

 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

**IAPRI-MSU Technical Training**

**Difference-in-differences**

Facilitated by Nicole M. Mason  
Michigan State University  
Department of Agricultural, Food, & Resource Economics

1 March 2018  
Indaba Agricultural Policy Research Institute  
Lusaka, Zambia

 **USAID**  
FROM THE AMERICAN PEOPLE

 INNOVATION LAB FOR  
FOOD SECURITY POLICY

 MICHIGAN STATE  
UNIVERSITY

 IAPRI

 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

**Recall from July/September trainings on  
introduction to impact evaluation (1)**

“An impact evaluation **assesses changes in the well-being** of individuals, households, communities or firms **that can be attributed** to a particular project, program or policy” Source: [World Bank](#)

 **USAID**  
FROM THE AMERICAN PEOPLE

 INNOVATION LAB FOR  
FOOD SECURITY POLICY

 MICHIGAN STATE  
UNIVERSITY

 IAPRI



The U.S. Government's Global Hunger & Food Security Initiative

## Recall from July/September trainings on introduction to impact evaluation (2)

- We did a brief overview of common methods of impact evaluation (IE)
  - Randomized evaluation
  - Propensity Score Matching
  - Difference-in-Differences
  - Instrumental Variables
  - Regression Discontinuity
- Today we'll focus on difference-in-differences
  - Reminder on basic concepts/theory
  - Applications in Stata







The U.S. Government's Global Hunger & Food Security Initiative

## Learning objectives

- By the end of today's session, you should be able to:
  1. Understand the key assumption of DID models
  2. Write down the regression equation for a DID model
  3. Identify which parameter in the regression equation represents the causal effect of interest
  4. Estimate a DID model in Stata and interpret the results






 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

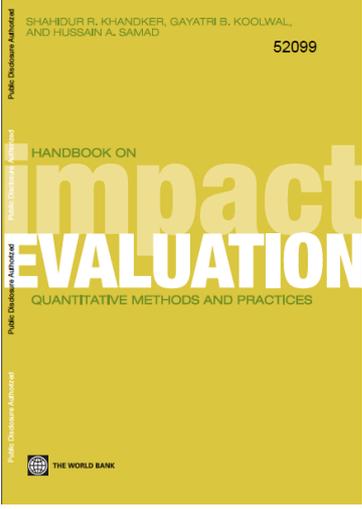
## References

**The training materials for this session (developed by Nicole Mason) draw heavily on:**

- Khandker, S.R., Koolwal, G.B. and Samad, H.A., 2010. *Handbook on impact evaluation: quantitative methods and practices*. Washington, DC: World Bank Publications. Available [here](#).

**Other suggested readings and references - DID sections in:**

- Angrist, J. D., & Pischke, J. S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Angrist, J. D., & Pischke, J. S. (2015). *Mastering 'metrics: The path from cause to effect*. Princeton University Press.
- Imbens, G. W., & Wooldridge, J. (2007). What's new in econometrics? Difference-in differences estimation, Lecture notes 10 for NBER summer 2007.
- Ravallion, M. (2008). Evaluating anti-poverty programs. *Handbook of development economics*, 4, 3787-3846.



 **USAID**  
FROM THE AMERICAN PEOPLE

 INNOVATION LAB FOR  
FOOD SECURITY POLICY

 MICHIGAN STATE  
UNIVERSITY

 IAPRI

 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

## REVIEW: With vs. Without

- The key comparison we want to make in IE is between outcomes **WITH VS. WITHOUT** the intervention (project/program/policy)
- Impact = "With" outcome – "without" outcome**

 **USAID**  
FROM THE AMERICAN PEOPLE

 INNOVATION LAB FOR  
FOOD SECURITY POLICY

 MICHIGAN STATE  
UNIVERSITY

 IAPRI

**FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

## Impact=With-Without

Participants  
Control  
Counterfactual  
Program

Income  
Time

Source: Khandker et al. (2010)

Participants' income **WITH** the program?  
•  $Y_4$   
Participants' income **WITHOUT** the program (counterfactual income)?  
•  $Y_2$   
Program impact?  
•  $Y_4 - Y_2$

