D = 1$), others do not ($D = 0$).

Example	**Outcome Y**	**Before ($t = 0$)**	**After ($t = 1$)**
Minimum wage	Employment	Before reform	After reform
Divorce reform	Female suicide rate	Before law	After law
LLMs	Case resolutions/hour	Before LLM access	After LLM access

Question: How much of the change in Y from $t = 0$ to $t = 1$ is caused by the treatment – and not just by other things that happened over time?

Potential Outcomes

For each unit, define two **potential outcomes**:

$$Y_1(0) \quad \text{outcome in } t=1 \text{ if untreated} \qquad Y_1(1) \quad \text{outcome in } t=1 \text{ if treated}$$

We only ever observe **one** of these:

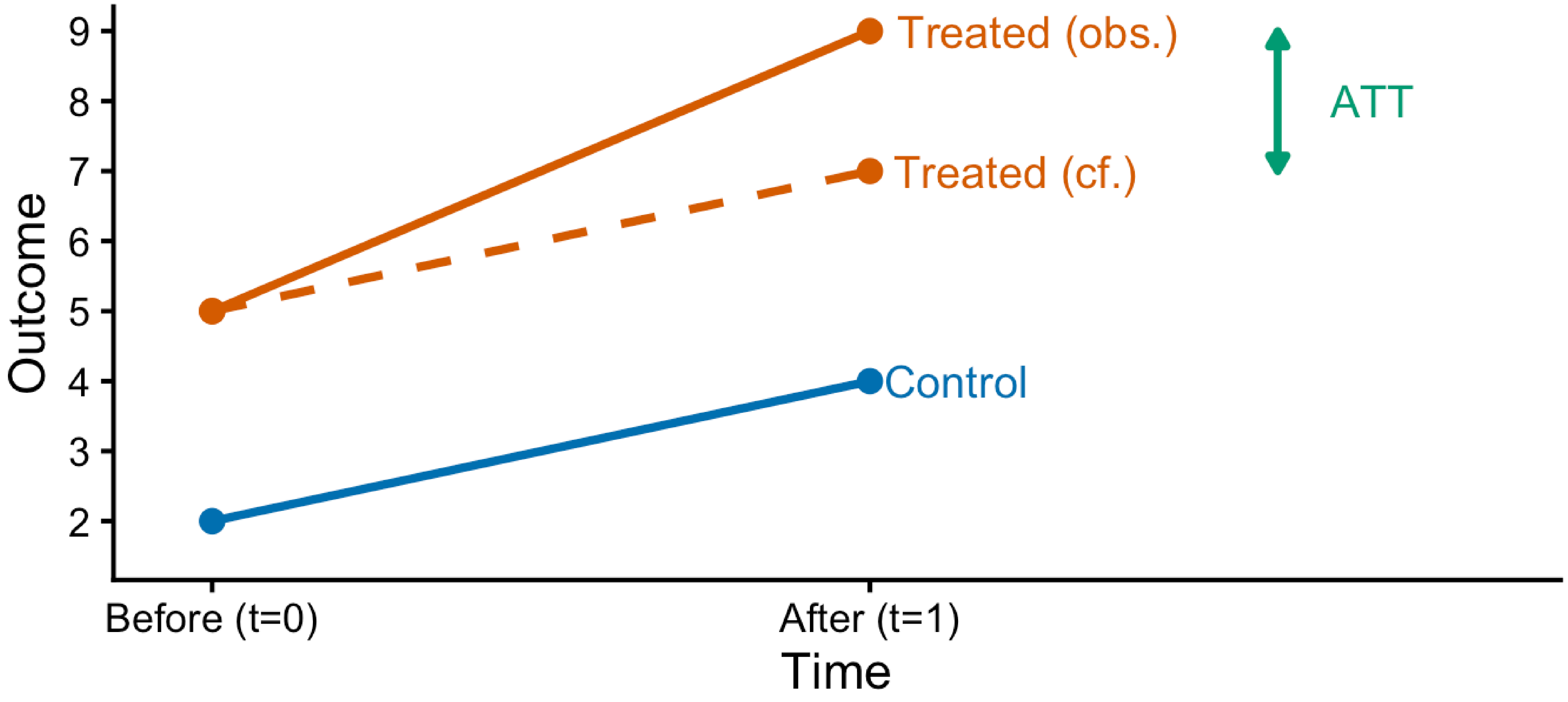
$$Y_1 = D \cdot Y_1(1) + (1-D) \cdot Y_1(0)$$

The **Average Treatment Effect on the Treated (ATT)** is what we want:

$$\text{ATT} = \mathbb{E}[Y_1(1) - Y_1(0) \mid D=1] = \underbrace{\mathbb{E}[Y_1(1) \mid D=1]}_{\text{observed (factual) outcome}} - \underbrace{\mathbb{E}[Y_1(0) \mid D=1]}_{\text{unobserved (counterfactual)}}$$

The problem: $\mathbb{E}[Y_1(0) \mid D=1]$ – what would have happened to treated units had they not been treated – is never observed.

The 2×2 Setup



Key idea: Use the trend of the untreated to impute the counterfactual for the treated.

When is this valid? Only if both groups would have trended the same way without the program — the **parallel trends assumption**.

The Parallel Trends Assumption

$$\mathbb{E}[Y_1(0) - Y_0(0) \mid D = 1] = \mathbb{E}[Y_1(0) - Y_0(0) \mid D = 0] \qquad \textbf{(PT)}$$

In words: *In the absence of the treatment, treated and untreated would have experienced the same change in the outcome.*

Is this plausible? The treated and untreated may differ in terms of some observable characteristics that also determine the trends.

We want to account for these observed characteristics.

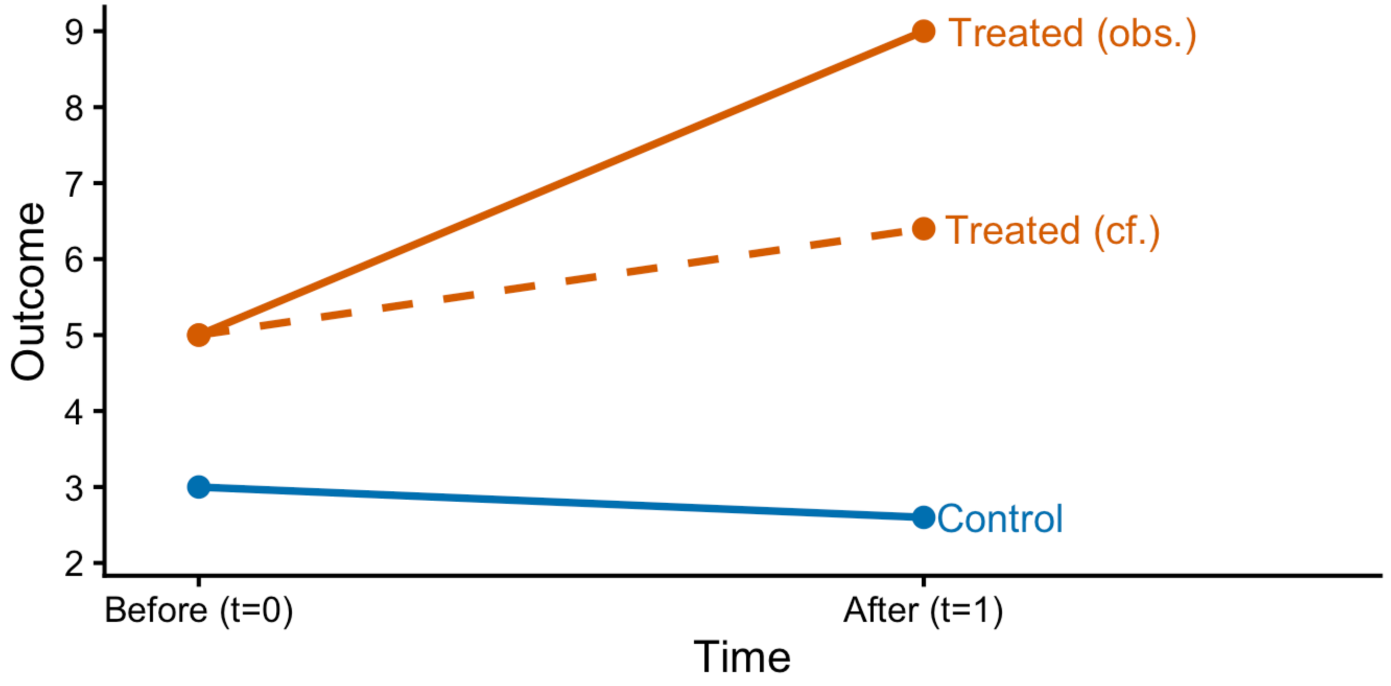
Stylized Example: Conditional Parallel Trends

Consider the example on the use of LLMs in customer service and the effect on productivity with the following adaptation:

- The workforce consists of Baby-Boomers ($X = 0$) and Non-Baby-Boomers ($X = 1$). Both groups are equally sized.
- Among the Non-Baby-Boomers, a larger share gets access to the LLM.
- Baby-Boomers and Non-Baby-Boomers also have different trends in outcomes without LLMs.

The following illustrates why parallel trends may fail here while conditional parallel trends holds.

Why Unconditional PT Fails



Treated and control units contain different shares of Baby Boomers ($X = 0$) – and trends differ by X.

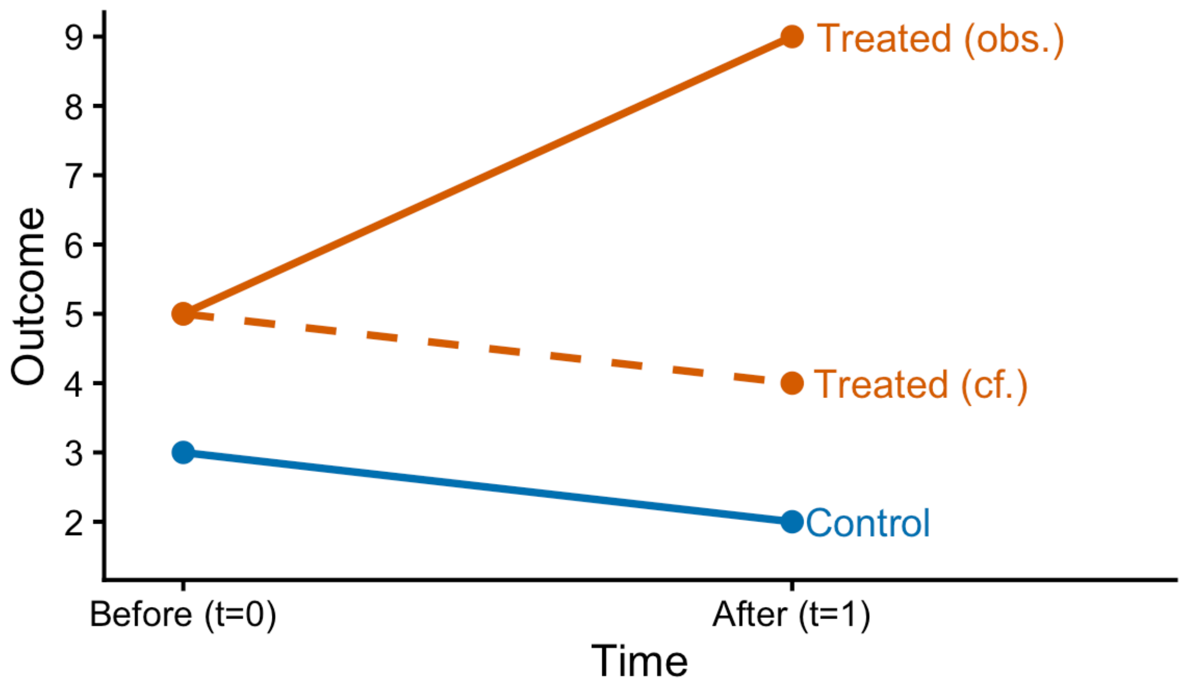
The aggregate counterfactual trend of the treated is steeper than the control trend.

→ Unconditional PT fails: simple DiD without control variables is biased.

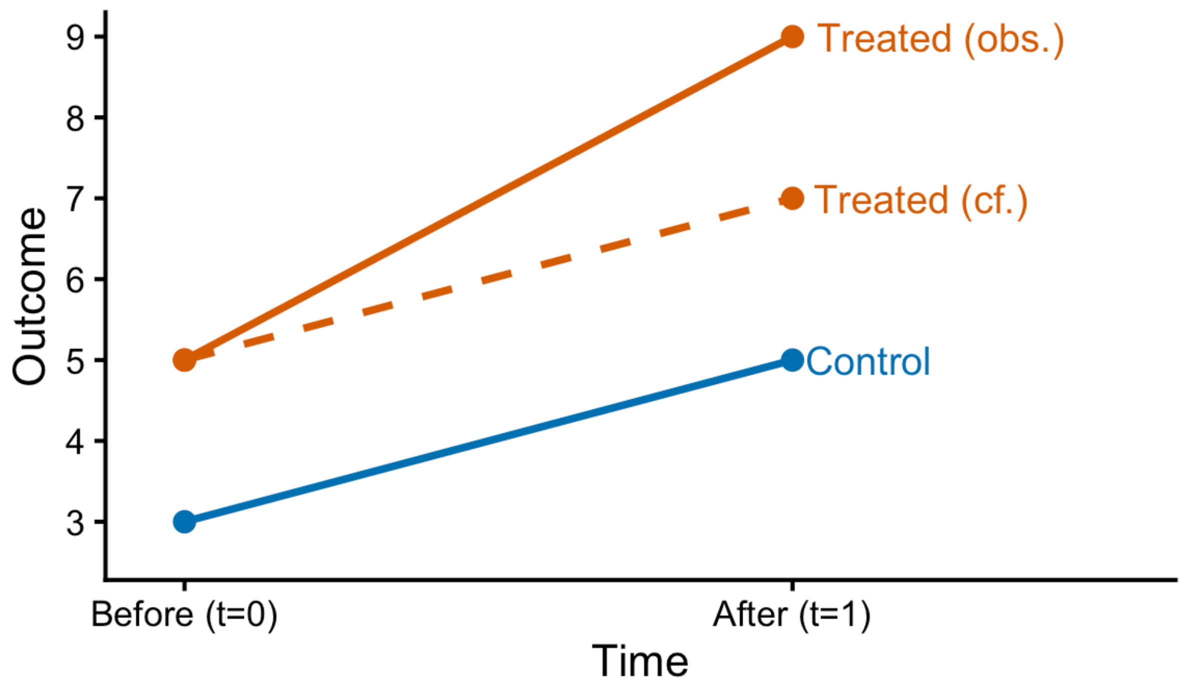
→ We need to condition on X.

CPT: Parallel Trends Within Each Group

$X = 0$: Baby Boomers



$X = 1$: Non-Baby Boomers



Within each group the dashed counterfactual runs parallel to the control — even though slopes differ between $X = 0$ and $X = 1$.

Conditional Parallel Trends

$$\mathbb{E}[Y_1(0) - Y_0(0) \mid D = 1, X] = \mathbb{E}[Y_1(0) - Y_0(0) \mid D = 0, X] \qquad \text{(CPT)}$$

Example: Among workers within the same generation, LLM users and non-LLM users would have had parallel productivity trends without the LLM.

New question: Which X should we condition on in practice?

- Age? Education? Previous wage?
- With time-varying covariates: from which period?

This is exactly what our graphical framework addresses — systematically and without algebra.

Identification Under CPT

Before we get to this, let's see how the CPT assumption helps to identify the ATT.

Goal: Express ATT using only observed quantities.

Step 1. Express the ATT in terms of trends:

$$\textbf{ATT} = \underbrace{\mathbb{E}[Y_1 - Y_0 \mid D = 1]}_{\textbf{observed trend of participants}} - \underbrace{\mathbb{E}[Y_1(0) - Y_0(0) \mid D = 1]}_{\textbf{unobserved counterfactual trend}}$$

Step 2. Apply the law of iterated expectations to the second term:

$$\mathbb{E}[Y_1(0) - Y_0(0) \mid D = 1] = \mathbb{E}[\mathbb{E}[Y_1(0) - Y_0(0) \mid X, D = 1] \mid D = 1]$$

Identification Under CPT (cont.)

Step 3. Apply CPT — replace conditioning on $D = 1$ with $D = 0$:

$$\mathbb{E}[\mathbb{E}[Y_1(0) - Y_0(0) \mid X, D = 1] \mid D = 1] = \mathbb{E}[\mathbb{E}[Y_1(0) - Y_0(0) \mid X, D = 0] \mid D = 1]$$

Step 4. For non-participants, $Y_t(0) = Y_t$ (they are never treated) — plug in observables:

$$\mathbb{E}[\mathbb{E}[Y_1(0) - Y_0(0) \mid X, D = 0] \mid D = 1] = \mathbb{E}[\mathbb{E}[Y_1 - Y_0 \mid X, D = 0] \mid D = 1]$$

Putting it together:

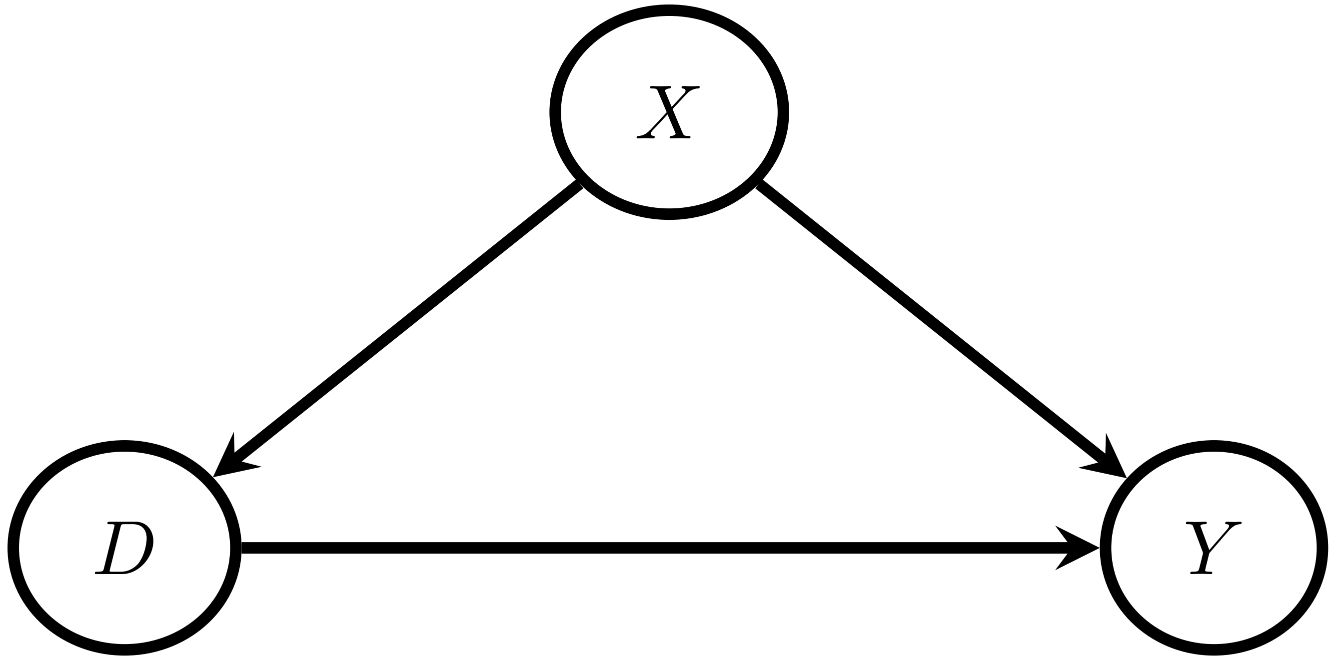
$$\mathrm{ATT} = \mathbb{E}[Y_1 - Y_0 \mid D = 1] - \mathbb{E}[\mathbb{E}[Y_1 - Y_0 \mid X, D = 0] \mid D = 1]$$

Effect of interest is now identified from observed data.

Part II: Causal Graphs

What Is a Directed Acyclic Graph (DAG)?

- DAGs represent the causal relations between different variables graphically
- **Nodes** = variables, **directed edges** ($\rightarrow$) = direct causal effects, **acyclic** — no variable can cause itself
- Each node has a **structural equation**: $V := f_V(\text{direct causes of } V,\, U_V)$



$$X := f_X(U_X)$$
$$D := f_D(X,\, U_D)$$
$$Y := f_Y(D,\, X,\, U_Y)$$

U_X, U_D, U_Y are independent background variables

Path Blocking: Chains and Forks

Two basic **path structures** transmit association and can be blocked by **conditioning**:

Chain: $A \to B \to C$

A affects C through another variable B. Conditioning on B **blocks** the path.

