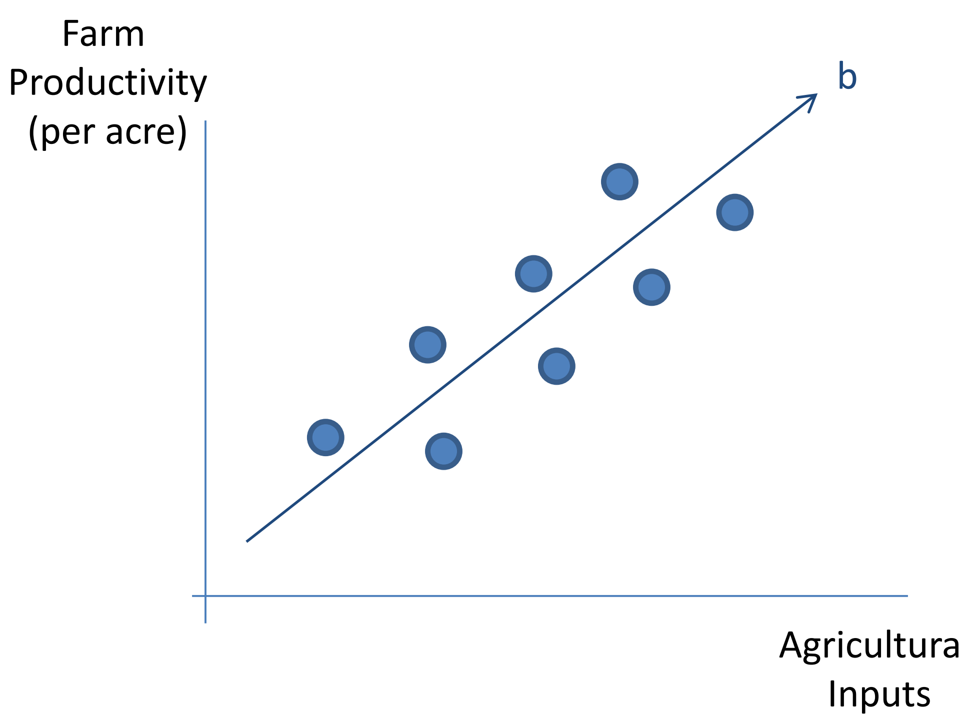
CORRECTING BIAS WITH PANEL DATA

Fundamentals of

PROGRAM EVALUATION

JESSE LECY

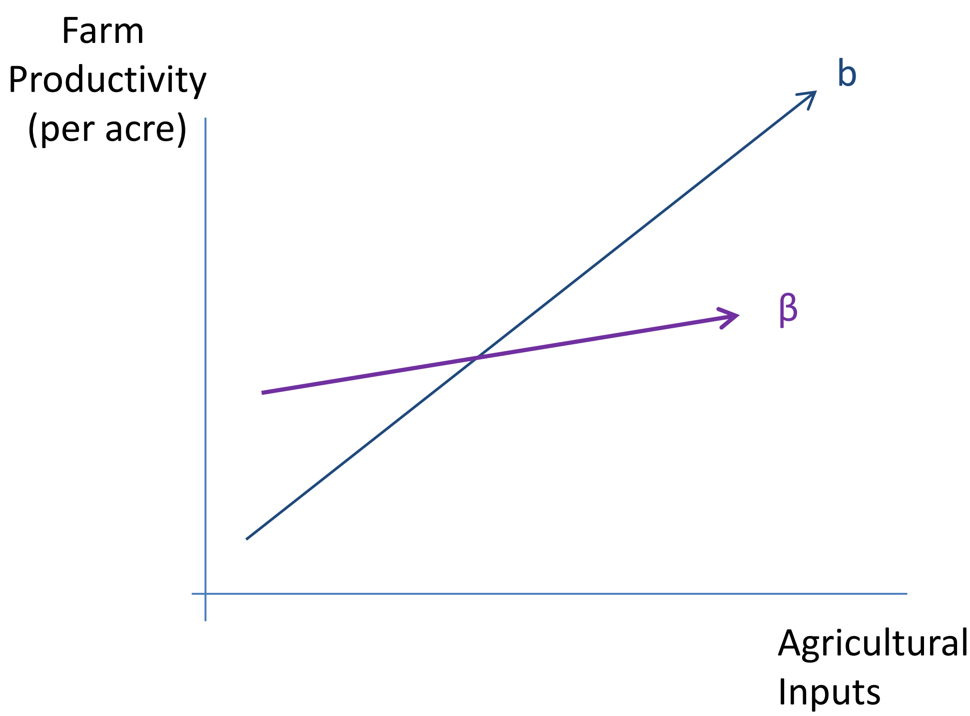
MANAGEMENT BIAS (MUNDLAK, 1965)