“The **key challenge** in impact evaluation is finding a group of people who did not participate, but closely resemble the participants had those participants not received the program. Measuring outcomes in this **comparison group** is as close as we can get to measuring ‘how participants would have been otherwise.’”  
[J-PAL Introduction to Evaluations](#)

**USAID** FROM THE AMERICAN PEOPLE INNOVATION LAB FOR FOOD SECURITY POLICY MICHIGAN STATE UNIVERSITY IAPRI

**FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

## Introducing Difference-in-Differences (DID)

- Who has used DID before and what were you studying?
- What is the DID approach to constructing a comparison group / approximating the counterfactual, and how is the DID treatment effect calculated?
  - “The DID estimator relies on a comparison of participants and non-participants before and after the intervention” (Khandker et al. 2010, p. 72)
  - DID Impact=(avg.  $\Delta Y$  participants)-(avg.  $\Delta Y$  non-participants)
    - (avg.  $Y^T$  after - avg.  $Y^T$  before) – (avg.  $Y^c$  after - avg.  $Y^c$  before)
    - → why it's called difference-in-differences or double difference

**USAID** FROM THE AMERICAN PEOPLE INNOVATION LAB FOR FOOD SECURITY POLICY MICHIGAN STATE UNIVERSITY IAPRI

**FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

## DID – visual representation

**DID impact = (avg.  $Y^T$  after - avg.  $Y^T$  before) – (avg.  $Y^c$  after - avg.  $Y^c$  before)**

*Change in participants' income?*

- $Y_4 - Y_0$

*Change in non-participants' (control) income?*

- $Y_3 - Y_1$

*DID impact?*

- $(Y_4 - Y_0) - (Y_3 - Y_1)$

Source: Khandker et al. (2010)

## DID key assumption: parallel (common) trends

**Parallel trends** = “unobserved characteristics affecting program participation do not vary over time with treatment status” (Khandker et al. 2010, p. 73)

- I.e., trends in the outcome variable would be the same in the two groups without treatment (Angrist & Pischke 2009); or
- “...treatment and control outcomes move in parallel in the absence of treatments” (Angrist & Pischke (2015, p. 178)
- Implies  $(Y_1 - Y_0) = (Y_3 - Y_2)$

*Change in participants' income (after – before)?*

- $Y_4 - Y_0$

*Change in non-participants' (control) income (after – before)?*

- $Y_3 - Y_1$

*DID impact?*

- $(Y_4 - Y_0) - (Y_3 - Y_1)$
- $= Y_4 - Y_0 - Y_3 + Y_1 = Y_4 - Y_3 + Y_1 - Y_0$

Substitute in  $(Y_1 - Y_0) = (Y_3 - Y_2)$  (parallel trend assumption):

- $= Y_4 - Y_3 + Y_3 - Y_2 = Y_4 - Y_2$

**Same as with vs. without!**

Source: Khandker et al. (2010)



## What happens if trends are not parallel?

- DID estimate is biased
- → See next 2 slides and handout based on figures in Ravallion (2008) for illustration



