Request that you should not refuse

- PLEASE SWITCH OFF AND PUT AWAY YOUR CELL PHONES
- LAPTOPS OK IF WORK IS ACADEMIC
- REMOVE BAGS AND OTHER MATERIALS THAT CAN CAUSE DISTRACTION
- STOP HAVING SIDE CONVERSATIONS
- PARTICIPATE IN CLASS

Class 3
Natural Disasters Implications - basic labor market model
Regressions and Experiments: Internal/External Validity – Lemar, Rosenbaum

- **Read for Wednesday's Class**
- **2.1 - Benefits of Immigration - Borjas**
- **Read in the upcoming Weekend**
- **2.2 - Effect of Immigration on Wages**
- Concepts you should be ready with: Worker's Surplus, Firm's Surplus; Shape of Demand and Supply Curves of Labor with alternative elasticities

Implications of the basic labor demand supply model

$$w = A - Bh^d \ldots(1)$$

$$w = C + Dh^s \ldots(2)$$

$$\exists \; w = w^*$$

$$\ni h^d = h^s = h^*$$

- Existence, Stability, Uniqueness of the equilibrium
- Competitive wages protect demanders **and** suppliers
- Competitive wages are different from different sectors
- Competitive wages can be different for jobs in the same sector
- **Full Information** versus **Imperfect information** models
- Why Pay Gap?

Introducing Randomized Control Trial (RCT)

X: Medicine

Y: Red Blood Cell Count

Want to Prove that

X ⟶ Y

To Eliminate other Factors

1) Choose Treatment and Control roughly having similar characteristics
2) **Randomize Intervention** between Treatment and Control
3) Calculate the **average difference** of outcomes of two groups after repeated experiments

Single difference

	Red Blood Cells in Treatment Group	Red Blood Cells in Control Group
After repeated trials	T2=21	C2=10
Single Difference Estimate = (Average Effect on Treatment Group) – (Average effect on Control Group) = (T2 –C2)		

The treatment is the pill to boost red blood cells

Difference in difference

After Repeated Trials	Red Blood Cells in Treatment Group	Red Blood Cells in Control Group
Before the Treatment	**T1=2**	**C1=2.3**
After the Treatment	**T2=7**	**C2=3**
Difference = After - Before	**T2-T1=7-2=+5**	**C2-C1=3-2.3=+0.7**
Difference in Difference Estimate (DD/DID) = (Difference for Treatment) – (Difference for Control) = **(T2-T1) – (C2-C1)** **=5 – 0.7 =+4.3**		

- What is parallel assumption?

Natural Disasters Paper:

Context: Mariel Boatlift paper (David Card): No impact on employment and wages in the way suggested by theory

Unit of Analysis Matters (county versus larger geograpahy)

Limit yourself to Eq 1-E10 & Table 1, 2 & 3

Analyzing the paper: Questions to Answer

- What are the dependent variable (s)?
- What is the main independent/explanatory variable?
- What are the control variables?
- What is the unit of analysis?
- What is the strategy to identify the effect of hurricanes? [Fixed Effects & DD/GDD: comparing neighboring counties before and after an exogenous shock]
- What are the sources of data? Are you happy with how the variables are measured ? How you got the data, what is the measurement process, secondary or primary data, construct validity??
- Which parameters from the regression are important to make the author's point?
- Did the results match the author's predictions/intended results?
- What is the purpose of Table 5-8? [robustness]
- What is the academic importance of this paper? What is its contribution?
- What can be some of the big issues with this paper? [Think: if this would be an experiment (RCT) which assumptions may be violated?]

Hurricane	Category	Monetary Damage to Florida (millions)	Number of Deaths in Florida	Windspeed at Landfall (mph)	Rainfall (inches)
Florence (1988)	1	$0.6	0	75	5–10
Andrew (1992)	5	$43,000	44	175	5–7
Allison (1995)	1	$1.2	0	75	4–6
Erin (1995)	1	$0.5	6	87	5–12
Opal (1995)	3	$4,400	1	115	5–10
Danny (1997)	1	$100 (total to U.S.)	0	80	2–7
Earl (1998)	1	$64.5	2	92	6–16
Georges (1998)	2	$392	0	103	8–25
Irene (1999)	1	$1,100	8	75	10–20
Gordon (2000)	1	$11.9	1	75	3–5
Charley (2004)	4	$15,100	29	150	5–8
Frances (2004)	2	$8,900	37	105	10–20
Ivan (2004)	2	$8,100	19	130	7–15
Jeanne (2004)	3	$6,900 (total to U.S.)	3	121	8–13
Dennis (2005)	3	$2,200	14	120	10–15
Katrina (2005)	1	$115,000 (total to U.S.)	14	81	5–15
Ophelia (2005)	1	$70 (total to U.S.)	1	80	3–5
Rita (2005)	1	$10,000 (total to U.S.)	2	80	2–4
Wilma (2005)	3	$12,200	35	120	7–12

