didint: A Stata package for Intersection Difference-in-differences

Sunny Karim and Matthew D. Webb
Package developed by Eric B. Jamieson
Carleton University

Canadian STATA Conference 2025

What is DID-INT?

- The Stata program was developed to implement Intersection Difference-in-differences (DID-INT).
- DID-INT was introduced by (Karim and Webb, 2024) in the paper *"Good Controls Gone Bad: Difference-in-differences With Covariates."*
- It helps estimate the **Average Treatment Effect on the Treated (ATT)** when the effect of covariates on outcome is different across regions and over time.
- DID-INT is applicable to both common and staggered adoption designs.
- DID-INT allows consistent estimates of the ATT using covariates whose effects on outcomes vary over time and by region.

Why covariates matter I

- Difference-in-differences (DID) is commonly used to evaluate policies/interventions implemented in a region.
- These policies are introduced in some regions (treated regions) and not in others (control regions).
    - Regions could be provinces in Canada, states in the US, cities, municipalities, health authorities, etc.
- DID looks at differences across treated and control regions and over time to measure the impact of policies.
- The method relies on the parallel trends assumption:
    - The trends between treated and control groups would have moved in a similar way (i.e. in parallel) without the policy.
- Parallel trends (without covariates) are unlikely to hold due to differences across regions.

Why covariates matter II

- Researchers use covariates to improve the plausibility of parallel trends, often referred to as the Conditional Parallel Trends (CPT).
- CPT is a relaxed version of the usual parallel trends assumption.
- It is known from DID Literature that time-varying covariates cause **methodological challenges**.
- Without additional assumptions, using time-varying covariates can cause **inconsistent estimates**.
- We introduce an additional assumption, previously implied in the literature called the **Common Causality of Covariates (CCC)** assumption.
    - The effect of covariates (X) on Outcome (Y) is the same across regions and time.
- We show that when the CCC is violated, the estimate of the ATT is inconsistent using traditional and modern DID methods.
- DID-INT uses previously hidden parallel trends when CCC is violated.

Common Causality of Covariates (Setup)

- Assume:
  - Two periods: 1 and 2,
  - Two regions: A and B.
  - B is treated in period 2.
- Also, let the untreated potential outcome of an individual i in region s at time t be:

$$Y_{i,s,t}(0) = \Theta^{\circ}_{s,t} + \gamma^{\circ}_{s,t} X_{i,s,t} + \epsilon_{i,s,t} \tag{1}$$

Common Causality of Covariates Assumption

- Two-way CCC implies:

$$\gamma^{\circ}_{A,1} = \gamma^{\circ}_{A,2} = \gamma^{\circ}_{B,1} = \gamma^{\circ}_{B,2} = \gamma \tag{2}$$

- Region-varying CCC implies:

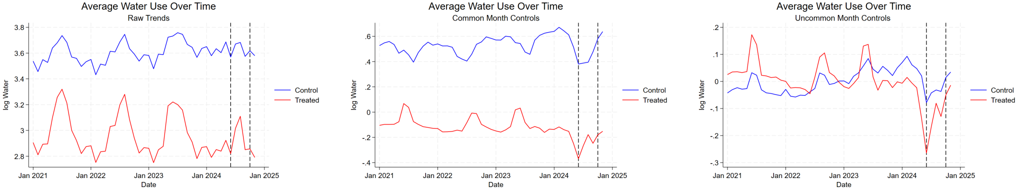
$$\gamma^{\circ}_{A,1} = \gamma^{\circ}_{B,1}; \gamma^{\circ}_{A,2} = \gamma^{\circ}_{B,2} \tag{3}$$

- Time-varying CCC implies:

$$\gamma^{\circ}_{A,1} = \gamma^{\circ}_{A,2}; \gamma^{\circ}_{B,1} = \gamma^{\circ}_{B,2} \tag{4}$$

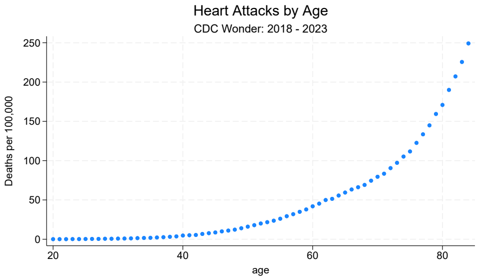
Motivating Example 1 - Uncommon Causality

- Example from Bray (2025) Bray (2025).
- The control group and the treated group have different seasonal trends.