**USAID**  
FROM THE AMERICAN PEOPLE

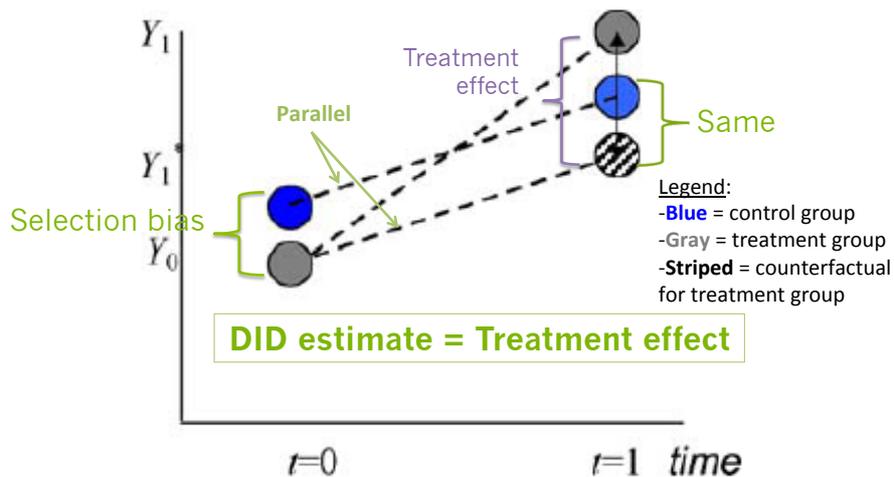


INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY

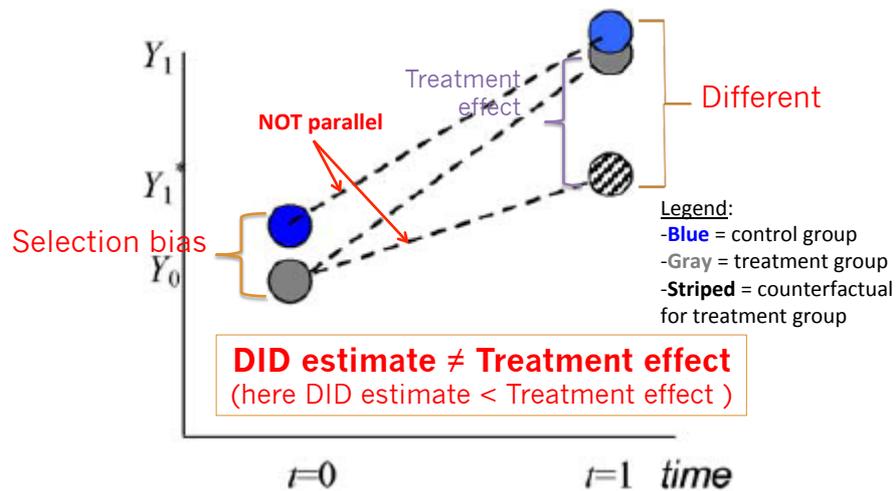


## Key assumption for DID: Parallel trends



Source: Ravallion (2008)

## Non-parallel trends cause DID to be biased



**FEED THE FUTURE**

The U.S. Government's Global Hunger & Food Security Initiative

## Can partially test for parallel trends if have multiple pre-treatment waves of data

### EX) Mason et al. (2017) study on effects of Kenya's NAAIAP input subsidy program on HH behavior and welfare

- Use 3 waves of data:
  - 2 waves prior to implementation of NAAIAP (2004 & 2007 TAPRA surveys)
  - 1 wave after (during) implementation (2010 TAPRA survey)
- Regress change in outcome variable (2007 minus 2004) on dummy for if HH was NAAIAP participant in 2010 ("Treated") and baseline (2004) controls
- While no guarantee that trends would have been parallel 2007 to 2010 in absence of treatment, if "Treated" dummy isn't stat. sig. above, it provides some evidence in support of parallel trends assumption



**USAID**  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

## DID simple numerical example

**DID impact = (avg.  $Y^T$  after - avg.  $Y^T$  before) – (avg.  $Y^c$  after - avg.  $Y^c$  before)**

The following table gives mean income during the pre- and postintervention period for a microfinance intervention in the rural Lao People's Democratic Republic:

	<i>Mean income (KN thousand)</i>	
	<i>Participants</i>	<i>Nonparticipants</i>
Preintervention period	80	90
Postintervention period	125	120

Impact of microfinance intervention on participants' income using DD is

(a) KN 45,000  
(b) KN 30,000  
(c) KN 15,000

Source: Khandker et al. (2010)

 **USAID** FROM THE AMERICAN PEOPLE
  INNOVATION LAB FOR FOOD SECURITY POLICY
  MICHIGAN STATE UNIVERSITY
  IAPRI

 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

## DID data requirements

- **Repeated cross-sections**
  - Separate random samples from the relevant population before and after the project/program/policy change
- OR
- **Panel data**
  - Random sample from the relevant population and data from the same cross-sectional units before and after the project/program/policy change