$$crop\_yield = b_0 + b_1 \cdot fertilizer + e$$

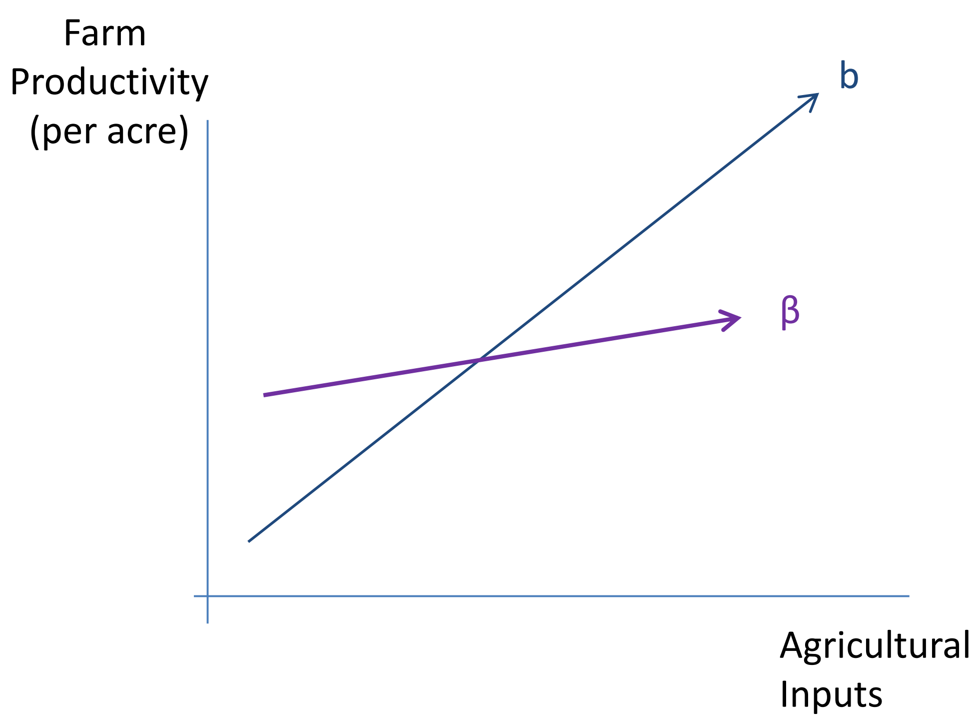
(each data represents a specific farm)

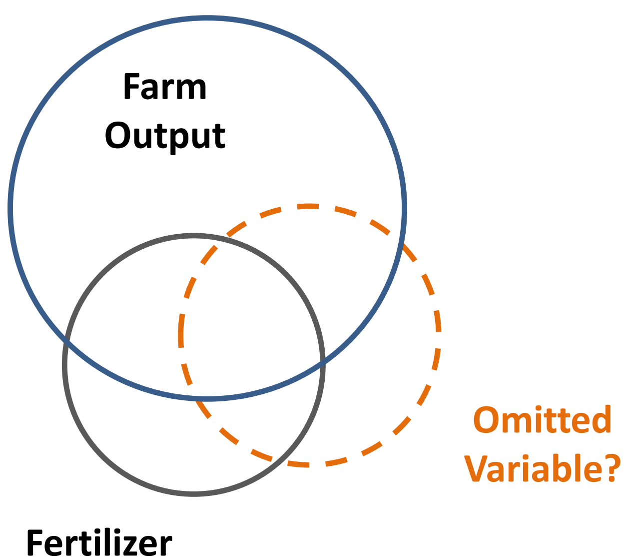
MANAGEMENT BIAS (MUNDLAK, 1965)