- In this example:
  - Unconditional parallel trends do not hold
  - Conditional parallel trends do not hold, with a common covariate adjustment
  - Parallel trends seem plausible with group specific adjustment
  - *Research Question: Which estimators can accommodate this situation?*
  - **Answer:** Many existing estimators can't adapt to this.

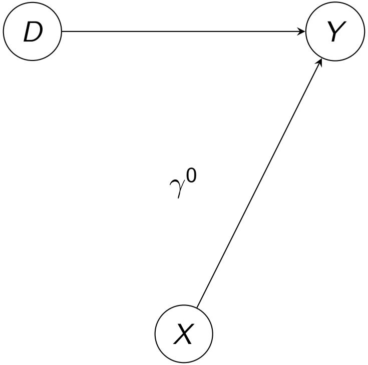
Example 2 - Time Varying Covariates



- Consider an intervention to reduce deaths from heart attacks

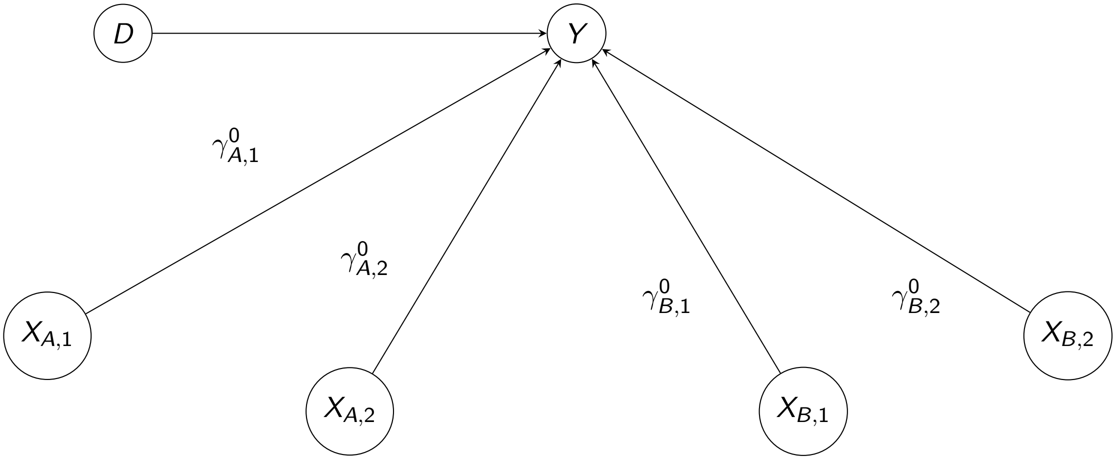
- We have a panel of individuals, say from 2018–2023
- Given the above, we want to include the age of the individual
    - **Problem:** Age is time varying
    - **Current Practice:** Use pre-treatment value of covariate
    - **Additional Problem:** This assumes that the relationship between the covariate and outcome is stable over time, clearly violated here
    - Age in 2010 is not the same as age in 2020

Illustration of Covariate Types: Good Controls



Good controls

Illustration of Covariate Types: Good Controls Gone Bad



Good controls gone bad

Intersection Difference-in-differences (DID-INT) I

- DID-INT is a multi-step semi-parametric method.
- Assume there are S regions and T periods.
- **Step 1:** We run the following regression:

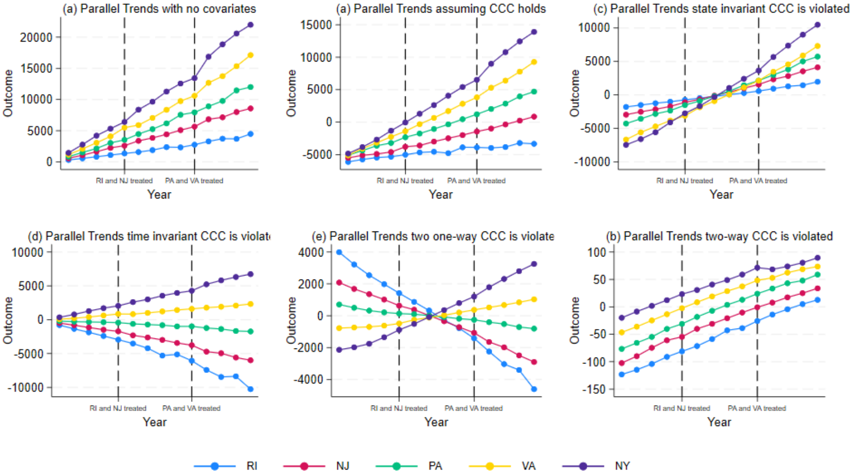
$$Y_{i,s,t} = \sum_{s \in S} \sum_{t=2}^{T} \lambda_{s,t} I(s,t) + f(X^k_{i,s,t}) + \epsilon_{i,s,t} \tag{5}$$

  - $I(s,t)$ is the group-by-time dummy (i.e. 1 if unit is in region s at time t).
  - $I(s)$ is a region dummy.
  - $I(t)$ is a time dummy.
  - k indexes covariates (K total).
- **Step 2:** Several "2×2" $ATT(s,t)$'s are estimated using the $\lambda_{s,t}$ terms.
  - DID-INT estimates ATTs **for each region**, rather than pooling regions by treatment timing (as in Callaway and Sant'Anna (2021) Callaway and Sant'Anna (2021)).

Intersection Difference-in-differences (DID-INT) II

  - Benefits of doing this is explored in more details in Karim et. al (2025) Karim et al. (2025a).
- **Step 3:** Overall ATT is calculated using a weighted average.
- Depending on the type of CCC violation, $f(X_{i,s,t}^{k})$ can take the following forms:
  - **Case 1:** $f(X_{i,s,t}^{k}) = \sum_{k=1}^{K} \gamma^{k} X_{i,s,t}^{k}$ if two-way CCC holds.
  - **Case 2:** $f(X_{i,s,t}^{k}) = \sum_{s=1}^{S^{T}} \sum_{k=1}^{K} \gamma_{s}^{k} I(s) X_{i,s,t}^{k}$ if region-varying CCC is violated.
  - **Case 3:** $f(X_{i,s,t}^{k}) = \sum_{t=1}^{T} \sum_{k=1}^{K} \gamma_{t}^{k} I(t) X_{i,s,t}^{k}$ if time-varying CCC is violated.
  - **Case 4:** $f(X_{i,s,t}) = \sum_{t=1}^{T} \sum_{s=1}^{S^{T}} \sum_{k=1}^{K} \gamma_{s,t}^{k} I(s) I(t) X_{i,s,t}^{k}$ if two-way CCC is violated.
  - **Case 5:** $f(X_{i,s,t}) = \sum_{s=1}^{S^{T}} \sum_{k=1}^{K} \gamma_{s,t}^{k} I(s) X_{i,s,t}^{k} + \sum_{t=1}^{T} \sum_{t=2}^{T} \sum_{k=1}^{K} \gamma_{s,t}^{k} I(t) X_{i,s,t}^{k}$ if two-way CCC is violated (but it enters linearly).
- Inference is done using Jackknife/Randomization Inference

Model Selection Algorithm



STATA Package

- **Requirements:**
  - `STATA` version 14.1 or later, `Julia` version 1.1.6 or later and David Roodman's `Julia` Package for STATA (v1.1.10).
- The wrapper uses the `julia` package (Roodman, 2025) which integrates the Julia programming language into `STATA`.
- The command help file can be accessed using: `help didintjl`
- DID-INT can be implemented using the following command:

```
didintjl outcome(string) state(string) time(string) treated_states(string) treatment_times(string) date
_format(string) [covariates(string) ccc(string) ref_column(string) ref_group(string) freq(string) freq_
multiplier(int 1) start_date(string) end_date(string) nperm(int 1000) verbose(int 1) seed(int 0) use_pr
e_controls(int 0)]
```

Syntax I

- `outcome` - Declares the dependent variable or the outcome of interest.
- `state` - Variable identifying region of observation.
- `time` - Variable identifying date of observation.
- `treated_states` A list of strings which lists the treated states.
  - Example: `treated_states` - should be entered as ("ON QC MN SK NL").
  - Other `states` - in variable state are treated as controls.
- `treated_times` - A list of strings of treatment time corresponding to the treated states.
  - Example: `treated_times` - should be entered as ("2001 2002 2003 2004 2005").
  - There is a one-to-one correspondence between the ordered lists.
  - So, ON is first treated in 2001, QC is first treated in 2002, and so on.
  - 2001 corresponds to ON, 2002 corresponds to QC.
- `date_format` - Variable used to define the date format used in the data.
  - Options include: yyyy, yyyy/mm/dd, ddmonyyyy, etc.