Although we can predict with confidence that the hurricane season will generate some hurricanes and that Florida will likely be hit by some of these hurricanes during the course of a decade, the exact timing and path of the hurricanes cannot be forecast. As a result, each of these hurricanes generates exogenous economic shocks to the Florida counties that are directly hit. The randomness of the path and intensity of the hurricane, therefore, provide a "natural experiment" that can be used to analyze how the economic shocks set off by such deadly storms alter labor market conditions. Because so many hurricanes have hit Florida in the past two decades, we can use the available data to estimate difference-in-differences models that examine the economic impact on the affected Florida counties relative to the economic events unfolding in the unaffected counties.

Supply shifts to the left

Demand can shift to the left or right

Difference in difference type model

After Repeated Trials	Employment/Wages in Counties with Hurricane	Employment/Wages in Counties without Hurricane
Before the Treatment	T1	C1
After the Treatment	T2	C2
Difference = After - Before	T2-T1	C2-C1
Difference in Difference Estimate (DD/DID) = (Difference for Treatment) – (Difference for Control) = (T2-T1) – (C2-C1) Fixed Effects= controlling for within variation		

- What is parallel assumption in this experiment?

Counties treated and not treated are following similar time paths with respect to the outcome variable

$$Q_{it} = f(Q_t, Z_i, S_t, H_{it}) + u_{it} \ldots.(1)$$

$$y_{it} = f(y_t, Z_i, S_t, H_{it}) + v_{it} \ldots.(2)$$

$$Q_{it} = \theta_{1i} Q_t + \theta_{2i} Z_i + \theta_{3i} Z_i t + \theta_{4i} S_t + \theta_{5i} H^{D}{}_{it} + \theta_{6i} H^{N}{}_{ijt} + u_{it} \ldots.(3)$$

$$y_{it} = \phi_{1i} y_t + \phi_{2i} Z_i + \phi_{3i} Z_i t + \phi_{4i} S_t + \phi_{5i} H^{D}{}_{it} + \phi_{6i} H^{N}{}_{ijt} + v_{it} \ldots.(4)$$

3 and 4 are the general forms of 1 and 2

Does Equation 3 and 4 look familiar?

$$Q_{it} = \theta_{1i} Q_t + \theta_{2i} Z_i + \theta_{3i} Z_i t + \theta_{4i} S_t + \theta_{5i} H^D{}_{it} + \theta_{6i} H^N{}_{ijt} + u_{it} \text{....(3)}$$

$$Q_{it-1} = \theta_{1i} Q_{t-1} + \theta_{2i} Z_i + \theta_{3i} Z_i (t-1) + \theta_{4i} S_{t-1} + \theta_{5i} H^D{}_{it-1} + \theta_{6i} H^N{}_{ijt-1} + u_{it-1} \text{....(3.1)}$$

$$y_{it} = \phi_{1i} y_t + \phi_{2i} Z_i + \phi_{3i} Z_i t + \phi_{4i} S_t + \phi_{5i} H^D{}_{it} + \phi_{6i} H^N{}_{ijt} + v_{it} \text{....(4)}$$

$$y_{it-1} = \phi_{1i} y_{t-1} + \phi_{2i} Z_i + \phi_{3i} Z_i (t-1) + \phi_{4i} S_{t-1} + \phi_{5i} H^D{}_{it-1} + \phi_{6i} H^N{}_{ijt-1} + v_{it-1} \text{....(4.1)}$$

$$\Delta x_{it} = x_{it} - x_{it-1}$$

$$\Delta Q_{it} = \theta'_{1i} \Delta Q_t + \theta'_{3i} Z_t + \theta'_{4i} \Delta S_t + \theta'_{5i} \Delta H^D{}_{it} + \theta'_{6i} \Delta H^N{}_{ijt} + \Delta u_{it} \text{....(5)}$$

$$\Delta y_{it} = \phi'_{1i} \Delta y_t + \phi'_{3i} Z_i + \phi'_{4i} \Delta S_t + \phi'_{5i} \Delta H^D{}_{it} + \phi'_{6i} \Delta H^N{}_{ijt} + \Delta v_{it} \text{....(7)}$$