Fork: $A \leftarrow B \to C$

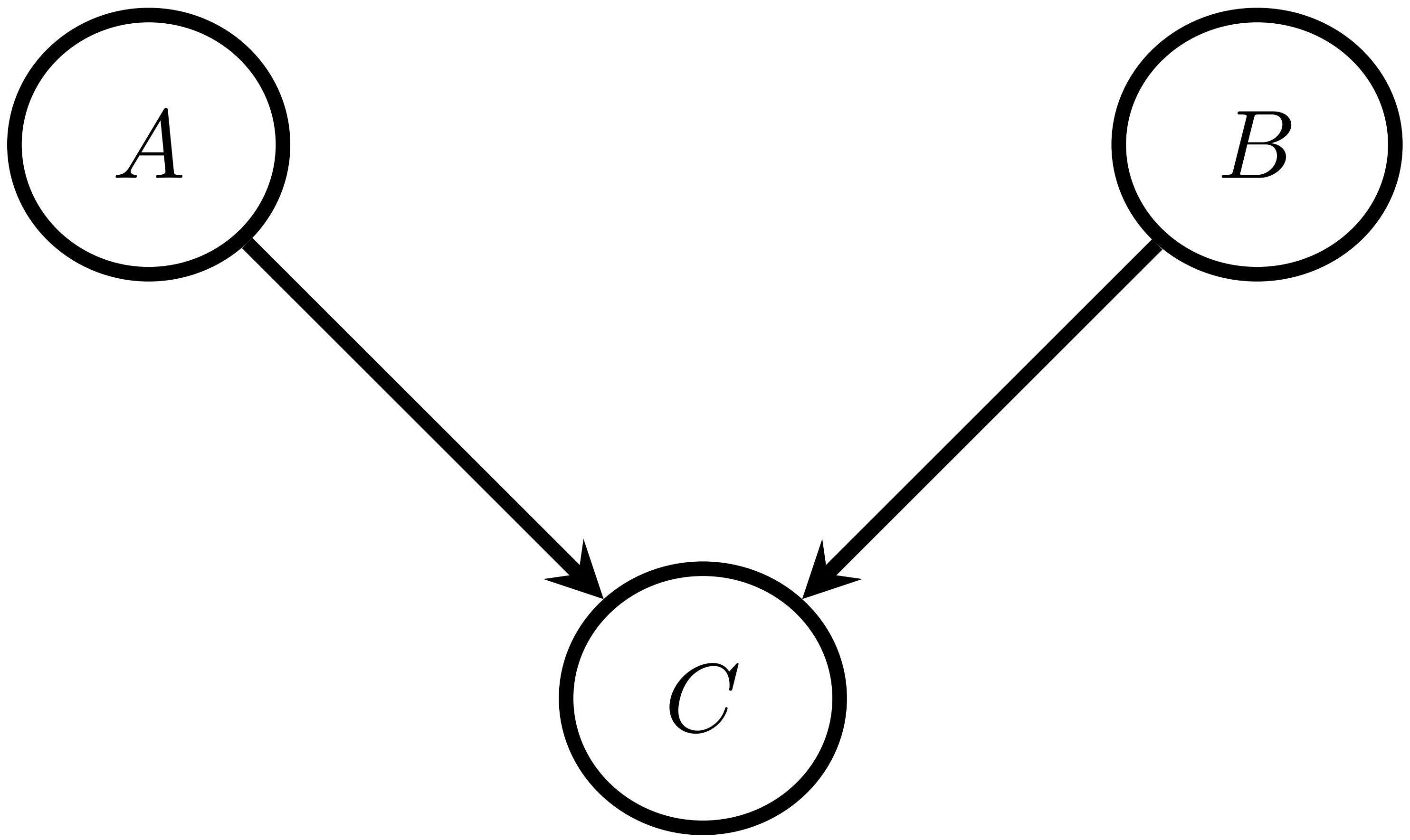
B causes both A and C – a **confounder**. Conditioning on B **blocks** the path.

Both structures are **open** by default — they transmit association between A and C.

They are **blocked** by conditioning on the middle node B.

Path Blocking: Colliders

Collider: $A \to C \leftarrow B$ – two arrows pointing *into* the same node.



- The path between A and B is **blocked**
- Conditioning on C **opens** the path

Path Blocking and Conditional Independence

A path is blocked by Z if it contains:

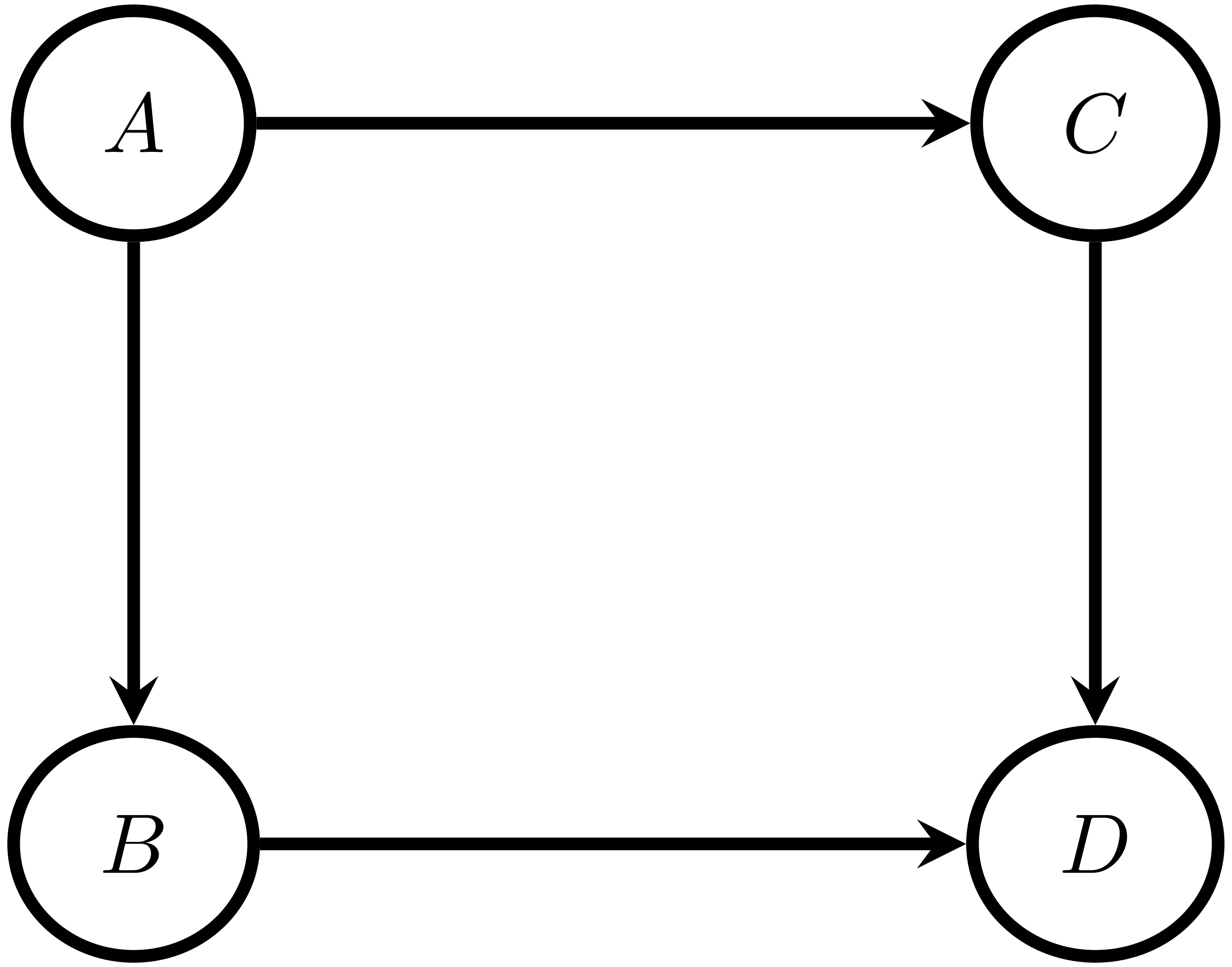
Structure	**Blocked when**
Chain $A \to C \to B$ or fork $A \leftarrow C \to B$	$C \in Z$
Collider $A \to C \leftarrow B$	$C \notin Z$ (and no descendant of C in Z)

If all paths between A and B are blocked by Z:

$$A \perp\!\!\!\perp B \mid Z$$

They are **conditionally independent** given Z — no information about A in B once we know Z.

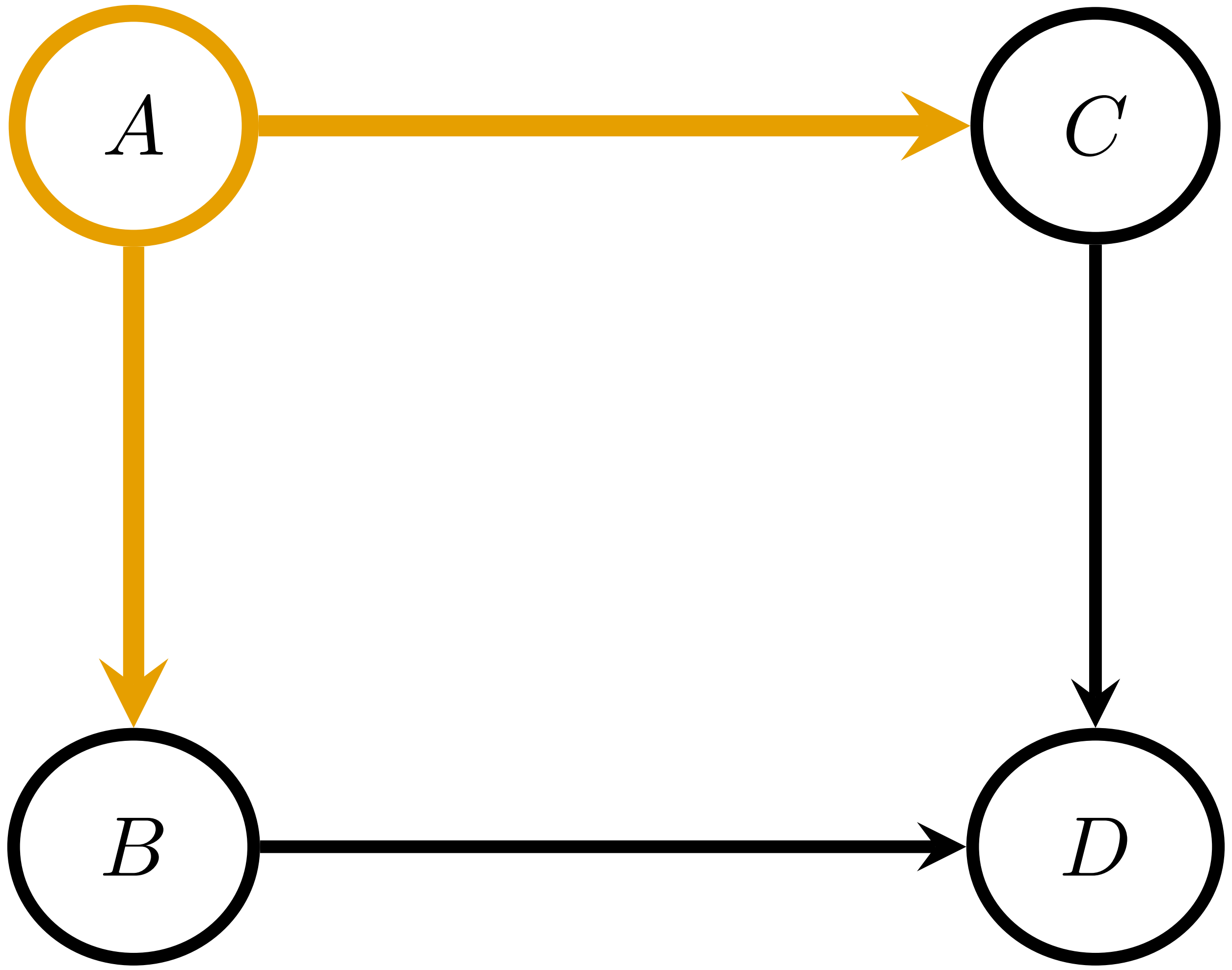
Path Blocking and Conditional Independence: Example



Conditional independencies:

- $B \perp\!\!\!\perp C \mid A$ – fork $B \leftarrow A \rightarrow C$ blocked; collider $B \rightarrow D \leftarrow C$ also blocked (D not conditioned on)

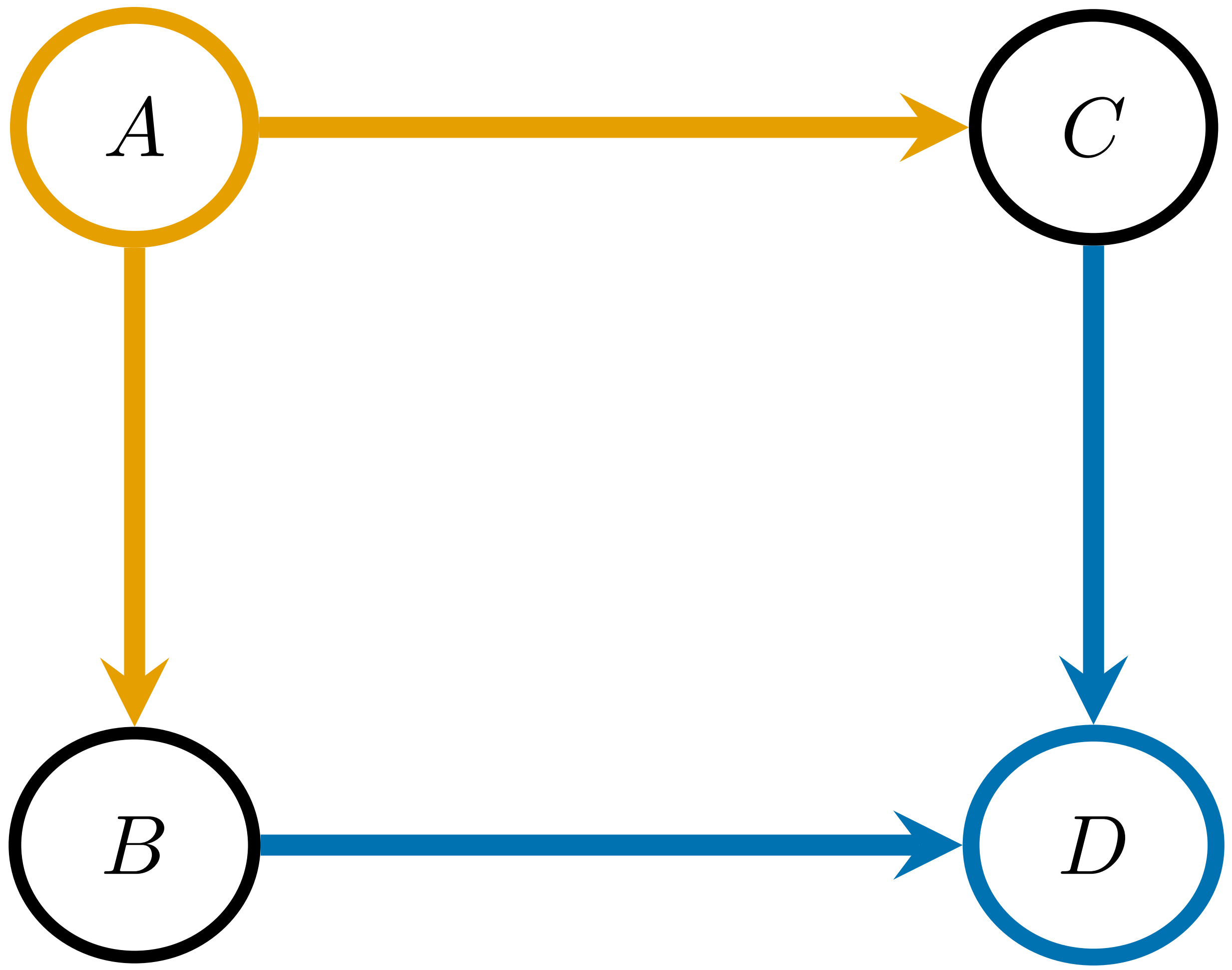
Path Blocking and Conditional Independence: Example



Conditional independencies:

- $B \perp\!\!\!\perp C \mid A$ – fork $B \leftarrow A \rightarrow C$ blocked; collider $B \rightarrow D \leftarrow C$ also blocked (D not conditioned on)

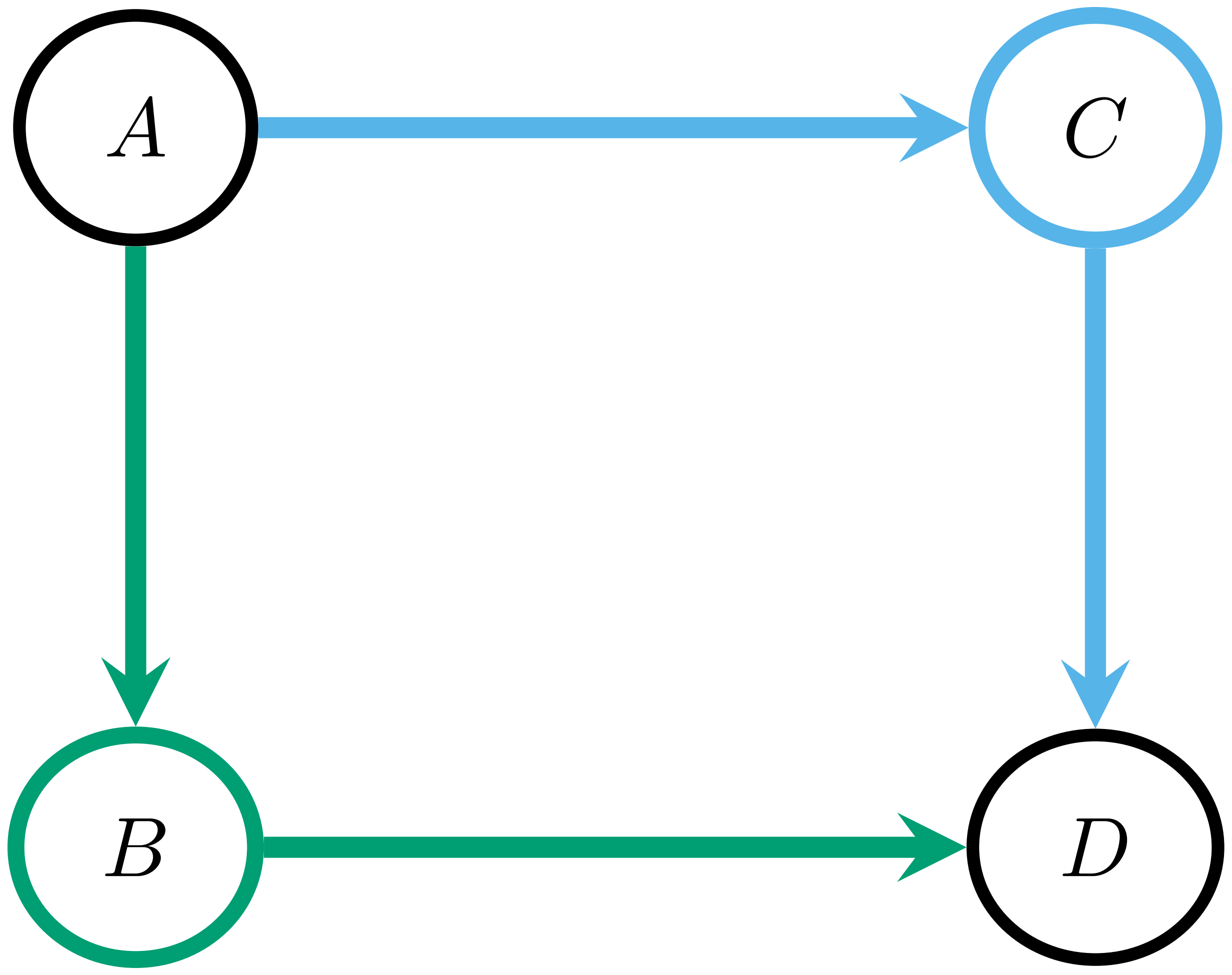
Path Blocking and Conditional Independence: Example



Conditional independencies:

- $B \perp\!\!\!\perp C \mid A$ – fork $B \leftarrow A \rightarrow C$ blocked; collider $B \rightarrow D \leftarrow C$ also blocked (D not conditioned on)