 **USAID** FROM THE AMERICAN PEOPLE
  INNOVATION LAB FOR FOOD SECURITY POLICY
  MICHIGAN STATE UNIVERSITY
  IAPRI



## Regression DID – Basic Setup

$$Y_{it} = \alpha + \gamma \text{Treated}_i + \lambda \text{After}_t + \delta (\text{Treated}_i \times \text{After}_t) + \varepsilon_{it}$$

Where:

- $i$  indexes the cross-sectional unit and  $t$  indexes time
- $\text{Treated} = 1$  if unit is ultimately treated (exposed to project/program/policy change), = 0 o.w. (specified as a time-constant variable)
- $\text{After} = 1$  if time period is after the project/program/policy change, = 0 before (changes over time but not across units in the dataset)
- $\text{Treated} \times \text{After}$  is the interaction of these two variables
- \*Note: Notation above is for when “treatment” or the project/program/policy change is at the same level as the outcome variable. We’ll look at higher level changes next.
- *Which parameter represents the causal effect of interest (assuming the key assumptions hold)?*  $\delta$  (the parameter on the Treated X After term)



**USAID**  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



## Regression DID – Basic Setup - with higher level project/program/policy change

Suppose the project/program/ policy change is at the district ( $d$ ) level but you have data at the household level ( $i$ ). Then the notation would be:

$$Y_{idt} = \alpha + \gamma \text{Treated}_{id} + \lambda \text{After}_t + \delta (\text{Treated}_{id} \times \text{After}_t) + \varepsilon_{idt}$$

- *Which parameter represents the causal effect of interest (assuming the key assumptions hold)?*  $\delta$  (the parameter on the Treated X After term)
- This is the more common instance in which DID is used
- Would want to cluster your standard errors at the district level



**USAID**  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY





The U.S. Government's Global Hunger & Food Security Initiative

## Regression DID – Basic Setup – with covariates

Can also control for additional covariates:

$$Y_{idt} = \alpha + \gamma \text{Treated}_d + \lambda \text{After}_t + \delta (\text{Treated}_d \times \text{After}_t) + \mathbf{X}_{idt} \boldsymbol{\beta} + \varepsilon_{idt}$$

- Bold represents vectors
- Which parameter represents the causal effect of interest (assuming the key assumptions hold)?  $\delta$  (the parameter on the Treated X After term)



USAID  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



The U.S. Government's Global Hunger & Food Security Initiative

Source: Imbens & Wooldridge (2007) with minor changes to notation

## Panel FE setup without control variables

EX) 2 periods:

$$Y_{it} = \alpha + \lambda \text{After}_t + \delta \text{Treated}_i + c_i + u_{it}$$

- First difference to remove  $c_i$  (or estimate via FE)

$$\Delta Y_i = \lambda + \delta \Delta \text{Treated}_i + \Delta u_i$$

- Which parameter is the causal effect of interest (assuming the key assumptions hold)?
- $\delta$



USAID  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

Source: Imbens & Wooldridge (2007) with minor changes to notation

**Panel FE setup with control variables**

**EX) 2 or more periods**

$$Y_{it} = \alpha + \mathbf{Year}_t \boldsymbol{\lambda} + \delta \text{Treated}_{it} + \mathbf{X}_{it} \boldsymbol{\beta} + c_i + u_{it}$$

- Where **Year** is a vector of year dummies
- Estimate via FE
- *Which parameter is the causal effect of interest (assuming the key assumptions hold)?*
  - $\delta$

 **USAID**  
FROM THE AMERICAN PEOPLE

 INNOVATION LAB FOR  
FOOD SECURITY POLICY

 MICHIGAN STATE  
UNIVERSITY

 IAPRI

 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

**Examples & tweaking the variable names/  
notation to fit your particular situation (1)**

General:  $Y_{it} = \alpha + \gamma \text{Treated}_i + \lambda \text{After}_t + \delta (\text{Treated}_i \times \text{After}_t) + \varepsilon_{it}$

**Example: Wooldridge (2002) - new garbage incinerator and effects on housing values in a city in Massachusetts**

- New garbage incinerator construction began in 1981
- Have housing value data from 1978 and 1981 plus info on distance from house to new incinerator (let "rprice" be the home value in real US\$)
- Let "nearinc" = 1 if house is near (within 3 miles/4.83 km) incinerator, =0 o.w. (far from incinerator)
- Let "y81" = 1 if year is 1981, and =0 o.w. (year is 1978)
- *How would you write the DID regression equation in this case? What is the key parameter of interest?*