$$\ln Q_{it} = f(\ln Q_t, Z_i, S_t, H_{it}) + u_{it}$$

$$\ln y_{it} = f(\ln y_t, Z_i, S_t, H_{it}) + v_{it}$$

$$\ln Q_{it} = \theta_{1i} \ln Q_t + \theta_{2i} Z_i + \theta_{3i} Z_i t + \theta_{4i} S_t + \theta_{5i} H^D{}_{it} + \theta_{6i} H^N{}_{ijt} + u_{it}$$

$$\ln y_{it} = \phi_{1i} \ln y_t + \phi_{2i} Z_i + \phi_{3i} Z_i t + \phi_{4i} S_t + \phi_{5i} H^D{}_{it} + \phi_{6i} H^N{}_{ijt} + v_{it}$$

$$\Delta \ln Q_{it} = \theta_{1i} \Delta \ln Q_t + \theta_{3i} Z_t + \theta_{4i} \Delta S_t + \theta_{5i} \Delta H^D{}_{it} + \theta_{6i} \Delta H^N{}_{ijt} + \Delta u_{it} \ldots (6)$$

$$\Delta \ln y_{it} = \phi_{1i} \Delta \ln y_t + \phi_{3i} Z_i + \phi_{4i} \Delta S_t + \phi_{5i} \Delta H^D{}_{it} + \phi_{6i} \Delta H^N{}_{ijt} + \Delta v_{it} \ldots (8)$$

We take logs because when there is measurement error this shrinks them!!!

Use basic demand and supply graphs for the county that is hit by hurricanes and for the county that is not hit by the hurricanes

Predict the theoretical implications on changes in wages and changes in employment (in the county that is hit by hurricane and in the county that is spared by hurricane).

Please draw the graphs – follow discussion in class

Assigning a sign to regression parameters using theory

	Percent Change in Employment	Percent Change in Earnings
1. Effect of category 1–3 hurricane on county directly hit	−1.5	+1.3
2. Effect of category 4–5 hurricane on county directly hit	−4.5	+4.4
3. Effect of category 1–3 hurricane on neighboring county	+0.2	−4.5
4. Effect of category 4–5 hurricane in neighboring county	+0.8	−3.3

Important Econometric insights from Lemar

1) Continuous variables can be turned discrete
2) Discrete vs continuous variables measure the margin of analysis

Lemar Questions
1) What are two sources of variance and estimates?
Selection and misspecification, can only get rid of the latter
2) Do Data tell us anything about the size of bias?
No, only theory can
3) How are anecdotes and experiments related?
Anecdotes lead to experiments, gives you a good idea
4) What is more important than the possibility of bias?
The direction of bias
5) What is the best way of overcoming the whimsy of the statistical inference process?
Sensitivity analysis (priors, functional form , inclusion of some variables)
Rpbustness

Problems of Randomized Experiments - Cost

- **Financial costs** : The **negative income tax (NIT) experiments** of the late 60s and 70s in the US.
- NIT: progressive income tax system where people earning below a certain amount receive supplemental pay from the government instead of paying taxes to the government
- **Ethical problems**: Treated and Control both face issues
- Memory Books, Medicines etc.

Problems of Randomized Experiments – Threats to Internal Validity

- **Non response bias**: attrition bias – NIT experiment – people moving away
- **Mix up of Treatment and Controls**: Separate allocation hard. Example: (Krueger 2000) evaluation of the Tennessee Star small class size experiment
- It is then important to use the **initial assignment** as the treatment, because it is the only variation that was randomly assigned. **It can then be used as an instrument for actual class size**.

Problems of Randomized Experiments – Threats to External Validity

- **Limited duration**: Behaviors different when something is not permanent
- **Experiment Specificity**: The NIT experiments – things that work in one part of the country may not work in other – hard to generalize
- **Hawthrone and John Henry effects**: Treatment and control may behave differently because they know they are being observed.
- **General Equilibrium effects**: small scale experiments do not generate general equilibrium

Problems of Randomized Experiments – Threats to Power of the Experiment

- **Small samples (individuals)**: Because experiments are difficult to administer, samples are often small, which makes it difficult to obtain significant results.
- **Experiment design and power of the experiment**: When the unit of randomization is a group (e.g. a school), we may need to collect data on a very large number of individuals to get significant results, if outcomes are strongly correlated within groups (see below how standard errors are corrected for the grouped structure).