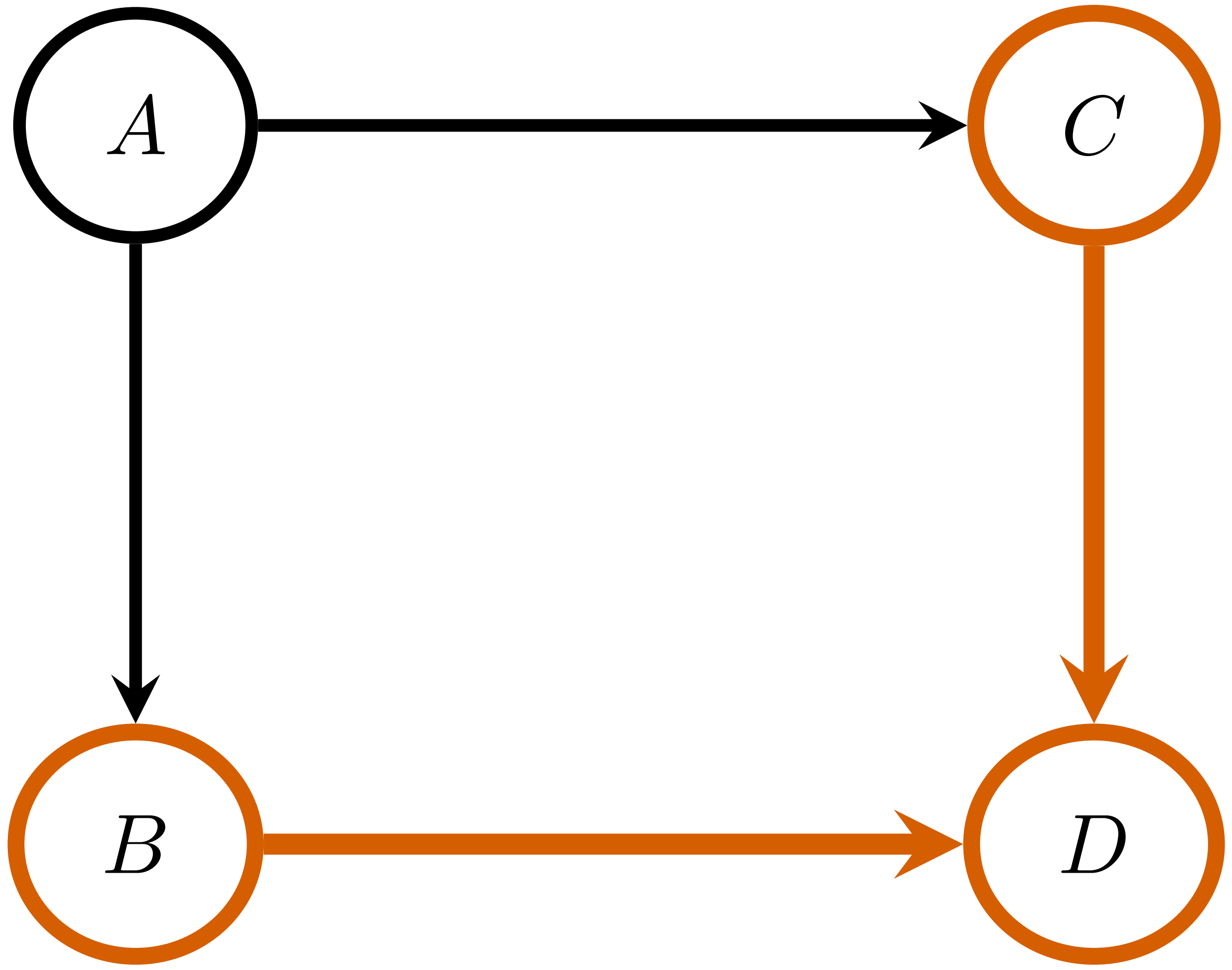
Path Blocking and Conditional Independence: Example



Conditional independencies:

- $B \perp\!\!\!\perp C \mid A$ – fork $B \leftarrow A \rightarrow C$ blocked; collider $B \rightarrow D \leftarrow C$ also blocked (D not conditioned on)
- $A \perp\!\!\!\perp D \mid B, C$ – both chains $A \rightarrow B \rightarrow D$ and $A \rightarrow C \rightarrow D$ blocked

Path Blocking and Conditional Independence: Example



Conditional independencies:

- $B \perp\!\!\!\perp C \mid A$ – fork $B \leftarrow A \rightarrow C$ blocked; collider $B \rightarrow D \leftarrow C$ also blocked (D not conditioned on)
- $A \perp\!\!\!\perp D \mid B, C$ – both chains $A \rightarrow B \rightarrow D$ and $A \rightarrow C \rightarrow D$ blocked

Collider warning:

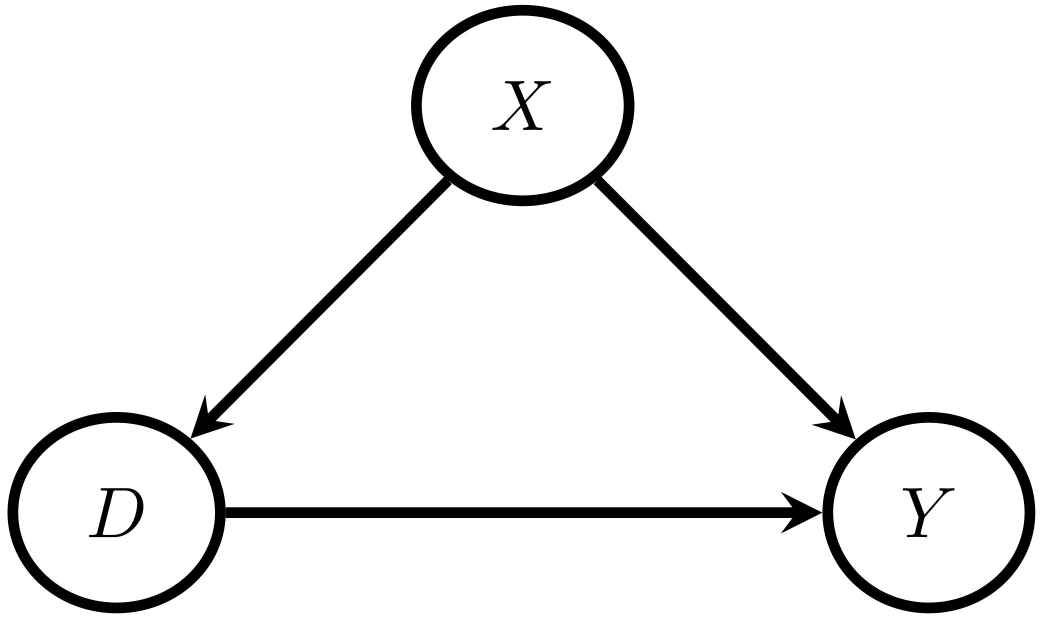
- B and C are independent given A
- Conditioning on D opens $B \rightarrow D \leftarrow C$ – suddenly B and C appear dependent!
- So, we do not have $B \perp\!\!\!\perp C \mid A, D$

From DAG to SWIG: Unconfoundedness

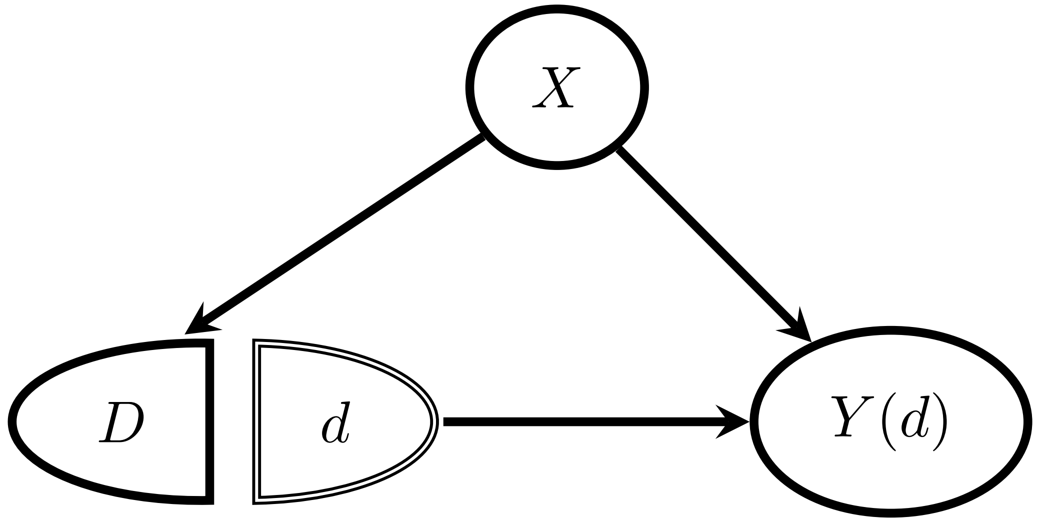
Shortcoming of plain DAGs: They represent Y, not potential outcomes $Y(d)$.

Solution: A **SWIG** (Single World Intervention Graph) — split the treatment node:

DAG



SWIG (intervene: set $D = d$)

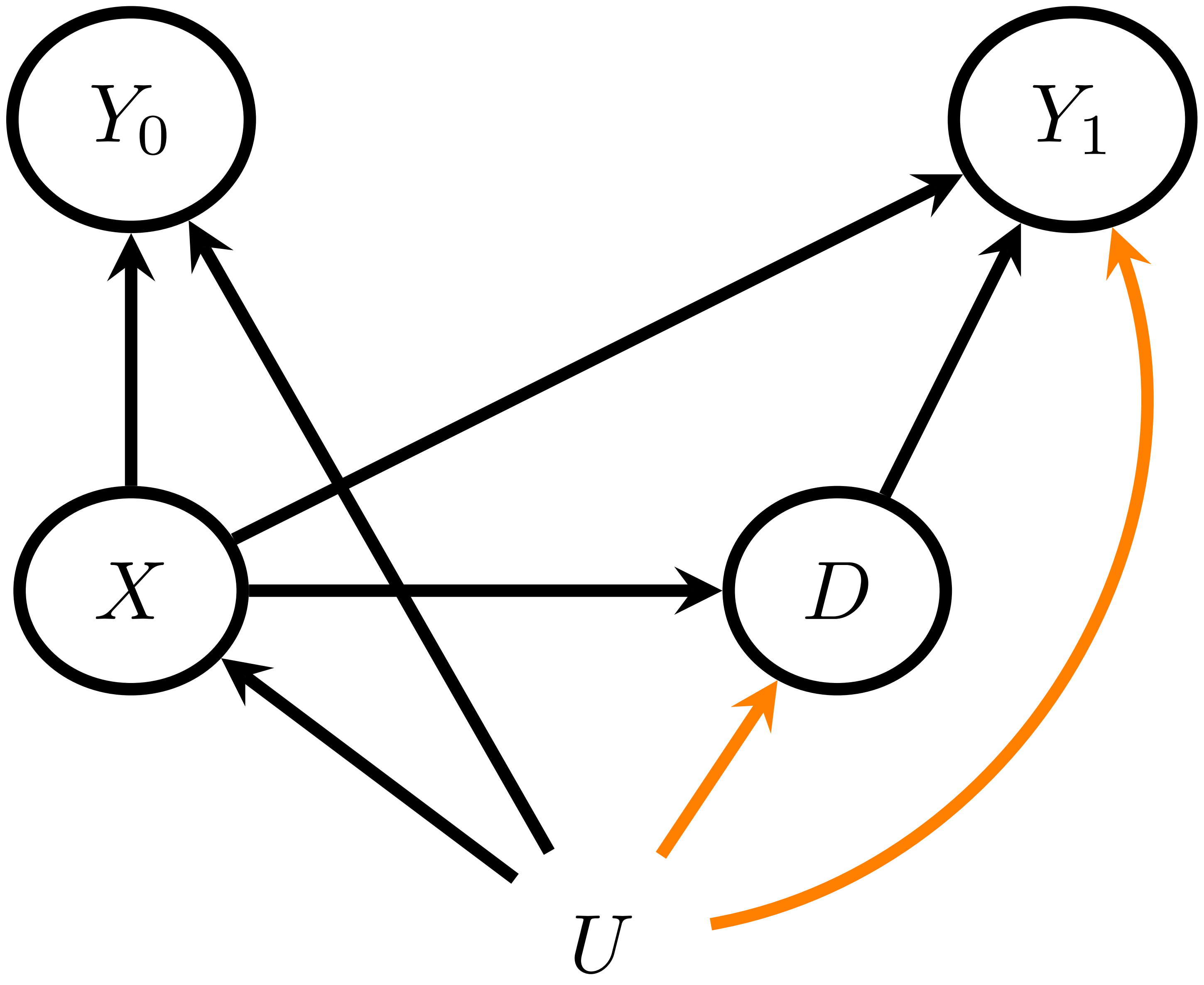


In the SWIG, all paths between D (the random variable) and $Y(d)$ are **blocked by** X. So: $Y(d) \perp\!\!\!\perp D \mid X$

This is the famous **unconfoundedness/conditional independence/exogeneity/ignorability/measured confounding/no omitted variables/no unmeasured confounding/selection-on-observables/…** assumption.

Part III: Δ-SWIGs

When Plain Path-Blocking Falls Short



U is an **unobserved confounder** that drives both D and outcomes

Even after conditioning on X…

the backdoor path $D \leftarrow U \rightarrow Y_1$ stays open

U is unobserved – X cannot block a path through it

We need a new tool: the Δ-SWIG

Single World Additive Separability

(Conditional) Parallel Trends are usually explicitly or implicitly justified by the following assumption on the structural equations for Y_0 and Y_1:

Assumption (Single World Additive Separability, SWAS)

$$Y_0 := f_{Y_0}(U, X, U_{Y_0}) = \underbrace{\alpha(U, X) + g_{Y_0}(X, U_{Y_0})}_{Y_0(0)}$$

$$Y_1 := f_{Y_1}(U, X, D, U_{Y_1}) = \underbrace{\alpha(U, X) + g_{Y_1}(X, U_{Y_1})}_{Y_1(0)} + D \cdot \underbrace{\tau(U, X, U_{Y_1})}_{Y_1(1) - Y_1(0)}$$

The important part becomes apparent when setting $D = 0$ in the structural equations. This is the operation that corresponds to the graphical node splitting:

$$Y_0(0) := \alpha(U, X) + g_{Y_0}(X, U_{Y_0})$$
$$Y_1(0) := \alpha(U, X) + g_{Y_1}(X, U_{Y_1})$$

U enters both potential outcomes via the same function $\alpha(U, X)$!

What Is $\alpha(U, X)$?

- $\alpha(U, X)$ is the same time-invariant function of time-invariant (unobserved and observed) variables in both time periods.
- It captures unobserved time-invariant individual-level heterogeneity.
- In the unconditional case, i.e., $X = \emptyset$, we can write $g_{Y_t}(U_{Y_t}) = \underbrace{\mathbb{E}[g_{Y_t}(U_{Y_t})]}_{\lambda_t} + \underbrace{g_{Y_t}(U_{Y_t}) - \mathbb{E}[g_{Y_t}(U_{Y_t})]}_{\varepsilon_t}$

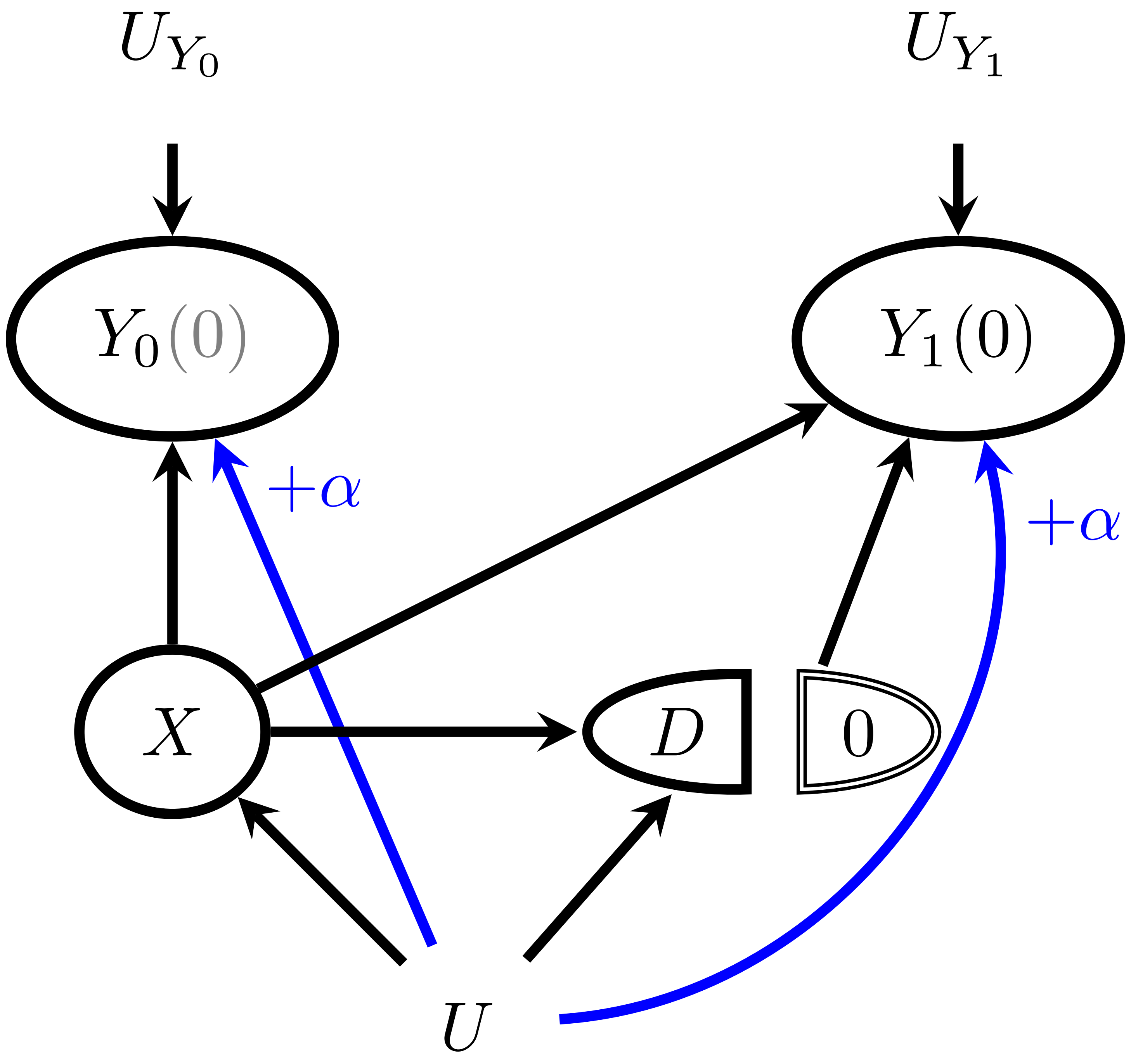
  such that

  $$Y_t(0) = \alpha(U) + \lambda_t + \varepsilon_t$$

  recovers the standard two-way fixed-effects structure.

The DiD SWIG

Apply the SWIG idea to the DiD setting: two time periods, a time-invariant confounder U, observed control X.



1. Split the treatment node.
2. Relabel outcomes.

We only care about the SWIG for $D = 0$, i.e., the "untreated world."

This is because: the Conditional Parallel Trends Assumption is an assumption on the untreated potential outcomes.

$$\mathbb{E}[Y_1(0) - Y_0(0) \mid D = 1, X] = \mathbb{E}[Y_1(0) - Y_0(0) \mid D = 0, X]$$

The grey potential outcome in $Y_0(0)$ is used to illustrate that relabeling it is optional.

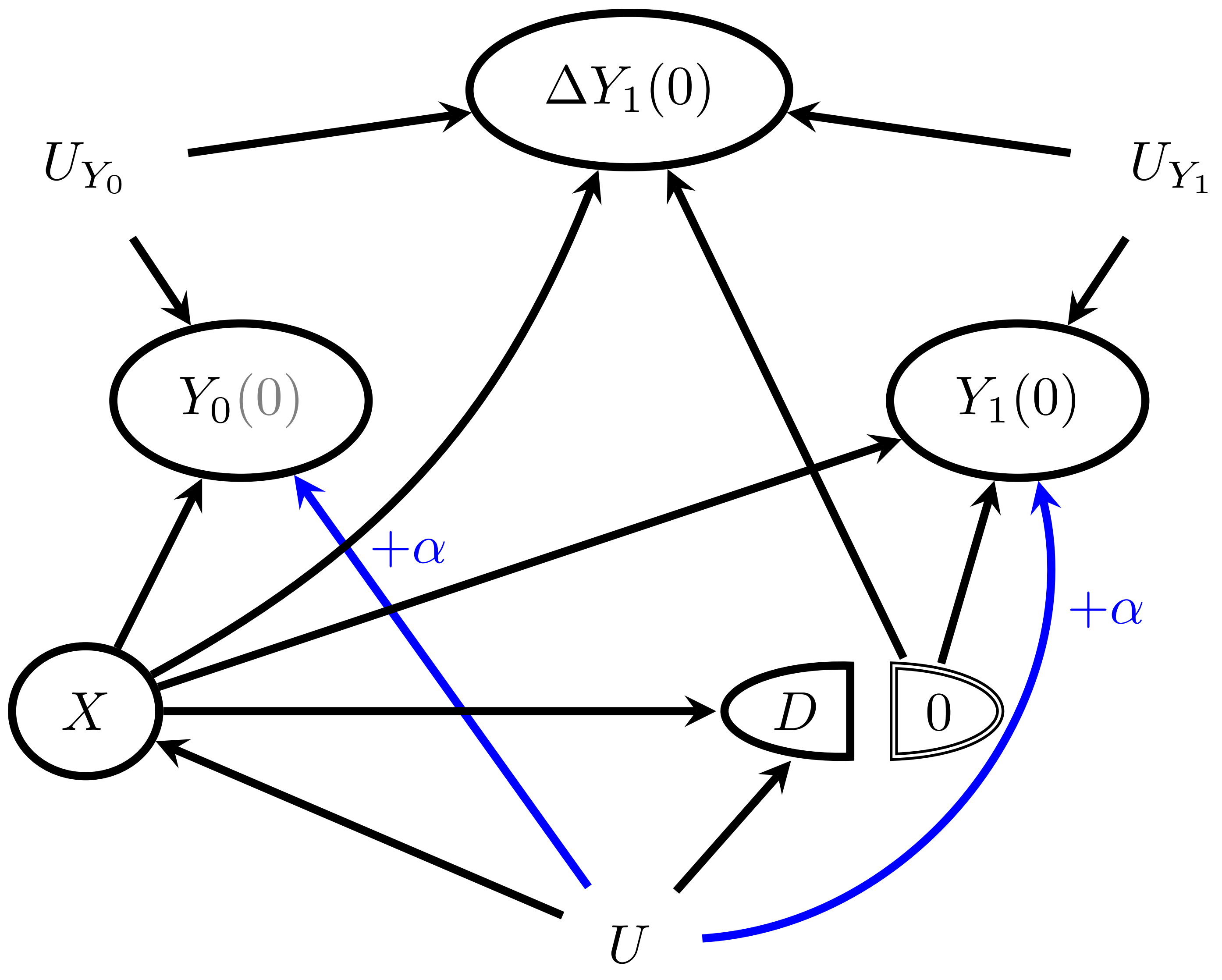
Building the $\boldsymbol{\Delta}$-SWIG

Add a new node $\Delta Y_1(0) \equiv Y_1(0) - Y_0(0)$ to the SWIG. It inherits all incoming arrows of $\boldsymbol{Y_1(0), Y_0(0)}$, except the arrows from $\boldsymbol{U}$ because, under SWAS, the $\boldsymbol{\alpha}$ terms cancel:

$$\Delta Y_1(0) = \underbrace{\alpha(U,X) + g_{Y_1}(U_{Y_1})}_{Y_1(0)} - \underbrace{\alpha(U,X) + g_{Y_0}(U_{Y_0})}_{Y_0(0)} = g_{Y_1}(U_{Y_1}) - g_{Y_0}(U_{Y_0})$$

Key consequence: $\boldsymbol{U}$ **does not directly influence** $\Delta Y_1(0)$ – the $\boldsymbol{U} \rightarrow \boldsymbol{\Delta Y_1(0)}$ arrows vanish from the graph.