Syntax II

- `covariates` (optional) - lists covariates to be used. If omitted, covariates are not used
- `ccc` (optional) - Specifies which version of DID-INT to use.
  - Options include: hom for Case 1, state for Case 2, time for Case 3, int for Case 4 (default) and add for Case 5.
- agg (optional) - Used to specify the aggregation scheme to be used in Step 3.
  - Options include: cohort (default), state, simple, sgt, time, none.

Example Output

```
. use "MeritExampleDataDiDIntjl.dta", clear
. didintjl, outcome("coll") state("state") time("year") ///
 treated_states("34 57 58 59 61 64 71 72 85 88") ///
 treatment_times("2000 1998 1993 1997 1999 1996 1991 1998 1997 2000") ///
 date_format("yyyy") covariates("asian male black") ccc("int")
```

DiDInt.jl Results

Cohort	ATT	SE	p-val	JKNIFE SE	JKNIFE p-val	RI p-val
1991-01-01	0.0044675	0.040	0.911	0.044	0.920	0.947
1993-01-01	0.0414129	0.039	0.284	0.044	0.344	0.676
1996-01-01	0.0637732	0.047	0.172	0.058	0.272	0.442
1997-01-01	0.0883183	0.034	0.010	0.040	0.027	0.333
1998-01-01	0.0301035	0.056	0.593	0.067	0.654	0.711
1999-01-01	0.1940669	0.022	0.000	0.035	0.000	0.300
2000-01-01	-0.0161931	0.034	0.638	0.095	0.865	0.893

```
Aggregation Method: Cohort

Aggregate Results:
Aggregate ATT: .05110589
Standard error: .01691945
p-value: .02338082
Jackknife SE: .02094461
Jackknife p-value: .05046831
RI p-value: .1911912
```

Code

- Code and example dataset can be found at: https://github.com/ebjamieson97/didintjl
- The Julia package can be found at: https://github.com/ebjamieson97/DiDInt.jl

Related Papers

- DID-INT is fully described in Karim and Webb (2024)
- Karim et al. (2025b) describes a related procedure for unpoolable data
- Cluster robust inference for DID-INT, and state level treatment effects are considering in Karim et al. (2025a)
  - Cluster robust inference is done using either randomization inference (RI β) (MacKinnon and Webb, 2020) or the jackknife (MacKinnon et al., 2023; MacKinnon et al., 2025)

References I

Bray, B. (2025). Estimating the effectiveness of the city of calgary's response to the bearspaw south feeder main rupture. mimeo.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Karim, S., Strumpf, E., Austin, N., and Webb, M. (2025a). Which policy works and where? estimation and inference of state level treatment effects using difference-in-differences. In *Slides presented at the Canadian Economics Association Annual Conference, accessed July*, volume 11, page 2025.

Karim, S. and Webb, M. D. (2024). Good controls gone bad: Difference-in-differences with covariates. *arXiv preprint arXiv:2412.14447*.

Karim, S., Webb, M. D., Austin, N., and Strumpf, E. (2025b). Difference-in-differences with unpoolable data.

MacKinnon, J. G., Nielsen, M. Ø., and Webb, M. D. (2023). Fast and reliable jackknife and bootstrap methods for cluster-robust inference. *Journal of Applied Econometrics*, 38:671–694.

References II

MacKinnon, J. G., Nielsen, M. Webb, M. D., and Karim, S. (2025). Improving inferences for callaway and sant'anna did using the cluster jackknife. mimeo.

MacKinnon, J. G. and Webb, M. D. (2020). Randomization inference for difference-in-differences with few treated clusters. *Journal of Econometrics*, 218(2):435–450.

Roodman, D. (2025). Julia as a universal platform for statistical software development. *The Stata Journal*, 25(2):255–284.