Zambia example: Proximity to farm block & land values

 **USAID**  
FROM THE AMERICAN PEOPLE

 INNOVATION LAB FOR  
FOOD SECURITY POLICY

 MICHIGAN STATE  
UNIVERSITY

 IAPRI



## Examples & tweaking the variable names/ notation to fit your particular situation (1)

**General:**  $Y_{it} = \alpha + \gamma \text{Treated}_i + \lambda \text{After}_t + \delta(\text{Treated}_i \times \text{After}_t) + \varepsilon_{it}$

**Example: Wooldridge (2002) - new garbage incinerator and effects on housing values in a city in Massachusetts**

**Specific:**  $rprice_{it} = \alpha + \gamma \text{nearinc}_i + \lambda \gamma 81_t + \delta(\text{nearinc}_i \times \gamma 81_t) + \varepsilon_{it}$

*Which parameter captures the effect of interest?  $\delta!$*

- Where
  - “rprice” is the home value in real US\$
  - “nearinc”= 1 if house is near (within 3 miles/4.83 km) incinerator, =0 o.w. (far from incinerator)
  - “ $\gamma 81$ ” = 1 if year is 1981, and =0 o.w. (year is 1978)








## Examples & tweaking the variable names/ notation to fit your particular situation (2)

**General:**  $Y_{it} = \alpha + \gamma \text{Treated}_i + \lambda \text{After}_t + \delta(\text{Treated}_i \times \text{After}_t) + \varepsilon_{it}$

**Example: Angrist & Pischke (2009) – understanding cholera transmission in mid-1800s**

- Have district-level data from London on death rates and water company in 1849 & 1852. Let:
  - “deathr” be the death rate in district d
  - “Lambeth”= 1 if district d gets its water from the Lambeth Company (which moved its water works to a cleaner section of the Thames River in 1852), = 0 if the district gets its water from the Southwark and Vauxhall Company
  - “ $\gamma 1852$ ” = 1 if year is 1852, and =0 o.w. (year is 1849)
- Write down the DID equation for this scenario and using these variables

John Snow (British physician) believed to be the pioneer of the DID idea!

**Specific:**  $\text{deathr}_{dt} = \alpha + \gamma \text{Lambeth}_d + \lambda \gamma 1852_t + \delta(\text{Lambeth}_d \times \gamma 1852_t) + \varepsilon_{dt}$






 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

## Examples & tweaking the variable names/ notation to fit your particular situation (3)

General:  $Y_{idt} = \alpha + \gamma \text{Treated}_d + \lambda \text{After}_t + \delta(\text{Treated}_d \times \text{After}_t) + \epsilon_{idt}$

Example: Angrist & Pischke (2009) – effect of ↑ min. wage on fast food employment

- Have data from neighboring states (NJ & PA). Both states have \$4.25 minimum wage in early 1992 but NJ raises minimum wage to \$5.05 in April 1992. You have employment data from individual fast food restaurants (i) in each state (s) before (Feb. 1992) and after (Nov. 1992) the policy change. Let:
  - “employ” be the employment level of each restaurant
  - “NJ” = 1 if state is NJ, = 0 if state is PA
  - “Nov92” = 1 if time is November 1992, and = 0 o.w. (time is February 1992)
- *Write down the DID equation for this scenario and using these variables. Think carefully about which subscripts to put on each variable.*

Specific:  $\text{employ}_{ist} = \alpha + \gamma \text{NJ}_s + \lambda \text{Nov92}_t + \delta(\text{NJ}_s \times \text{Nov92}_t) + \epsilon_{ist}$

 **USAID**  
FROM THE AMERICAN PEOPLE

 INNOVATION LAB FOR  
FOOD SECURITY POLICY

 MICHIGAN STATE  
UNIVERSITY

 IAPRI

 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

## DID good to consider if have natural experiment

**EX) Dillon, B. (2016). Selling crops early to pay for school: A large-scale natural experiment in Malawi. Working Paper No. 243. Abidjan, Côte d'Ivoire: African Development Bank.**