The Δ-SWIG



Notice: U has **no direct arrow** into $\Delta Y_1(0)$

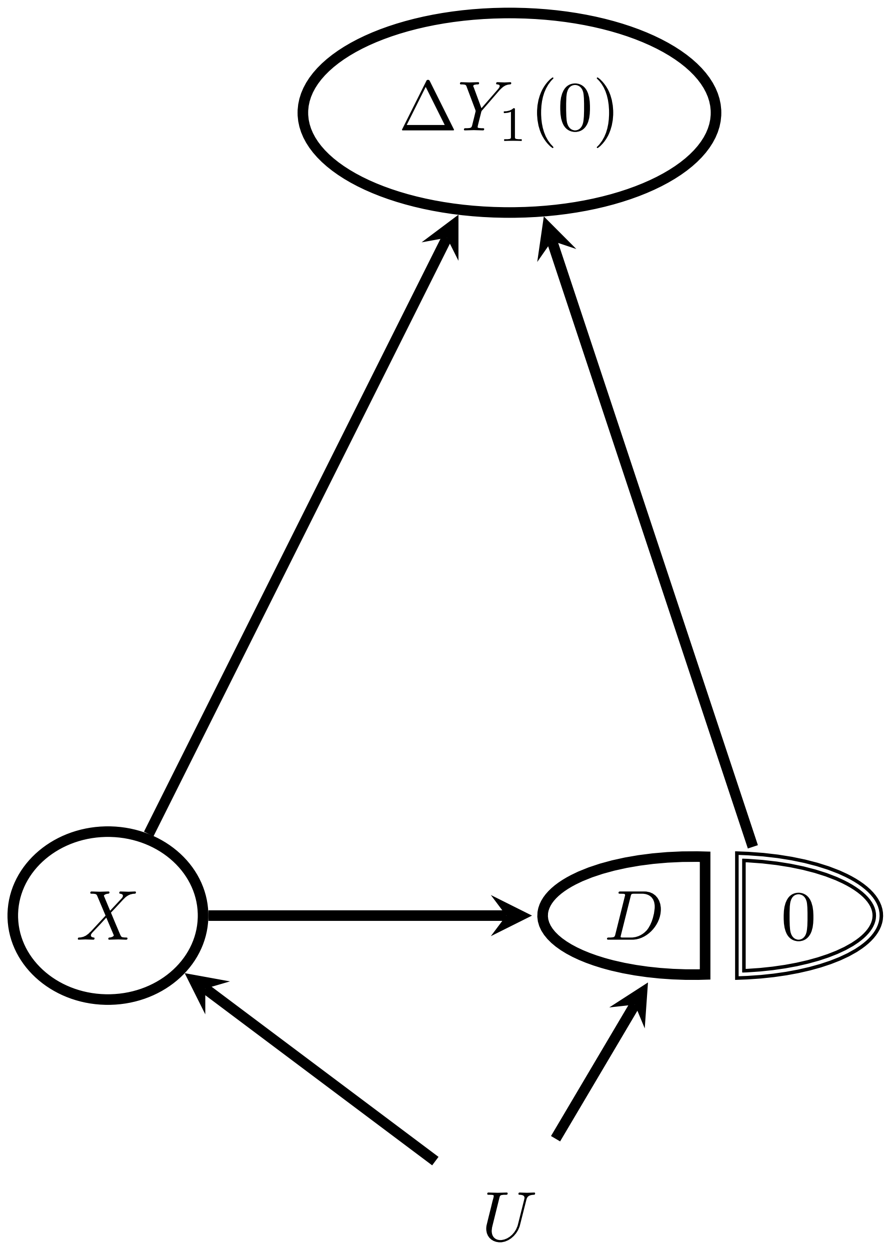
The $+\alpha$ terms cancelled!

Other than that: All direct causes of either Y_0 or Y_1 are also direct causes of $\Delta Y_1(0)$.

U can still influence $\Delta Y_1(0)$ – but only **through** X

Reading Off CPT: The Pruned Δ-SWIG

For CPT, we only need to ask: Is $\Delta Y_1(0) \perp\!\!\!\perp D \mid X$? Drop all nodes irrelevant to this question ($Y_0(0)$, $Y_1(0)$, U_{Y_0}, U_{Y_1}):



Check path blocking between $\Delta Y_1(0)$ and D given X:

Path $\Delta Y_1(0) \leftarrow X \rightarrow D$: blocked by X ✓

Path $\Delta Y_1(0) \leftarrow X \leftarrow U \rightarrow D$: blocked by X ✓

Result: $\Delta Y_1(0) \perp\!\!\!\perp D \mid X$

We get Conditional Parallel Trends

The $\boldsymbol{\Delta}$-SWIG implies:

$$\Delta Y_1(0) \perp\!\!\!\perp D \mid X.$$

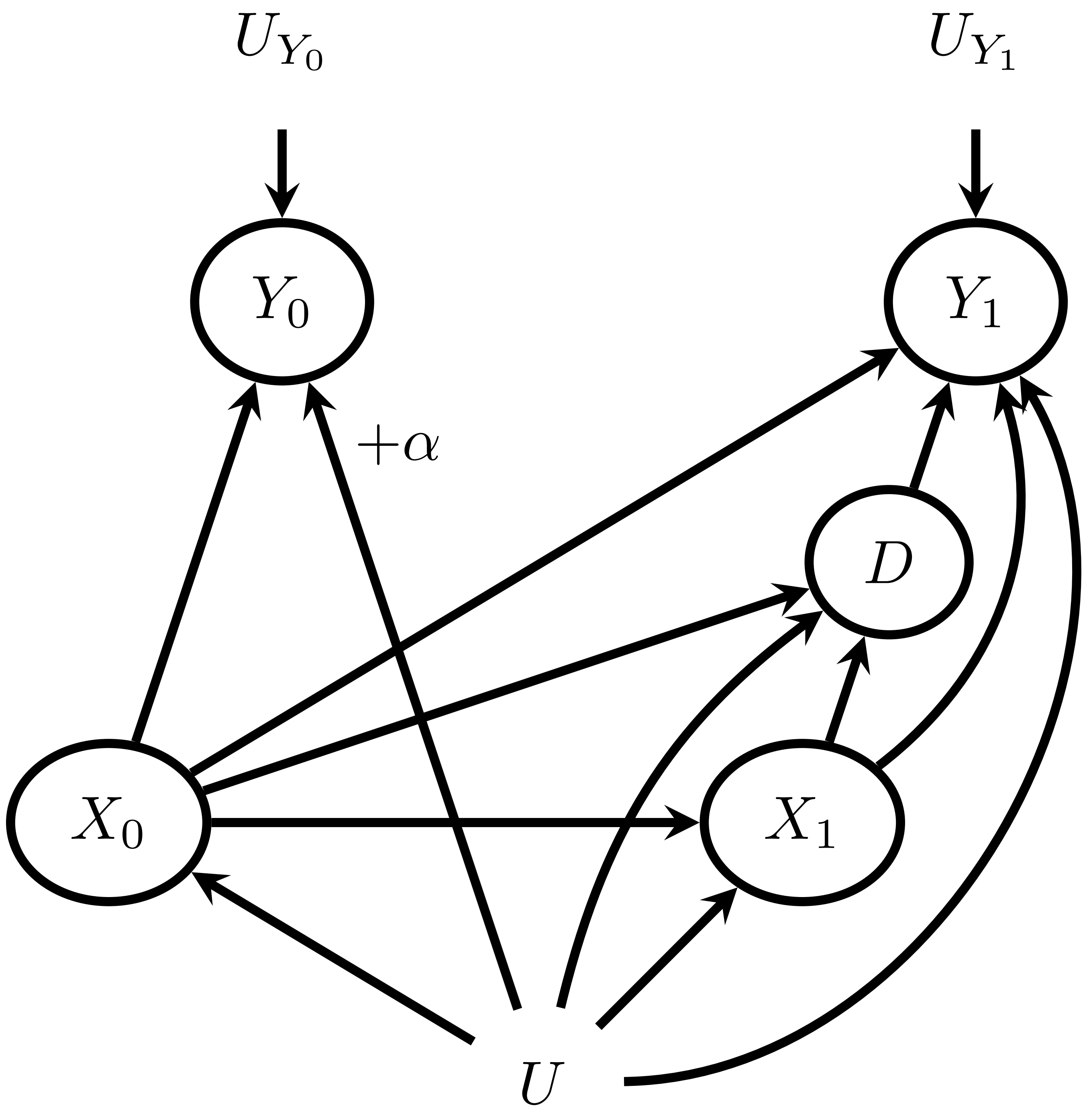
Thus, we can write:

$$\begin{aligned}\mathbb{E}[Y_1(0) - Y_0(0) \mid D = 1, X] &= \mathbb{E}[\Delta Y_1(0) \mid D = 1, X] \\ &= \mathbb{E}[\Delta Y_1(0) \mid D = 0, X] \\ &= \mathbb{E}[Y_1(0) - Y_0(0) \mid D = 0, X]\end{aligned}$$

This is Conditional Parallel Trends and shows how it can be derived graphically.

Part IV: More Results in Two Time Periods

Time-Varying Covariates: The DAG



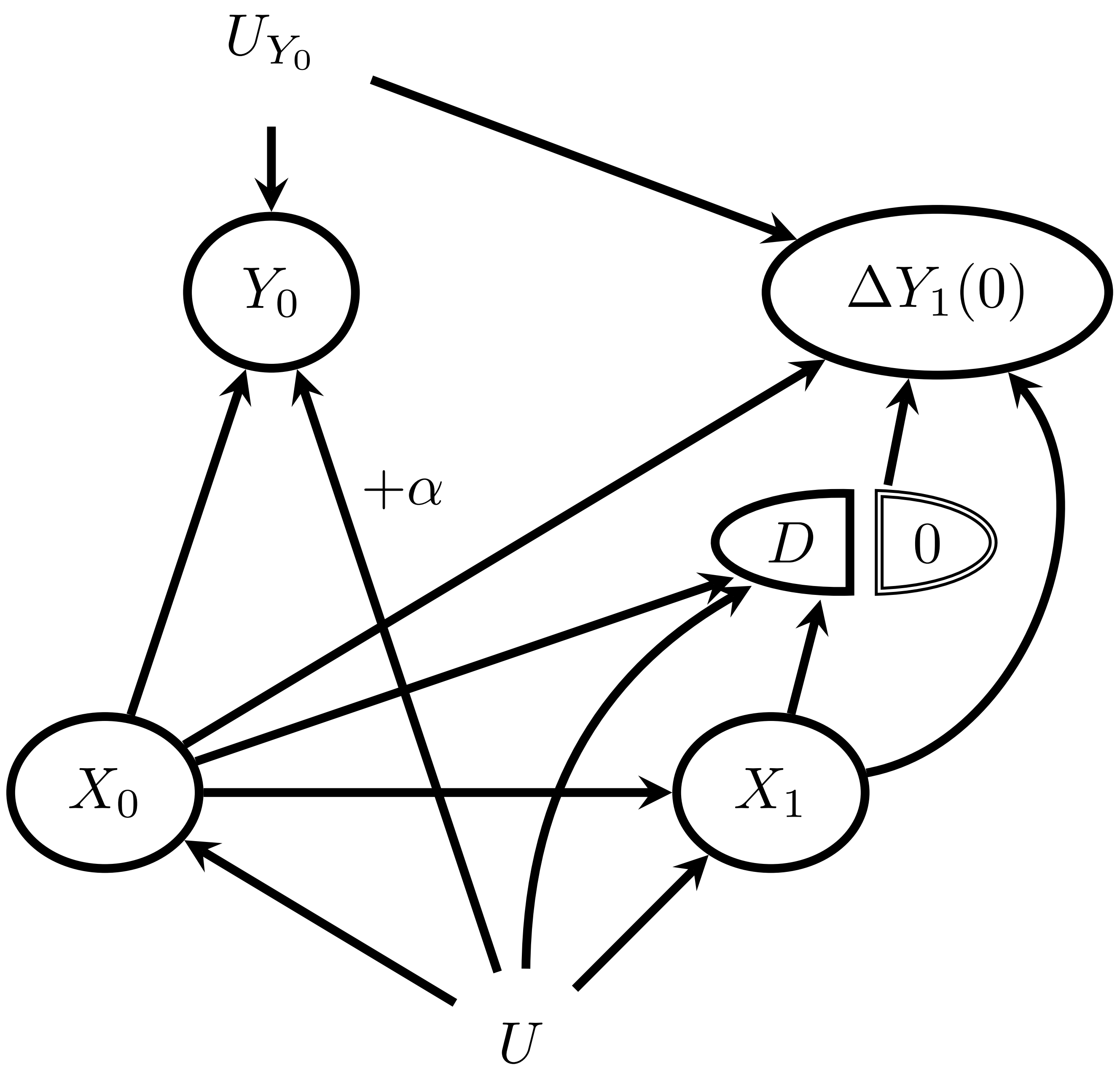
Setting: X_1 is a *time-varying covariate* — observed after the baseline, before the outcome

Both X_0 and X_1 affect treatment selection and Y_1

What CPT can be read-off here? Let's check using a Δ-SWIG.

Time-Varying Covariates: The Δ-SWIG

From now on, we skip the intermediate SWIG and go directly to the final Δ-SWIG.



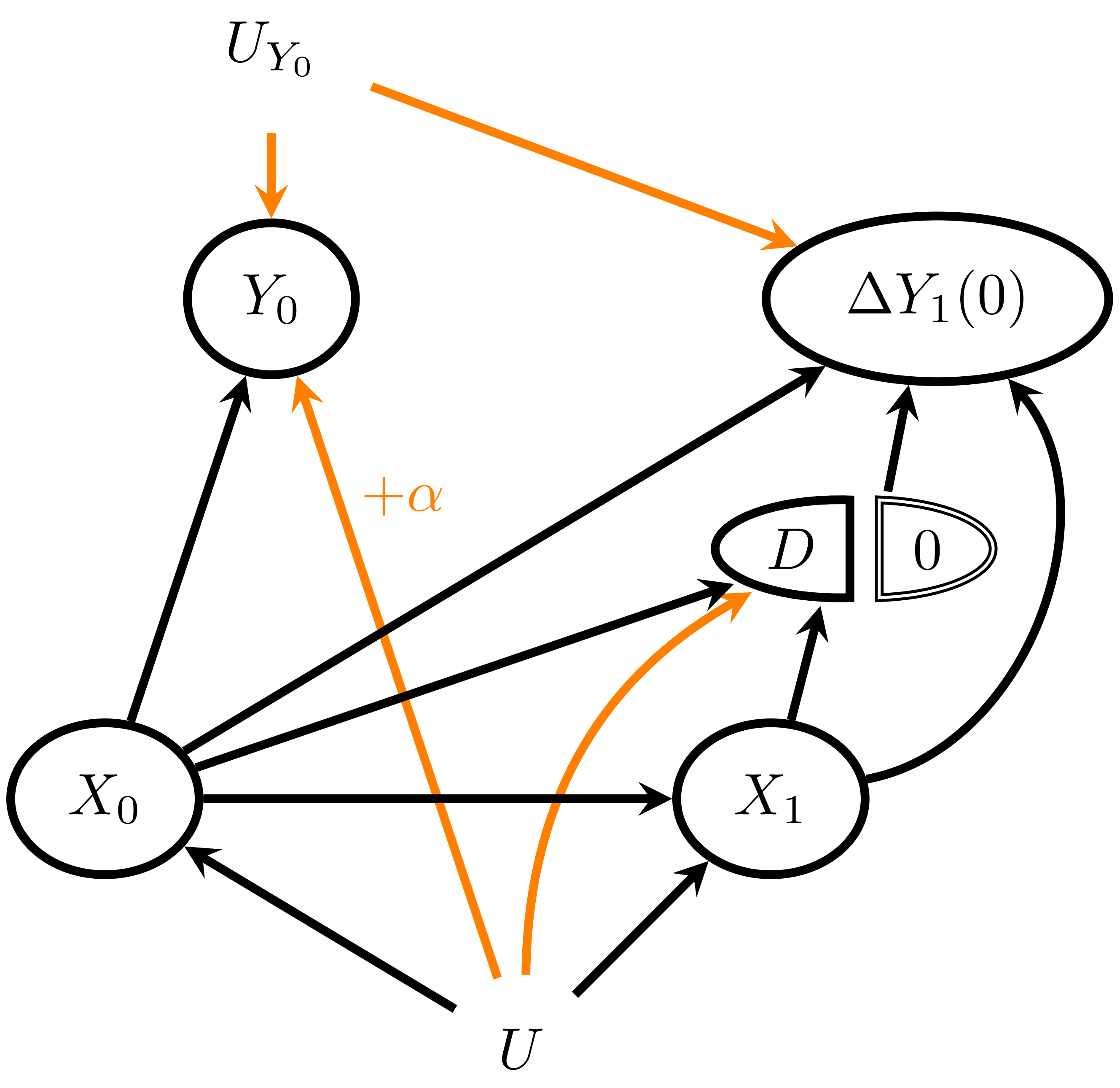
Check the path blocking between $\Delta Y_1(0)$ and D:

The paths $\Delta Y_1(0) \leftarrow X_0 \rightarrow D$ and $\Delta Y_1(0) \leftarrow X_0 \leftarrow U \rightarrow D$ are blocked by X_0.

The paths $\Delta Y_1(0) \leftarrow X_1 \rightarrow D$ and $\Delta Y_1(0) \leftarrow X_1 \leftarrow U \rightarrow D$ are blocked by X_1.

What about the paths via Y_0?

Time-Varying Covariates: Pre-Treatment Outcomes



Check the orange path.

Y_0 is a collider.

The path is blocked without conditioning on Y_0!

All other paths are blocked by X_0, X_1. Thus: $\Delta Y_1(0) \perp\!\!\!\perp D \mid X_0, X_1$.

If we also condition on Y_0, the orange path is suddenly unblocked. We can NOT read off:

$$\Delta Y_1(0) \perp\!\!\!\perp D \mid X_0, X_1, Y_0.$$

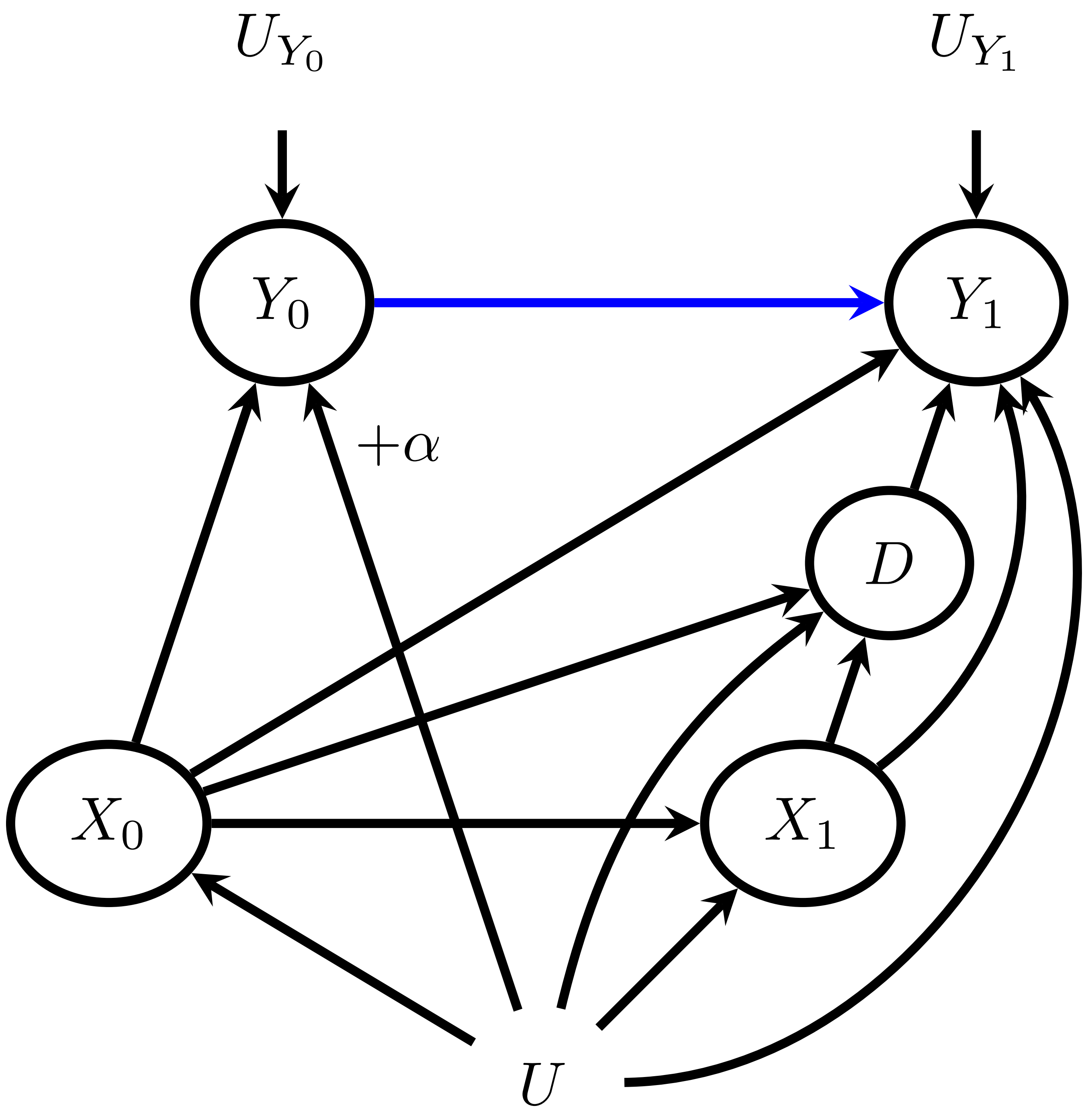
Y_0 is a bad control.

Outcome Dynamics - State dependence

Consider the case of state dependence:

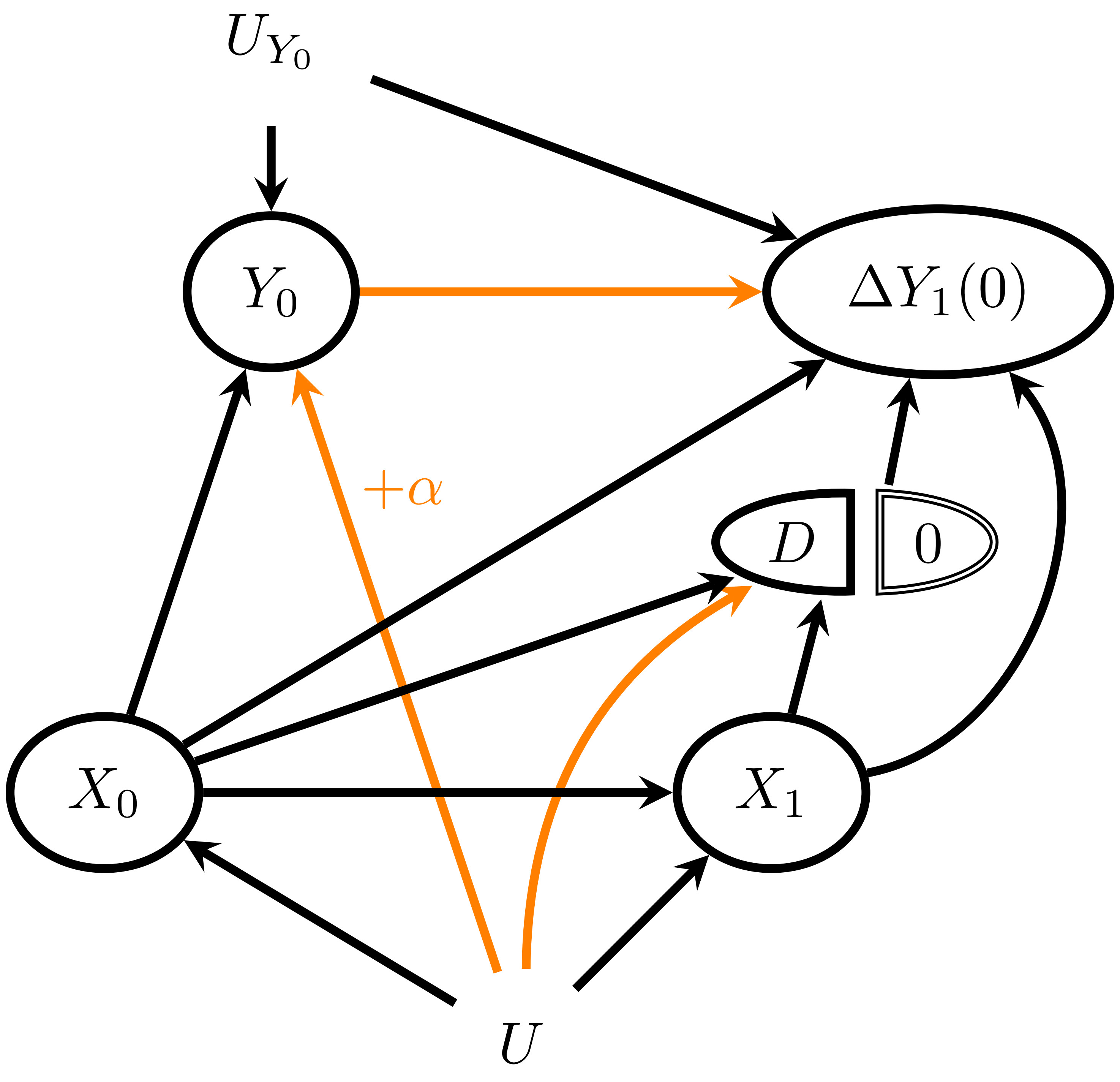
Outcome Dynamics - State dependence

Consider the case of state dependence:



Outcome Dynamics - State dependence

Consider the case of state dependence:



Check the orange path.

The only way to block it is by conditioning on Y_0.

But conditioning on Y_0 unblocks the same path as before:

$$\Delta Y_1(0) \leftarrow U_{Y_0} \rightarrow Y_0 \leftarrow U \rightarrow D.$$

We cannot block all paths between $\Delta Y_1(0)$ and D.

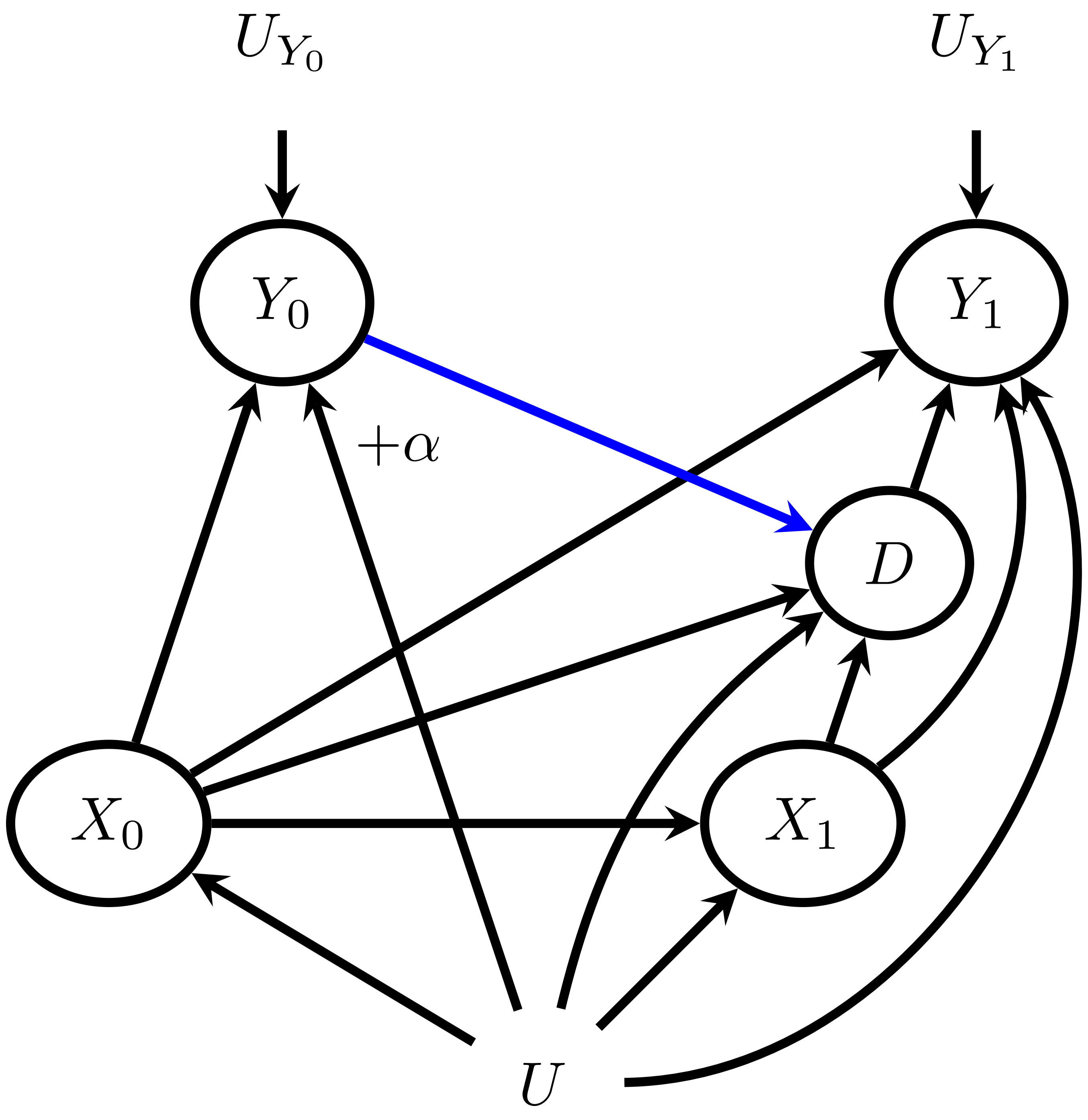
No CPT from this graph.

Outcome Dynamics - Selection on past outcomes

There is a similar story for selection on past outcomes:

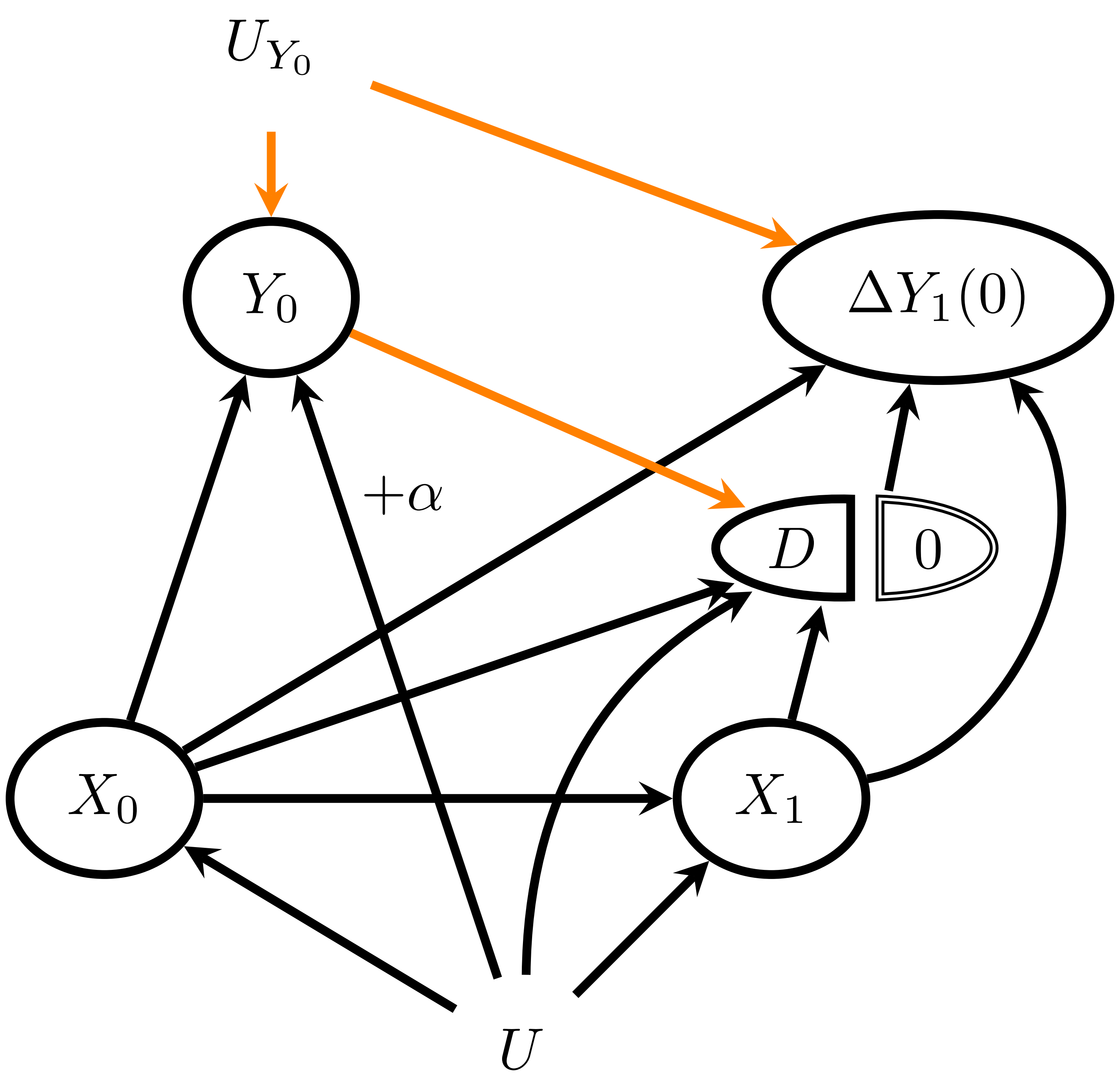
Outcome Dynamics - Selection on past outcomes

There is a similar story for selection on past outcomes:



Outcome Dynamics - Selection on past outcomes

There is a similar story for selection on past outcomes:



Check the orange path.

The only way to block it is by conditioning on Y_0.

Again: conditioning on Y_0 opens the same path as before:

$$\Delta Y_1(0) \leftarrow U_{Y_0} \rightarrow Y_0 \leftarrow U \rightarrow D.$$

We cannot block all paths between $\Delta Y_1(0)$ and D.

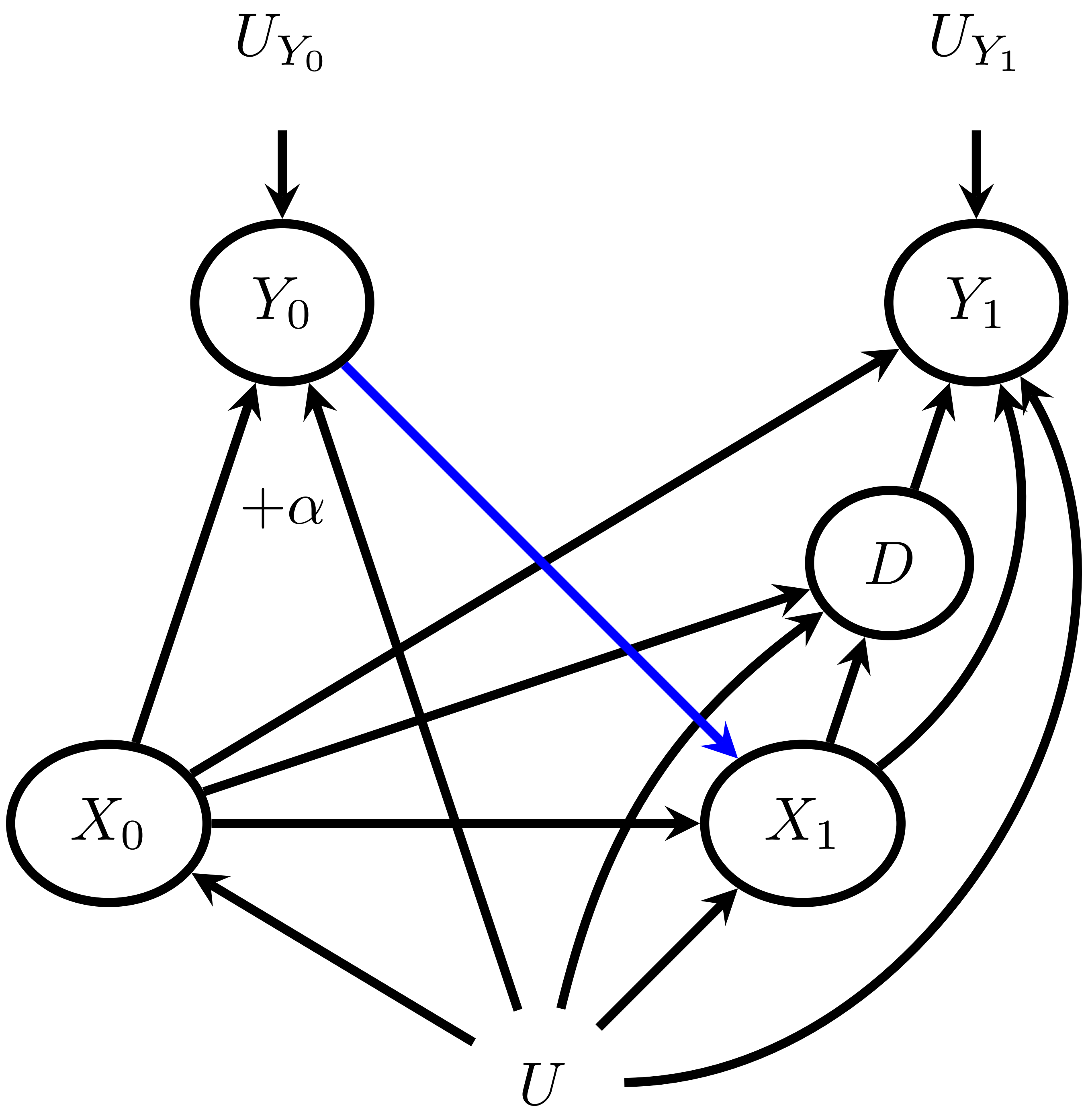
No CPT from this graph.

Outcome Dynamics - Outcome Covariate Feedback

The case of outcome-covariate feedback is slightly more difficult:

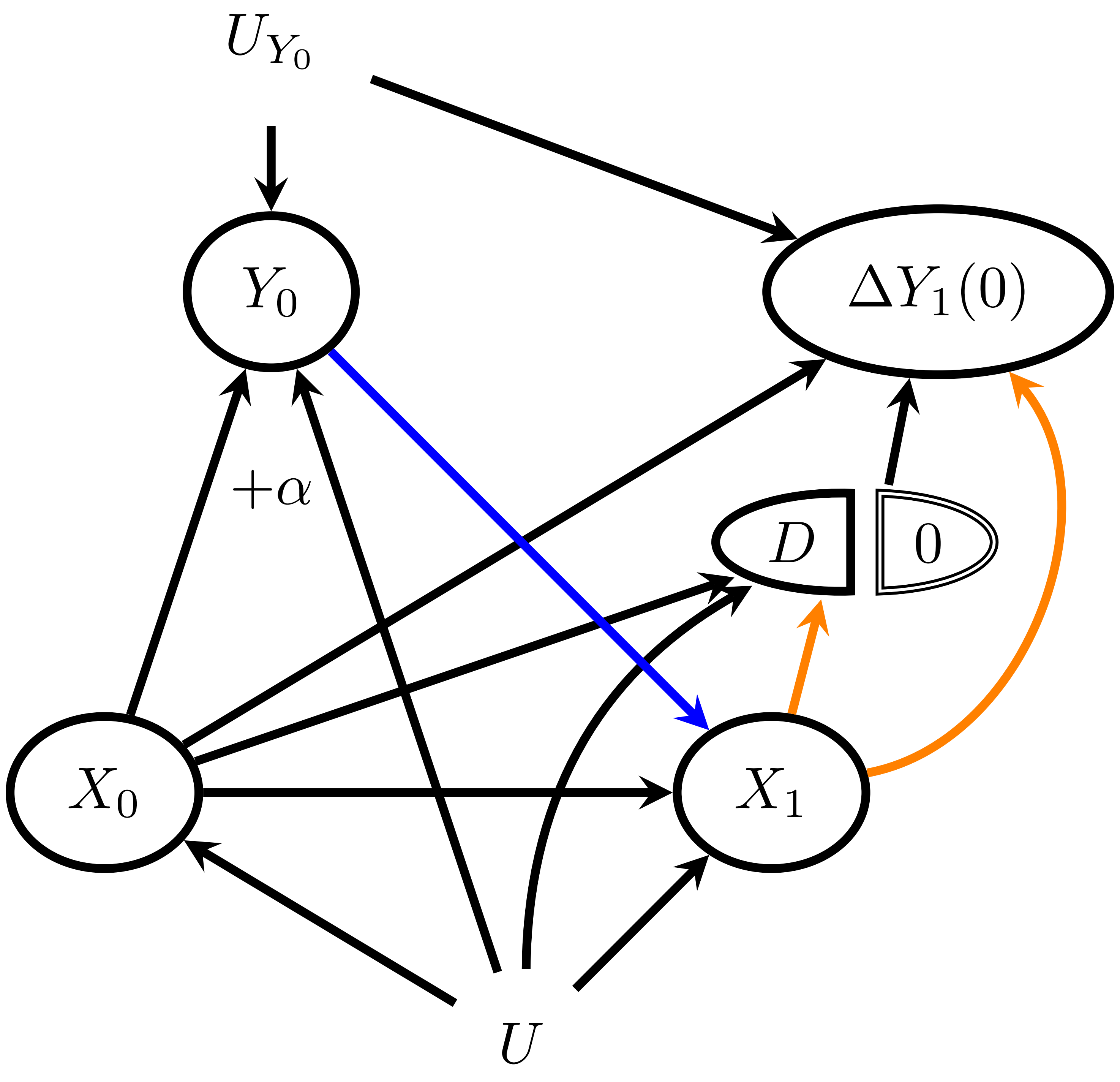
Outcome Dynamics - Outcome Covariate Feedback

The case of outcome-covariate feedback is slightly more difficult:



Outcome Dynamics - Outcome Covariate Feedback

The case of outcome-covariate feedback is slightly more difficult:



Check the orange path.