- Big picture question: why do many HHs sell low, buy high (w.r.t. crop prices)?
- Hypothesis: “short-term expenditure needs force poor households to sell crops early, when output prices are well below their peak” (p. 4). I.e., “farming households that are credit-constrained sell crops early to finance immediate needs” (p. 12)
- Natural experiment: Malawi changed primary school calendar
  - 2009: start in December. 2010: start in September.
  - HHs have to make school-related expenditures much earlier in 2010 than 2009. HHs with school-aged children are the main ones we expect to change their behavior in response to the school calendar change.

 **USAID**  
FROM THE AMERICAN PEOPLE

 INNOVATION LAB FOR  
FOOD SECURITY POLICY

 MICHIGAN STATE  
UNIVERSITY

 IAPRI



The U.S. Government's Global Hunger & Food Security Initiative

## DID & natural experiment example (cont'd)

**Dillon, B. (2016). Selling crops early to pay for school: A large-scale natural experiment in Malawi. Working Paper No. 243. Abidjan, Côte d'Ivoire: African Development Bank.**

### DID regression:

General:  $Y_{it} = \alpha + \gamma \text{Treated}_i + \lambda \text{After}_t + \delta(\text{Treated}_i \times \text{After}_t) + \epsilon_{it}$

Specific:  $\text{Cropsales}_{it} = \alpha + \gamma \text{Children}_i + \lambda y_{2010_t} + \delta(\text{Children}_i \times y_{2010_t}) + \epsilon_{it}$

Where

- Cropsales = cumulative value of HH crop sales through August of year t
- Children = # of children in primary school (0, 1, 2, 3, etc.). Or could do 0/1
- y2010 = 1 if year is 2010; =0 if year is 2009
- Estimate separately for HHs above vs. below the poverty line



USAID  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



The U.S. Government's Global Hunger & Food Security Initiative

## What other situations/examples appropriate for DID can you think of?

- *And how would you set up your DID regression equation?*
- General:  $Y = \alpha + \gamma \text{Treated} + \lambda \text{After} + \delta(\text{Treated} \times \text{After}) + \epsilon$



USAID  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

Paraphrased from Khandker et al. (2010)

## DID vs. other methods

- **Key difference between PSM and DID:** **PSM** assumes **selection on observables only**, **DID** allows selection to be a function of **time-constant unobserved factors** (a.k.a. time invariant unobserved heterogeneity)
  - *Where have you heard this term before?*
  - *What if selection is a function of time-varying unobservables?*
- **Another key difference:**
  - **Randomized evaluations & PSM** – cross-sectional data sufficient (although panel data better – baseline/endline)
  - **Need repeated cross sections or panel data for DID**

 **USAID** FROM THE AMERICAN PEOPLE  INNOVATION LAB FOR FOOD SECURITY POLICY  MICHIGAN STATE UNIVERSITY  IAPRI

 **FEED THE FUTURE**  
The U.S. Government's Global Hunger & Food Security Initiative

## Theory wrap-up

(Source: Angrist & Pischke 2015, pp. 203-204)

**Master Stevefu:** Wrap it up for me, Grasshopper.

**Grasshopper:** Treatment and control groups may differ in the absence of treatment, yet move in parallel. This pattern opens the door to DD estimation of causal effects.

**Master Stevefu:** Why is DD better than simply two-group comparisons?

**Grasshopper:** Comparing changes instead of levels, we eliminate fixed differences between groups that might otherwise generate omitted variables bias.

...

**Master Stevefu:** On what does the fate of DD estimates turn?

**Grasshopper:** Parallel trends, the claim that in the absence of treatment, treatment and control group outcomes would indeed move in parallel.”