The only way to block it is by conditioning on X_1.

But: X_1 is descendant of Y_0. Conditioning on X_1 has the same effect as conditioning on collider Y_0.

It unblocks the path:

$$\Delta Y_1(0) \leftarrow U_{Y_0} \rightarrow Y_0 \leftarrow U \rightarrow D.$$

We cannot block all paths between $\Delta Y_1(0)$ and D.

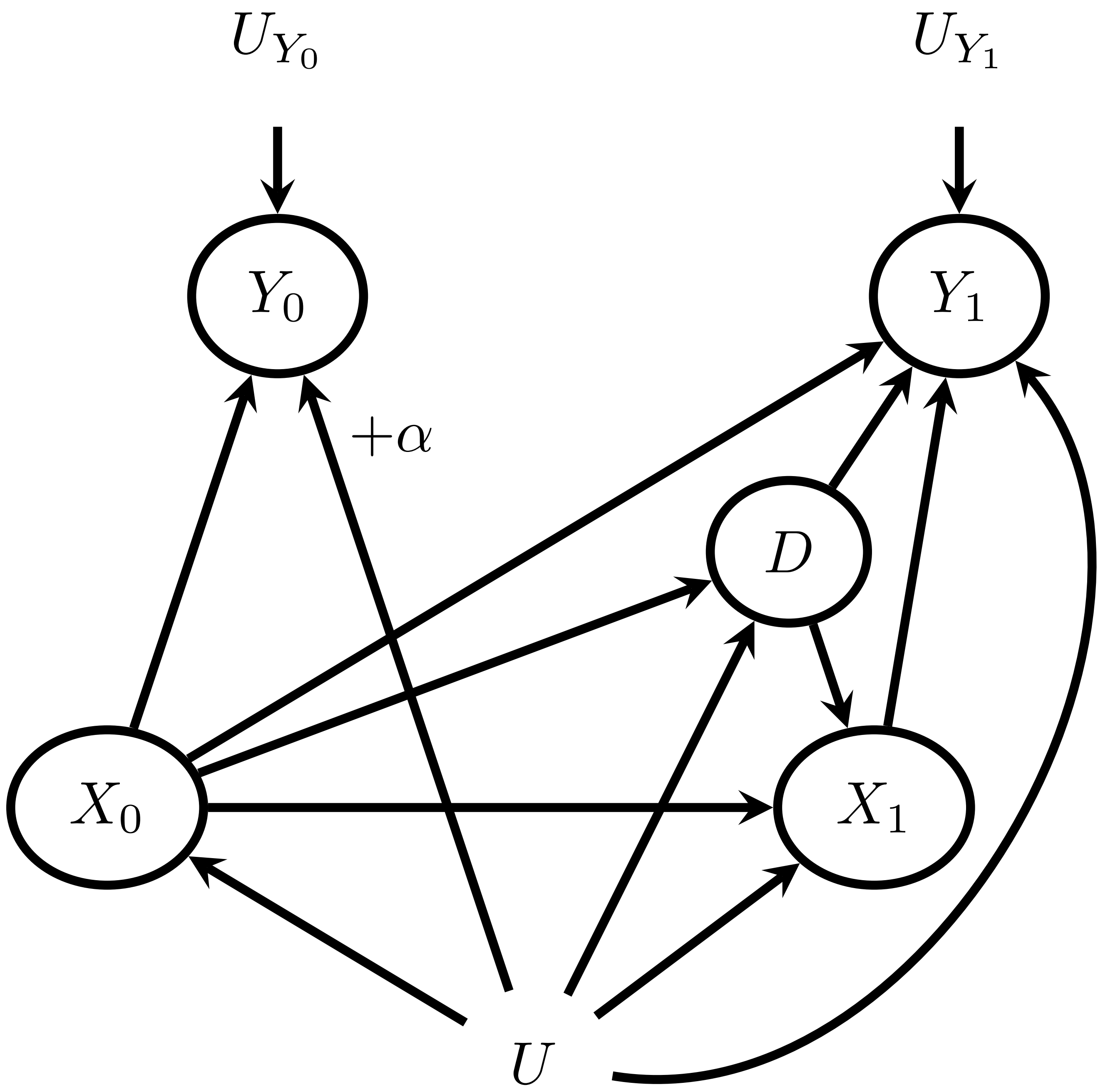
No CPT from this graph.

Post-Treatment Controls

Let's make the following change: $D \to X_1$ instead of $X_1 \to D$.

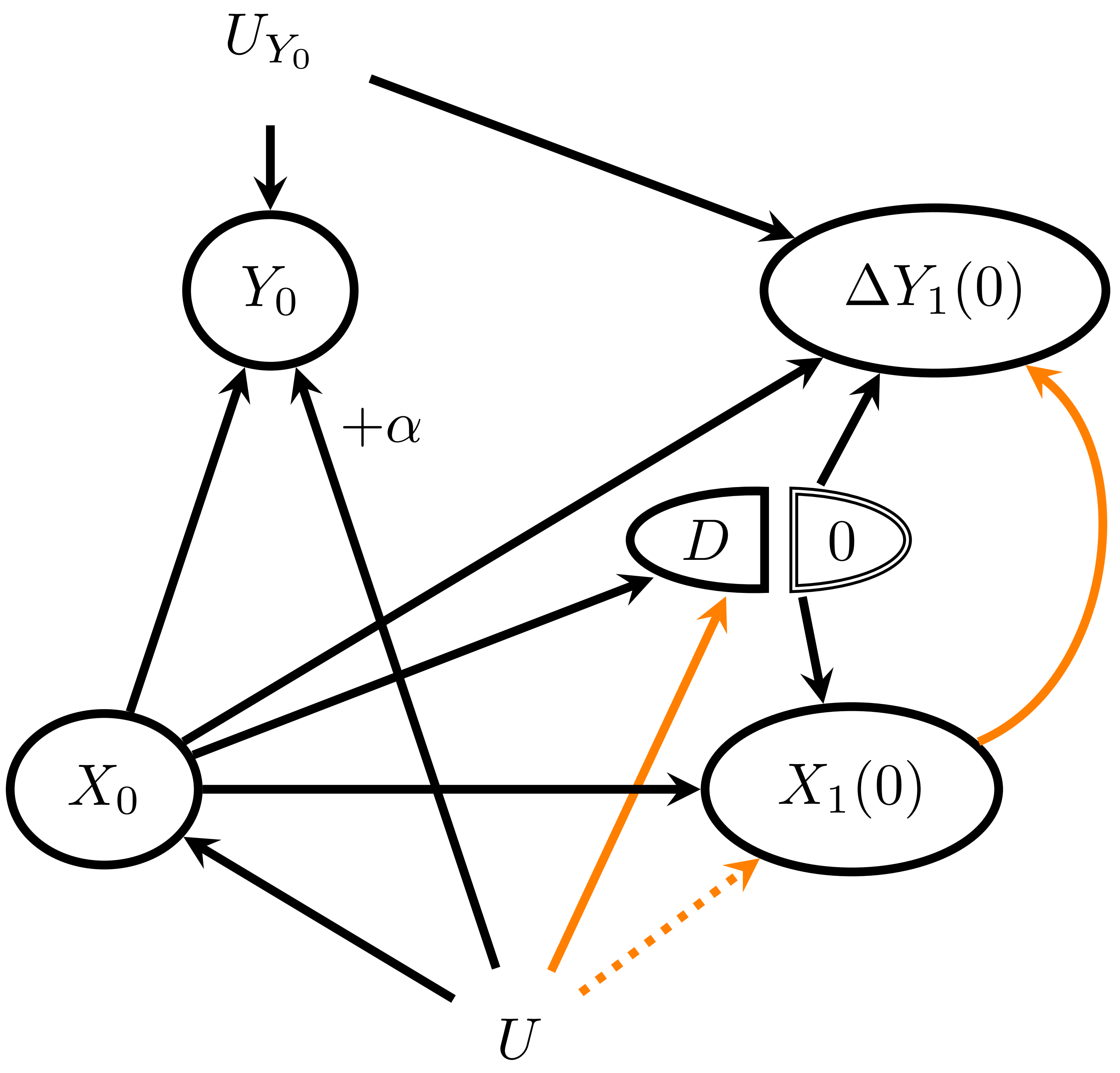
Post-Treatment Controls

Let's make the following change: $D \to X_1$ instead of $X_1 \to D$.



Post-Treatment Controls

Let's make the following change: $D \to X_1$ instead of $X_1 \to D$.



Check the orange path. The only way to block it is by conditioning on $X_1(0)$.

But: $X_1(0)$ is unobserved and not equal to the observed X_1 for those that are treated. We cannot condition on it.

Potential solution: assume covariate unconfoundedness, i.e., absence of the dotted edge $U \to X_1(0)$.

Then, all paths are blocked by X_0.

We get $\Delta Y_1(0) \perp\!\!\!\perp D \mid X_0$, i.e., parallel trends conditional on X_0.

Paths via the fixed nodes are blocked.

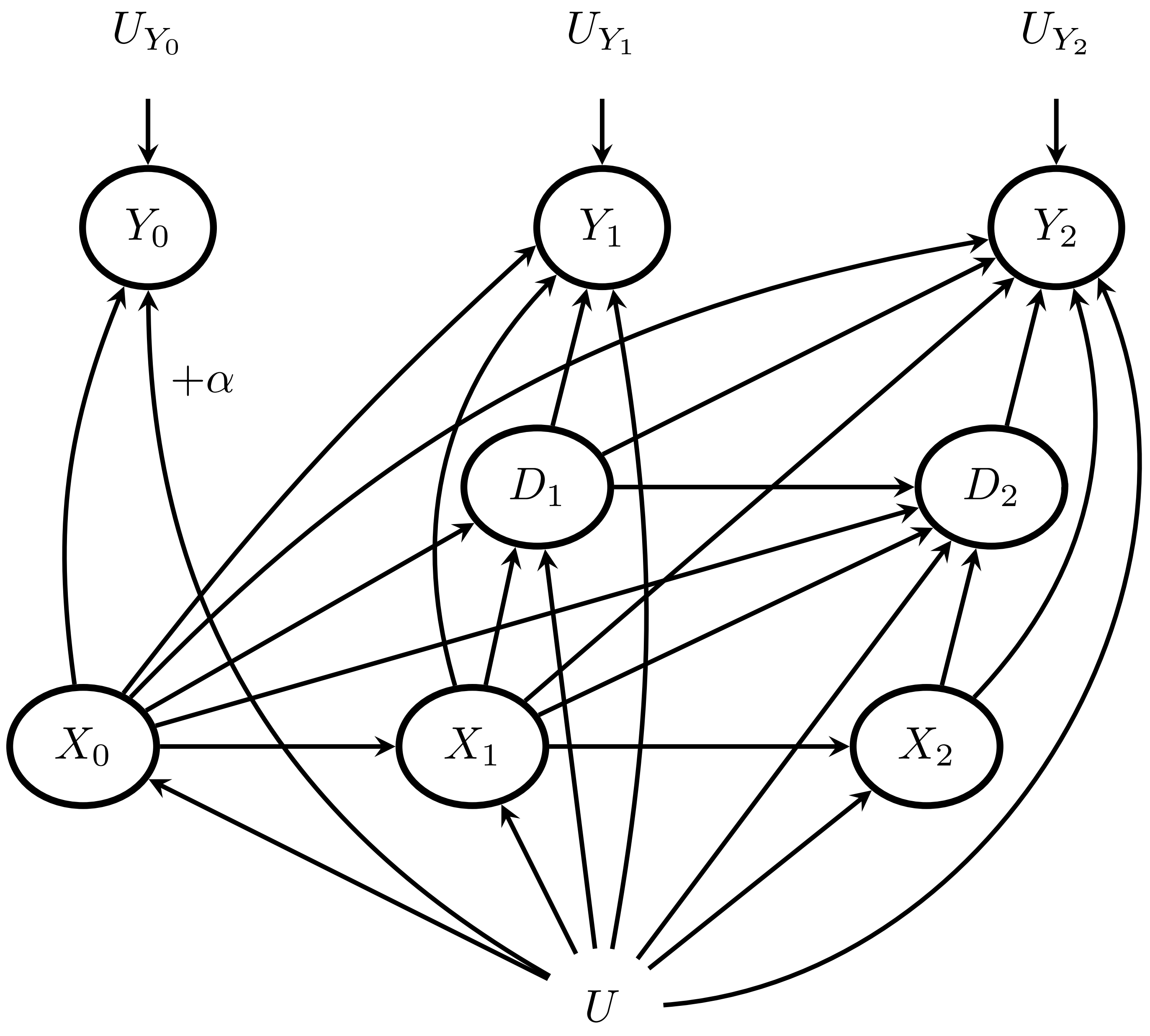
Part V: Multiple Time Periods

Multiple Time Periods: $\Delta Y_1(0,0)$ and D_1

The framework introduced so far extends to more than two time periods as discussed in (Callaway & Sant'Anna, 2021). The main insights can be illustrated in three time periods.

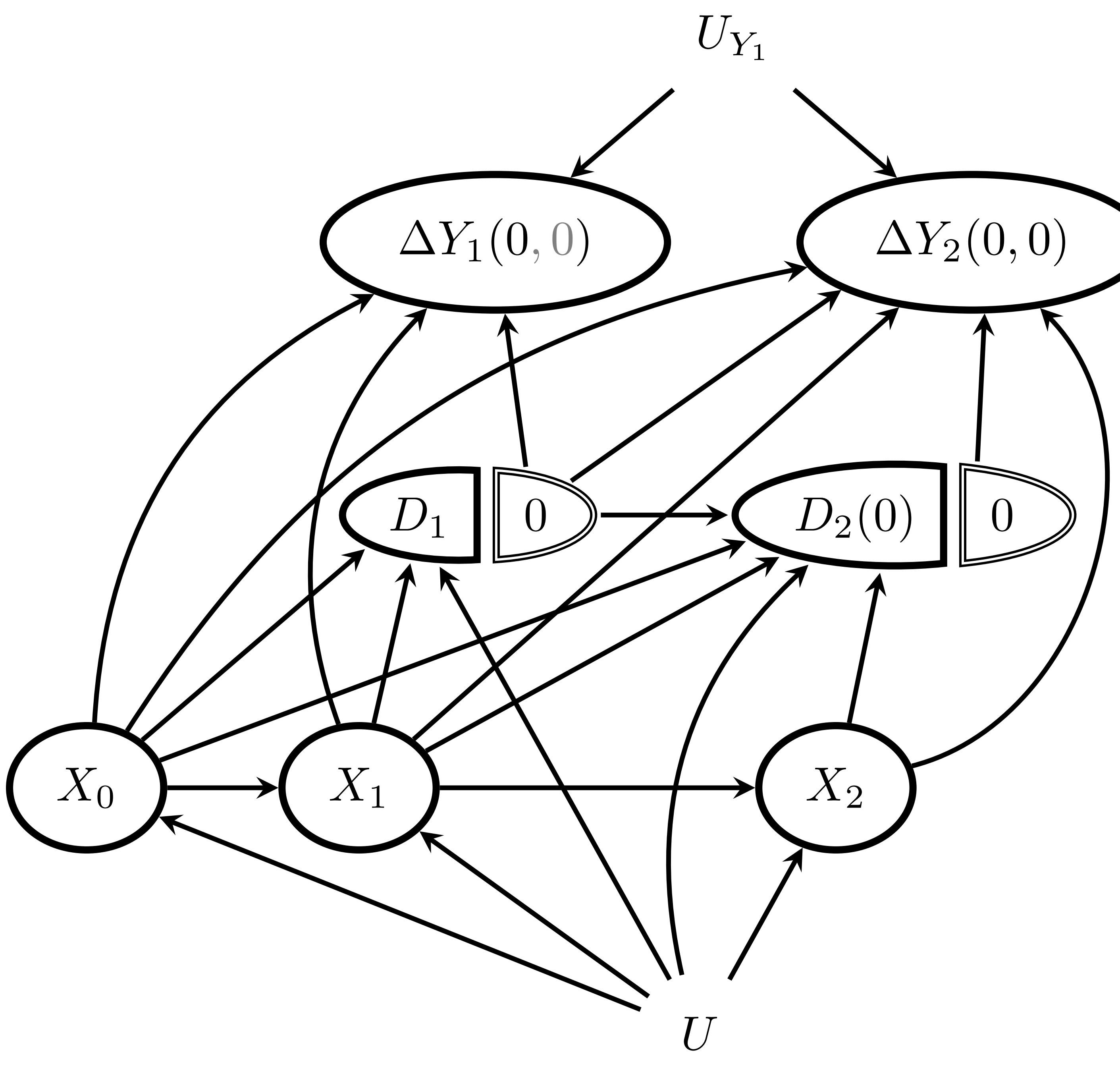
Multiple Time Periods: $\Delta Y_1(0,0)$ and D_1

The framework introduced so far extends to more than two time periods as discussed in (Callaway & Sant'Anna, 2021). The main insights can be illustrated in three time periods.



Multiple Time Periods: $\Delta Y_1(0,0)$ and D_1

The framework introduced so far extends to more than two time periods as discussed in (Callaway & Sant'Anna, 2021). The main insights can be illustrated in three time periods.



Paths via the fixed nodes are blocked.

Check the paths connecting $\Delta Y_1(0,0)$ and D_1.

$$\Delta Y_1(0,0) \leftarrow X_0 \rightarrow D_1,$$
$$\Delta Y_1(0,0) \leftarrow X_1 \rightarrow D_1,$$
$$\Delta Y_1(0,0) \leftarrow X_0 \leftarrow U \rightarrow D_1,$$
$$\Delta Y_1(0,0) \leftarrow X_1 \leftarrow U \rightarrow D_1 ...$$

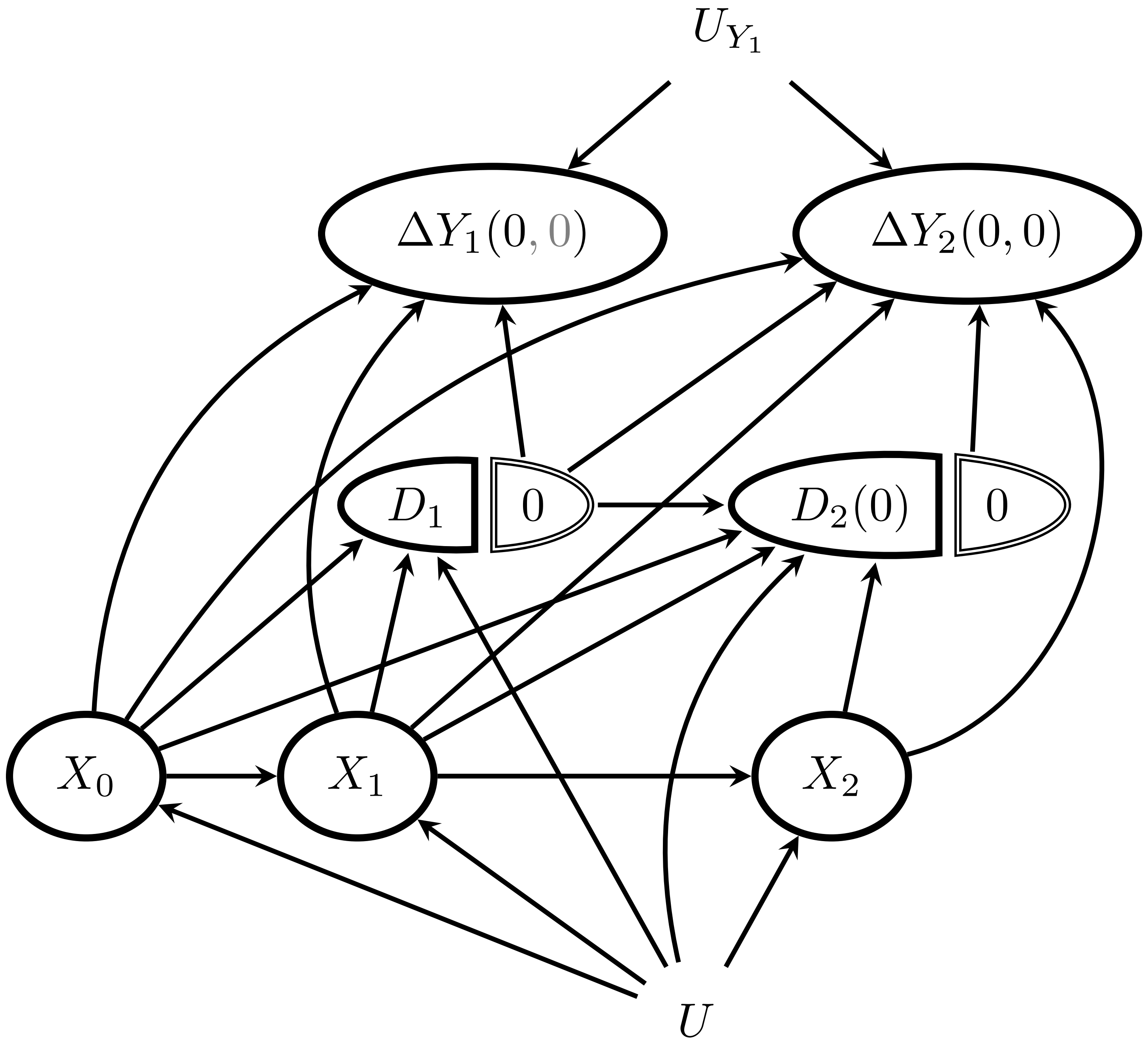
Contain (and can be blocked by) X_0, X_1. Conditioning on X_2 does not unblock any path.

The paths $\Delta Y_1(0,0) \leftarrow U_{Y_1} \rightarrow \Delta Y_2(0,0) \leftarrow ...$ are blocked by the collider $\Delta Y_2(0,0)$.

This implies: $\Delta Y_1(0,0) \perp\!\!\!\perp D_1 \mid X_0, X_1, X_2$

Gray variables in the conditioning set means that including them is optional.

Multiple Time Periods: $\Delta Y_1(0,0)$ and $D_2(0)$



Paths via the fixed nodes are blocked.

Check the paths connecting $\Delta Y_1(0,0)$ and $D_2(0)$.

$$\Delta Y_1(0,0) \leftarrow X_0 \rightarrow D_2(0),$$
$$\Delta Y_1(0,0) \leftarrow X_1 \rightarrow D_2(0),$$
$$\Delta Y_1(0,0) \leftarrow X_0 \leftarrow U \rightarrow D_2(0),$$
$$\Delta Y_1(0,0) \leftarrow X_1 \leftarrow U \rightarrow D_2(0)...$$

Again, they can be blocked by X_0, X_1. Conditioning on X_2, D_1 does not unblock any path. The paths $\Delta Y_1(0,0) \leftarrow U_{Y_1} \rightarrow \Delta Y_2(0,0) \leftarrow ...$ are blocked by the collider $\Delta Y_2(0,0)$.

This implies:

$$\Delta Y_1(0,0) \perp\!\!\!\perp D_2(0) \mid X_0, X_1, X_2, D_1$$

Conditional Parallel Trends

From the previous slides we get:

$\Delta Y_1(0,0) \perp\!\!\!\perp D_1 \mid X_0, X_1, X_2$ and $\Delta Y_1(0,0) \perp\!\!\!\perp D_2(0) \mid X_0, X_1, X_2, D_1$

These jointly imply by contraction:

$$\Delta Y_1(0,0) \perp\!\!\!\perp D_1, D_2 \mid X_0, X_1, X_2$$

And thus:

$$\mathbb{E}[\Delta Y_1(0,0) \mid D_1 = 1, D_2 = 1, X_0, X_1, X_2] = \mathbb{E}[\Delta Y_1(0,0) \mid D_1 = 0, D_2 = 0, X_0, X_1, X_2]$$

This is conditional parallel trends in $t = 1$ between those treated in periods $t = 1, 2$ ($D_1 = D_2 = 1$) and those never treated ($D_1 = D_2 = 0$).

Pre-Trends

The joint independence:

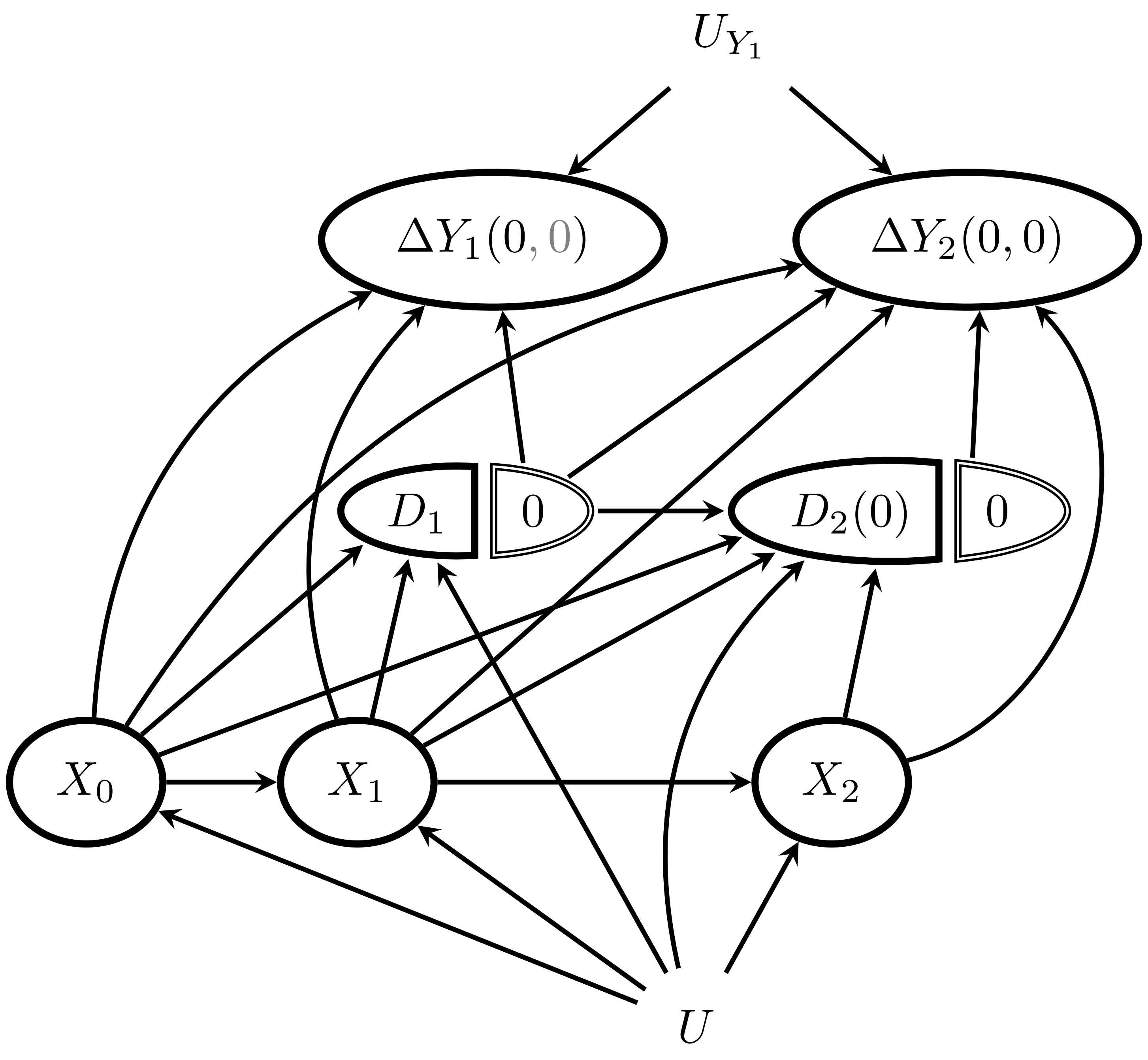
$$\Delta Y_1(0,0) \perp\!\!\!\perp D_1, D_2 \mid X_0, X_1, X_2$$

also implies:

$$\mathbb{E}[\Delta Y_1(0,0) \mid D_1 = 0, D_2 = 1, X_0, X_1, X_2] = \mathbb{E}[\Delta Y_1(0,0) \mid D_1 = 0, D_2 = 0, X_0, X_1, X_2].$$

This means parallel trends in $t = 1$ between those treated in $t = 2$ and those untreated throughout, i.e., parallel pre-trends in $t = 1$, conditional on X_0, X_1 and optionally X_2.

Multiple Time Periods: $\Delta Y_2(0,0)$ and D_1



Paths via the fixed nodes are blocked.

Check the paths connecting $\Delta Y_2(0,0)$ and D_1.

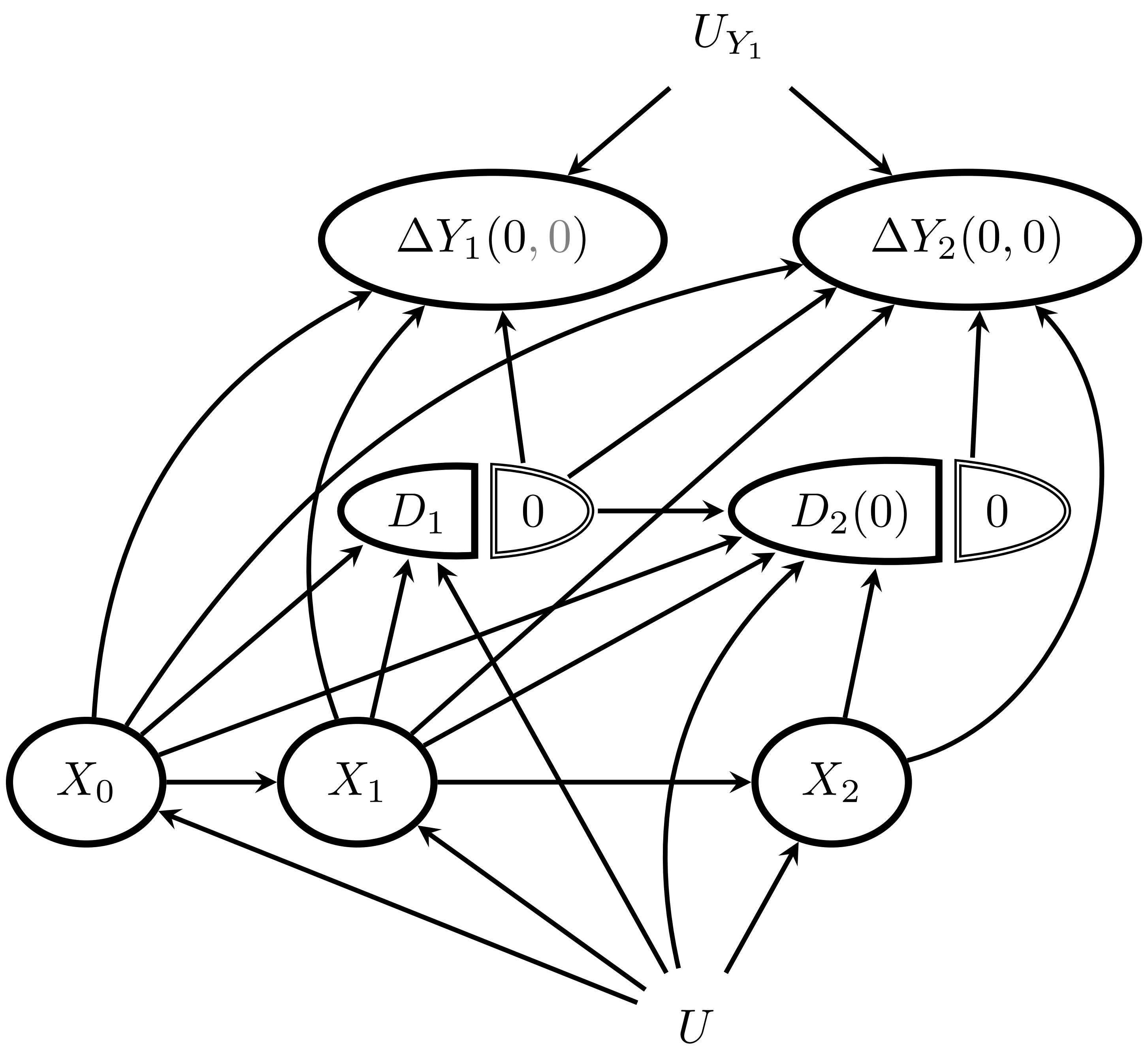
$$\Delta Y_2(0,0) \leftarrow X_0 \rightarrow D_1,$$
$$\Delta Y_2(0,0) \leftarrow X_1 \rightarrow D_1,$$
$$\Delta Y_2(0,0) \leftarrow X_0 \leftarrow U \rightarrow D_1,$$
$$\Delta Y_2(0,0) \leftarrow X_1 \leftarrow U \rightarrow D_1,$$
$$\Delta Y_2(0,0) \leftarrow X_2 \leftarrow U \rightarrow D_1 ...$$

These can be blocked by X_0, X_1, X_2. The paths $\Delta Y_2(0,0) \leftarrow U_{Y_1} \rightarrow \Delta Y_1(0,0) \leftarrow ...$ are blocked by the collider $\Delta Y_1(0,0)$.

This implies:

$$\Delta Y_2(0,0) \perp\!\!\!\perp D_1 \mid X_0, X_1, X_2$$

Multiple Time Periods: $\Delta Y_2(0,0)$ and $D_2(0)$



Paths via the fixed nodes are blocked.

Check the paths connecting $\Delta Y_2(0,0)$ and $D_2(0)$.

$$\Delta Y_2(0,0) \leftarrow X_0 \rightarrow D_2(0),$$
$$\Delta Y_2(0,0) \leftarrow X_1 \rightarrow D_2(0),$$
$$\Delta Y_2(0,0) \leftarrow X_2 \rightarrow D_2(0),$$
$$\Delta Y_2(0,0) \leftarrow X_0 \leftarrow U \rightarrow D_2(0),$$
$$\Delta Y_2(0,0) \leftarrow X_1 \leftarrow U \rightarrow D_2(0),$$
$$\Delta Y_2(0,0) \leftarrow X_2 \leftarrow U \rightarrow D_2(0)...$$

These can be blocked by X_0, X_1, X_2. Conditioning on D_1 does not unblock any path.

This implies:

$$\Delta Y_2(0,0) \perp\!\!\!\perp D_2(0) \mid X_0, X_1, X_2, D_1$$

Conditional Parallel Trends

From the previous slides we get:

$\Delta Y_2(0,0) \perp\!\!\!\perp D_1 \mid X_0, X_1, X_2$ and $\Delta Y_2(0,0) \perp\!\!\!\perp D_2(0) \mid X_0, X_1, X_2, D_1$

These jointly imply by contraction:

$$\Delta Y_2(0,0) \perp\!\!\!\perp D_1, D_2 \mid X_0, X_1, X_2$$

And thus, e.g.:

$$\mathbb{E}[\Delta Y_2(0,0) \mid D_1 = 1, D_2 = 1, X_0, X_1, X_2] = \mathbb{E}[\Delta Y_2(0,0) \mid D_1 = 0, D_2 = 0, X_0, X_1, X_2]$$

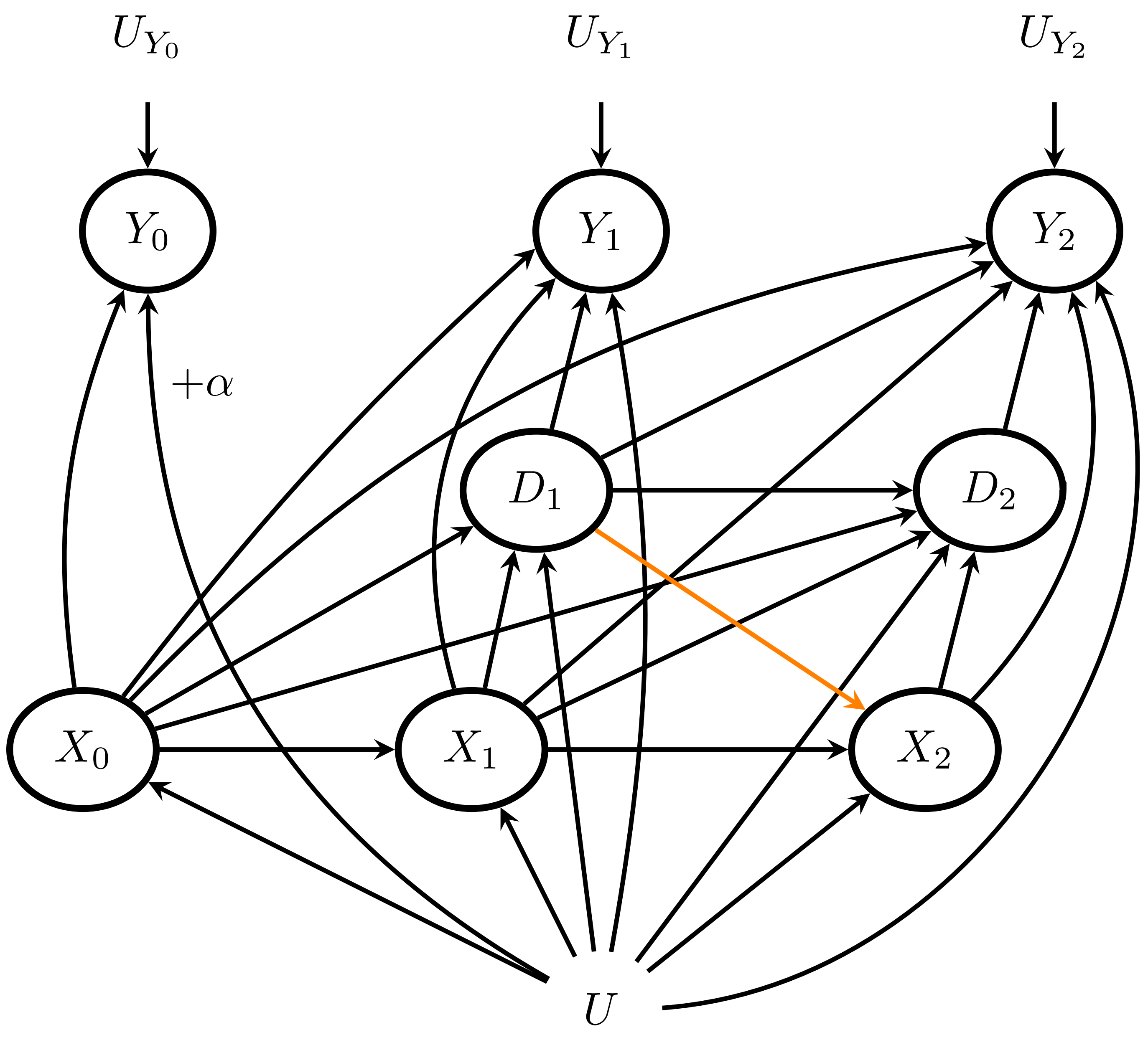
This is conditional parallel trends in $t = 2$ between those treated in periods $t = 1, 2$ ($D_1 = D_2 = 1$) and those never treated ($D_1 = D_2 = 0$).

Multiple Time Periods with $D \to X$: $\Delta Y_1(0,0)$ and D_1

Now, consider the following setting where the covariate X_2 is allowed to be affected by earlier treatment:

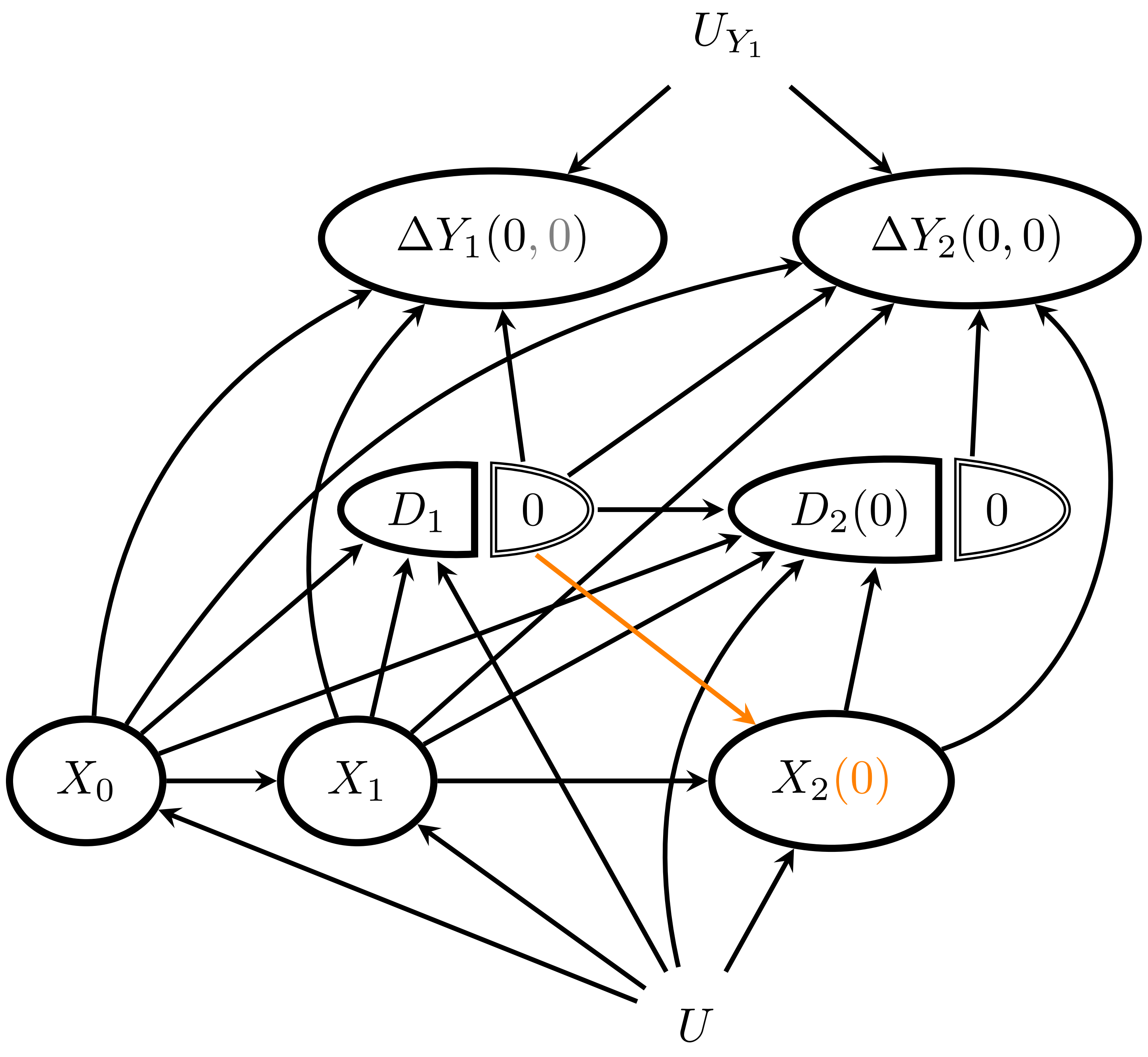
Multiple Time Periods with $D \to X$: $\Delta Y_1(0,0)$ and D_1

Now, consider the following setting where the covariate X_2 is allowed to be affected by earlier treatment:



Multiple Time Periods with $D \to X$: $\Delta Y_1(0,0)$ and D_1

Now, consider the following setting where the covariate X_2 is allowed to be affected by earlier treatment:



The paths connecting $\Delta Y_1(0,0)$ and D_1 still can be blocked by X_0, X_1.