 **USAID** FROM THE AMERICAN PEOPLE  INNOVATION LAB FOR FOOD SECURITY POLICY  MICHIGAN STATE UNIVERSITY  IAPRI



The U.S. Government's Global Hunger & Food Security Initiative

## In Stata – DID (panel or pooled cross-sections)

General:  $Y_{it} = \alpha + \gamma \text{Treated}_i + \lambda \text{After}_t + \delta(\text{Treated}_i \times \text{After}_t) + \varepsilon_{it}$

### In Stata?

reg Y i.Treated i.After i.Treated#i.After (where “Treated” here is time-constant)

*For panel data or if have repeated cross-sections and policy change is at higher level than data, consider clustering s.e.'s (see Angrist & Pischke 2015 DID chapter for details)*

*Which coefficient is the DID estimate of the causal effect of interest (assuming the key assumptions hold)?*

- The one on the i.Treated#i.After variable



USAID  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



The U.S. Government's Global Hunger & Food Security Initiative

## In Stata – FE (panel data)

General:  $Y_{it} = \alpha + \lambda \text{After}_t + \delta \text{Treated}_{it} + c_i + u_{it}$

### In Stata?

xtreg Y i.After i.Treated, fe (where “Treated” here is time-varying)

*Consider clustering s.e.'s (see Angrist & Pischke 2015 DID chapter for details)*

*Which coefficient is the FE estimate of the causal effect of interest (assuming the key assumptions hold)?*

- The one on the i.Treated variable



USAID  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY





The U.S. Government's Global Hunger & Food Security Initiative

## Stata exercises

### 1. Wooldridge (2002) new garbage incinerator & housing values

- a. Use KIELMC.DTA in “data” folder to estimate:  

$$rprice_{it} = \alpha + \gamma nearinc_i + \lambda y81_t + \delta(nearinc_i \times y81_t) + \epsilon_{it}$$
 Recall: reg Y i.Treated i.After i.Treated#i.After
- b. Interpret the key coefficient of interest



USAID  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



The U.S. Government's Global Hunger & Food Security Initiative

## Stata exercises

### 2a. Khandker et al. (2010) example on effects of a microcredit program on HH welfare (expenditure) – DID

Data from Bangladesh in 1991 & 1998 on log total HH expenditure per capita (lexptot) & participation in microcredit program in 1998 (dfmfd98). Assume no microcredit program in 1991.

- a. Use hh\_9198.dta in “data/Khandker et al 2010 data files” folder to estimate:  

$$lexptot_{it} = \alpha + \gamma dfmfd98_i + \lambda year_t + \delta(dfmfd98_i \times year_t) + \epsilon_{it}$$
 (where year=1 if 1998, =0 if 1991)  
 Recall: reg Y i.Treated i.After i.Treated#i.After
- b. Interpret the key coefficient of interest



USAID  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY





## Stata exercises

### 2b. Khandker et al. (2010) example on effects of a microcredit program on HH welfare (expenditure) – FE

These data are actually HH panel survey data. Now estimate via FE instead. Time-varying variable for participation in microcredit program is `dfmfdyr`. Recall for HH panel data, can write FE model as:

$$Y_{it} = \alpha + \lambda \text{After}_t + \delta \text{Treated}_{it} + c_i + u_{it}$$

- a. Continue using the same dataset and estimate via FE:

$$\text{lexptot}_{it} = \alpha + \lambda \text{year}_t + \delta \text{dfmfdyr}_{it} + c_i + u_{it}$$

Recall: `xtreg Y i.After i.Treated, fe`

- b. Compare the key coefficient estimates between DID & FE



**USAID**  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



Thank you for your attention & participation!

Nicole M. Mason ([masonn@msu.edu](mailto:masonn@msu.edu))

Assistant Professor

Department of Agricultural, Food, & Resource Economics

Michigan State University



**USAID**  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY





## Acknowledgements

This training was supported by funding from the U.S. Agency for International Development through the Zambia Mission Buy-In to the Innovation Lab for Food Security Policy.



**USAID**  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY



# FEED THE FUTURE

The U.S. Government's Global Hunger & Food Security Initiative

[www.feedthefuture.gov](http://www.feedthefuture.gov)



**USAID**  
FROM THE AMERICAN PEOPLE



INNOVATION LAB FOR  
FOOD SECURITY POLICY

MICHIGAN STATE  
UNIVERSITY