The same is true for paths connecting $\Delta Y_1(0,0)$ and $D_2(0)$, where we can additionally condition on D_1.

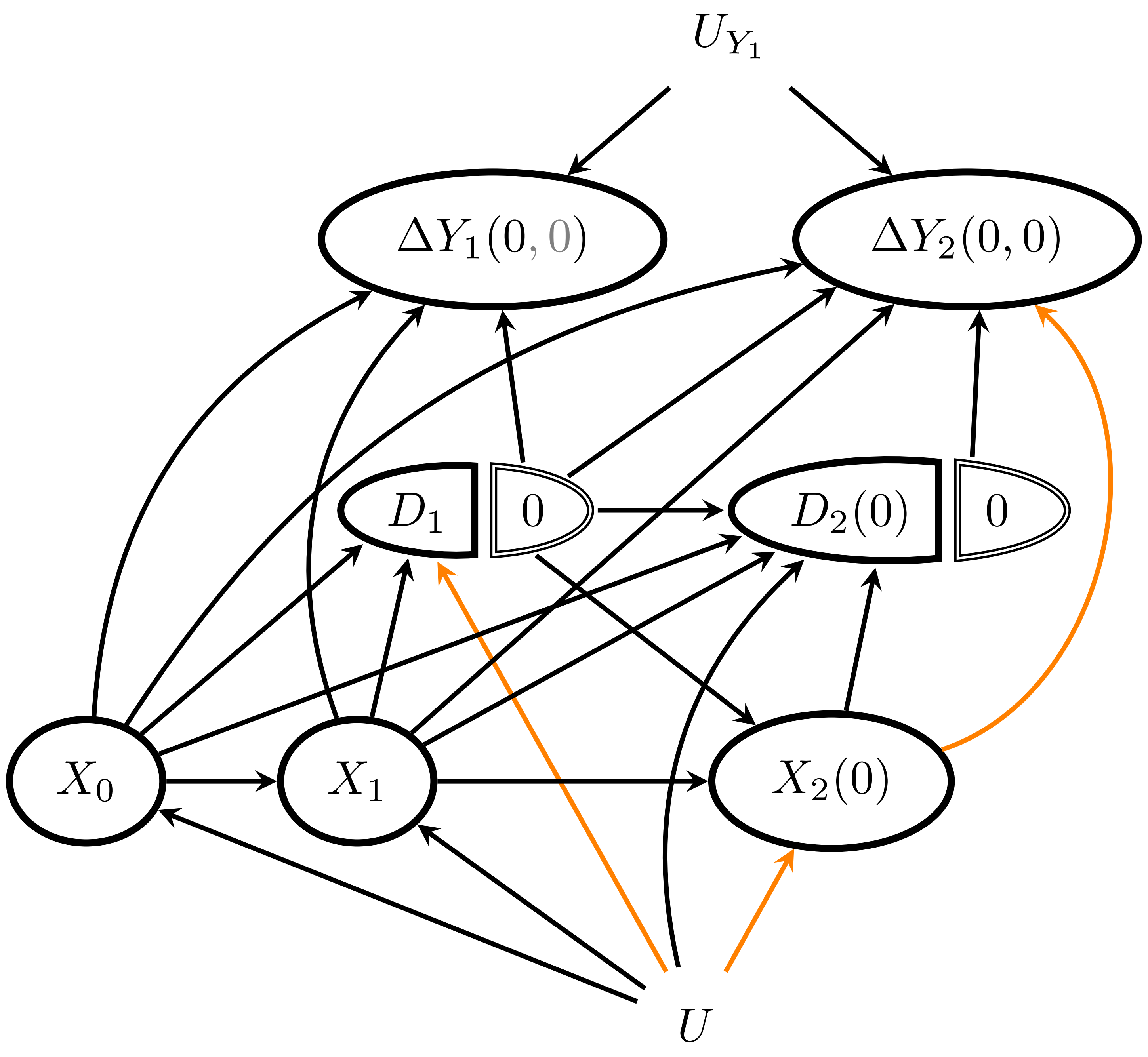
Conditioning on $X_2(0)$ does not unblock any path but it is unobserved. We cannot condition on it!

This implies: $\Delta Y_1(0,0) \perp\!\!\!\perp D_1, D_2 \mid X_0, X_1$.

CPT in $t = 1$ conditional on X_0, X_1.

Paths via the fixed nodes are blocked.

Multiple Time Periods with $D \to X$: $\Delta Y_2(0,0)$ and D_1



Paths via the fixed nodes are blocked.

Check the paths connecting $\Delta Y_2(0,0)$ and D_1. The orange path is problematic:

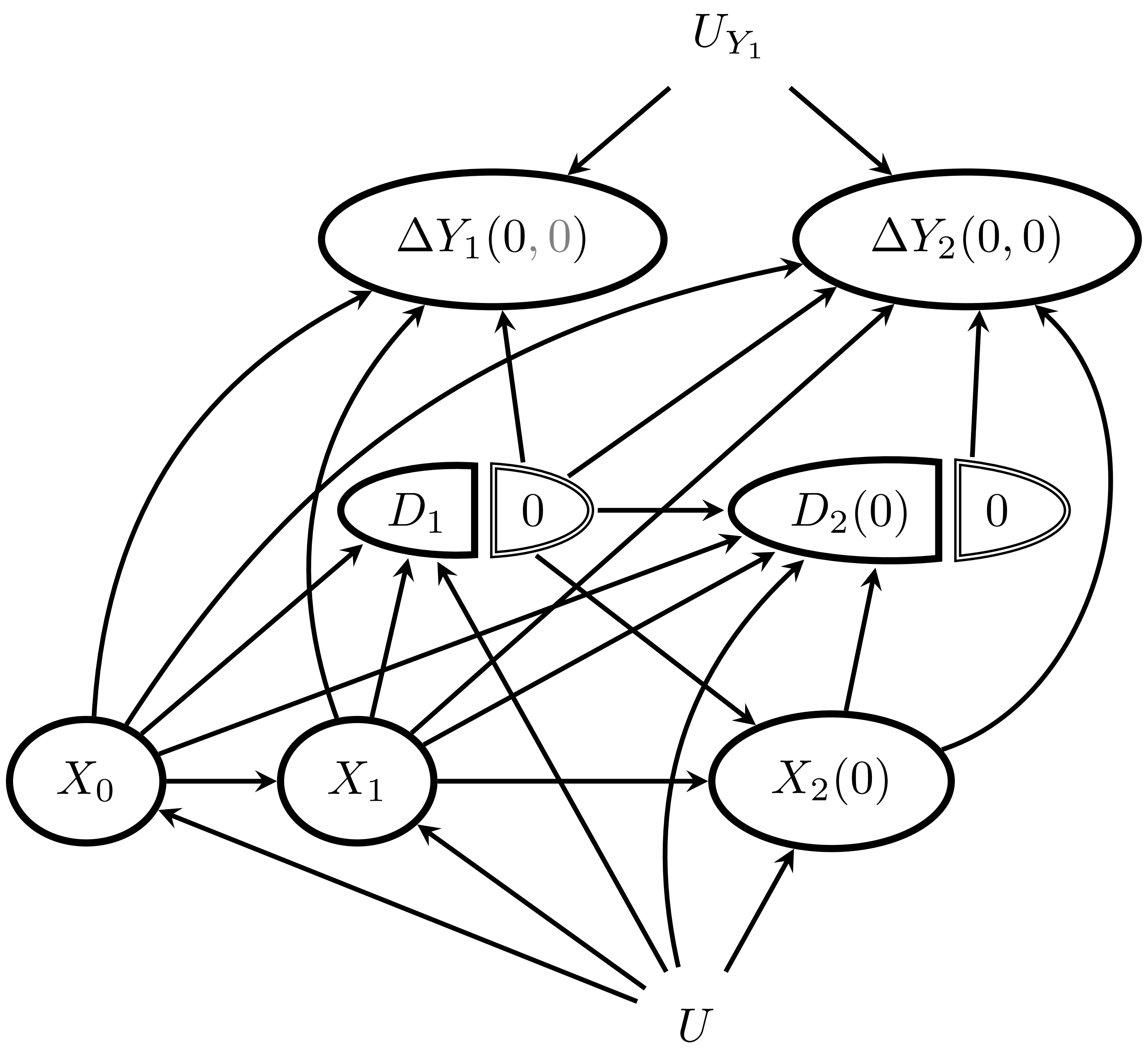
$$\Delta Y_2(0,0) \leftarrow X_2(0) \leftarrow U \rightarrow D_1$$

This can only be blocked by conditioning on $X_2(0)$. But we do not observe it.

We can not read off a conditional independence between $\Delta Y_2(0,0)$ and D_1 with observable variables in the conditioning set.

We can not read off CPT in $t=2$ between those treated in $t=1$ and those untreated throughout.

Multiple Time Periods with $D \to X$: $\Delta Y_2(0,0)$ and $D_2(0)$



Paths via the fixed nodes are blocked.

Check the paths connecting $\Delta Y_2(0,0)$ and $D_2(0)$. They are very similar to the setting without the $D_1 \to X_2$ edge.

The paths now contain $X_2(0)$ instead of X_2.

This is not problematic here, because we can always additionally condition on D_1. We get

$$\Delta Y_2(0,0) \perp\!\!\!\perp D_2(0) \mid X_0, X_1, X_2(0), D_1$$

and, by consistency:

$$\Delta Y_2(0,0) \perp\!\!\!\perp D_2 \mid X_0, X_1, X_2, D_1 = 0$$

We can read off CPT in $t = 2$ between those treated in $t = 2$ and those untreated throughout.

Pre-Trends

It can be seen that the Δ-SWIG also implies:

$$\Delta Y_1(0,0) \perp\!\!\!\perp X_2(0) \mid X_0, X_1, D_1, D_2(0).$$

and by consistency:

$$\Delta Y_1(0,0) \perp\!\!\!\perp X_2 \mid X_0, X_1, D_1 = 0, D_2.$$

This means we can not only write:

$$\mathbb{E}[\Delta Y_1(0,0) \mid X_0, X_1, D_1 = 0, D_2 = 1] = \mathbb{E}[\Delta Y_1(0,0) \mid X_0, X_1, D_1 = 0, D_2 = 0]$$

But also

$$\mathbb{E}[\Delta Y_1(0,0) \mid X_0, X_1, X_2, D_1 = 0, D_2 = 1] = \mathbb{E}[\Delta Y_1(0,0) \mid X_0, X_1, X_2, D_1 = 0, D_2 = 0]$$

We get parallel pre-trends regardless of whether we control for X_2 or not. X_2 is a neutral control for the pre-trends.

Valid Adjustment Sets

To generalize these results, we define

Valid Adjustment Set (VAS) for $ATT(g,t)$: a covariate set $\mathbf{Z}$ such that

$$
\begin{aligned}
ATT(g,t) &:= \mathbb{E}\left[Y_t(\overline{0}_{g-1}, \underline{1}_g) - Y_t(\overline{0}) \;\middle|\; \overline{D}_{g-1} = \overline{0}_{g-1},\, \underline{D}_g = \underline{1}_g\right] \\
&= \mathbb{E}\left[\Delta Y_{g-1,t} \;\middle|\; \overline{D}_{g-1} = \overline{0}_{g-1},\, \underline{D}_g = \underline{1}_g\right] \\
&\quad - \underbrace{\mathbb{E}\left[\mathbb{E}\left[\Delta Y_{g-1,t} \;\middle|\; \mathbf{Z},\, \overline{D} = \overline{0}\right] \;\middle|\; \overline{D}_{g-1} = \overline{0}_{g-1},\, \underline{D}_g = \underline{1}_g\right]}_{\textbf{covariate-adjusted never-treated trend}} \\
&=: DiD^{NT}_{g,t}(\mathbf{Z})
\end{aligned}
$$

In words: $\mathbf{Z}$ is a VAS if conditioning on it makes the covariate-adjusted never-treated trend a valid counterfactual for the treated group — so the conditional DiD recovers $ATT(g,t)$.

Which Covariates to Control For?

Our paper characterizes the **minimal valid adjustment set** under progressively stronger structural restrictions.

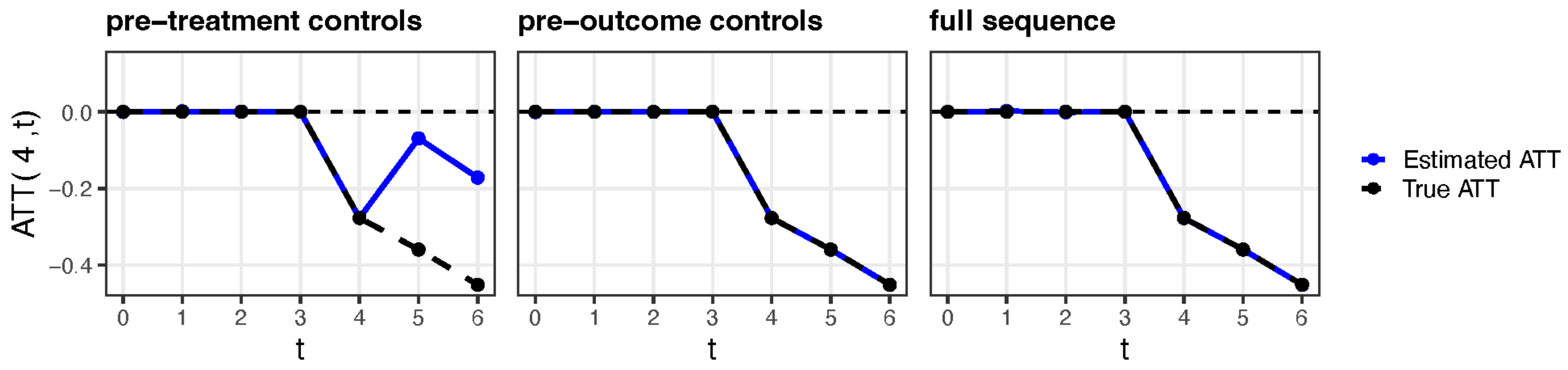
Structural assumptions					Minimal VAS for…		
SWAS + No Y-dyn	No D→X (within)	No D→X (feedback)	No X dynamics	No X→Y effect	Pre-trends	Short-term $ATT(g,g)$	Dynamic $ATT(g,t)$
✗	~	~	~	~	—	—	—
✓	✗	✗	✗	✗	$\overline{X}_{g-1}$	—	—
✓	✓	✗	✗	✗	$\overline{X}_{g-1}$	$\overline{X}_g$	—
✓	✓	✓	✗	✗	$\overline{X}_{g-1}$	$\overline{X}_g$	$\overline{X}_t$
✓	✗	✗	✓	✗	$X_t,\ X_{g-1}$	—	—
✓	✓	✗	✓	✗	$X_t,\ X_{g-1}$	$X_{g-1},\ X_g$	—
✓	✓	✓	✓	✗	$X_t,\ X_{g-1}$	$X_{g-1},\ X_g$	$X_{g-1},\ X_t$
✓	~	~	✓	✓	$\varnothing$	$\varnothing$	$\varnothing$

✓/✗ = restriction holds / does not hold; ~ = irrelevant; — = no VAS exists; $\overline{X}_t$ = all covariates up to period t; X_{g-1}, X_t = just these two period values.

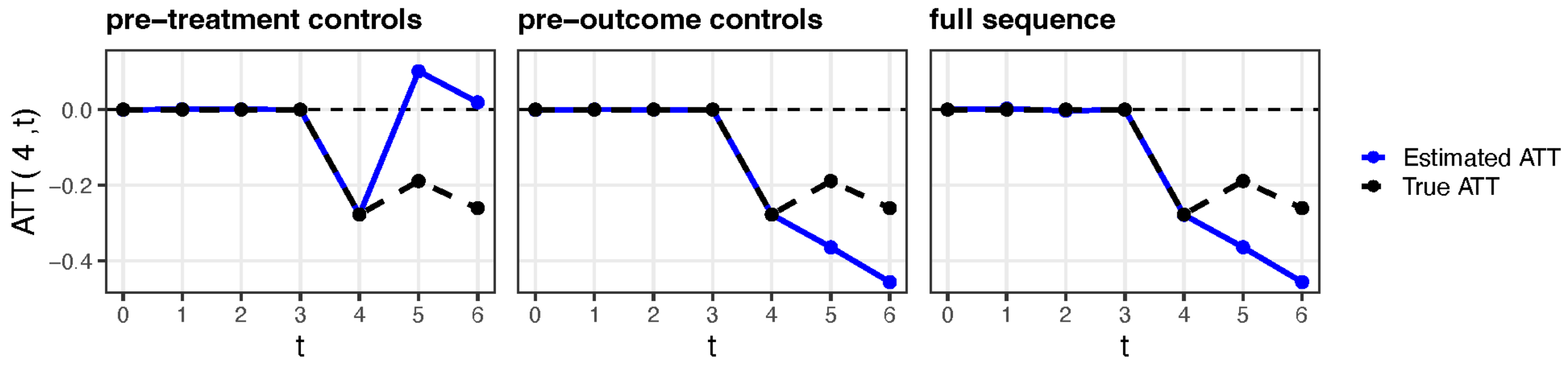
Additional Result: For cases where a minimal VAS exists, any set of covariates that contains the minimal VAS is a VAS. In particular, the full sequence $\overline{X}$ is the maximal VAS.

Illustration in Simulations

The event study plots below show conditional DiD estimates for the group first treated in period 4 using different sets of controls. The simulations build on DGPs that incorporate the restrictions in line 3 and 4 of the table respectively.



No treatment-covariate feedback



Treatment-covariate feedback

Pre-trends are parallel across all displayed (settings) × (sets of control variables).

Generalization

The simulation results illustrate the following insights:

- Controlling only for pre-treatment covariates yields parallel pre-trends and unbiased short-term effects but biased dynamic effects.
- Controlling for pre-outcome controls provides the same results as using the full sequence of covariates.
- In the absence of treatment–covariate feedback, strategies involving post-treatment variables are unbiased.
- In the presence of treatment–covariate feedback, all dynamic effect estimates are biased, reflecting omitted variable bias or “wrong world control bias.”
- Pre-trends do not diagnose post-treatment violations of CPT, while short-term effects remain unbiased even with treatment–covariate feedback.

What Can Pre-Trend Tests Tell Us?

Pre-trend tests ask: were treated and control groups moving in parallel **before** treatment?

What a rejection tells you

- Pre-trends fail with all pre-treatment covariates $\longrightarrow$ the most fundamental DiD assumptions are violated — no specification is credible
- Pre-trends fail only without covariates, but pass with them $\longrightarrow$ direct covariate–outcome effects are likely present

What pre-trend tests cannot detect

- Treatment-covariate feedback — the effect of treatment on future covariates unfolds *after* treatment starts. Pre-treatment periods are entirely silent about this.
- So: dynamic effect estimates can be biased even when all pre-trend tests pass

Important: Non-rejections may be due to a lack of power rather than validity of the underlying assumptions (Roth, 2022).

Practical Recommendations

1. Draw a causal graph: Before looking at data, sketching how treatment, covariates, and outcomes are connected can determine which covariates are valid controls and which are not.

2. Test pre-trends with all pre-treatment covariates: This is the most informative diagnostic. A rejection is fatal — it rules out all standard DiD specifications.

3. Run three benchmark specifications

- No time-varying controls
- Pre-outcome covariates (all X up to period t)
- Minimal pair: just X_{g-1} and X_t

Comparing results across these highlights which assumptions are driving the estimates.

4. Defend treatment-covariate feedback with economic arguments: Pre-trend tests cannot do this. With post-treatment covariates, substantive reasons why treatment does not affect them are needed.

Want more?

Find the paper on arXiv: arxiv.org/abs/2604.12818

Any feedback is welcome.

References

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics, Themed Issue: Treatment Effect 1*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Knaus, M. C., & Pfleiderer, H. (2026). *Causal Graphs for Conditional Parallel Trends*. arXiv. https://doi.org/10.48550/arXiv.2604.12818

Roth, J. (2022). Pretest with Caution: Event-Study Estimates after Testing for Parallel Trends. *American Economic Review: Insights*, *4*(3), 305–322. https://doi.org/10.1257/aeri.20210236