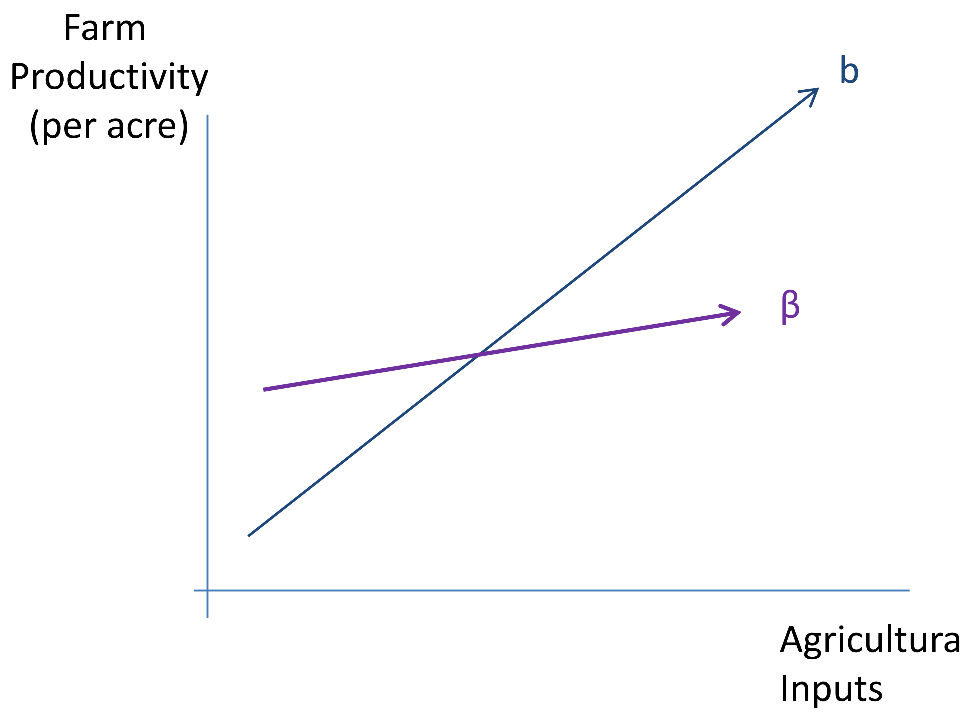
Policy impact did not meet expectations based upon the data.

MANAGEMENT BIAS (MUNDLAK, 1965)

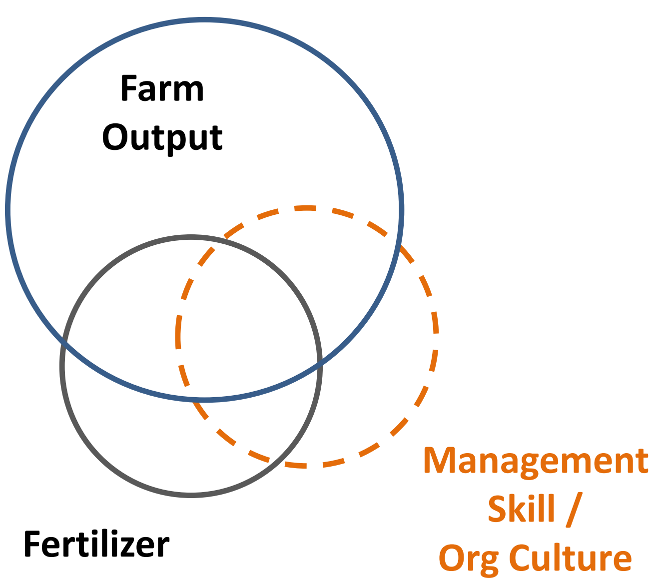


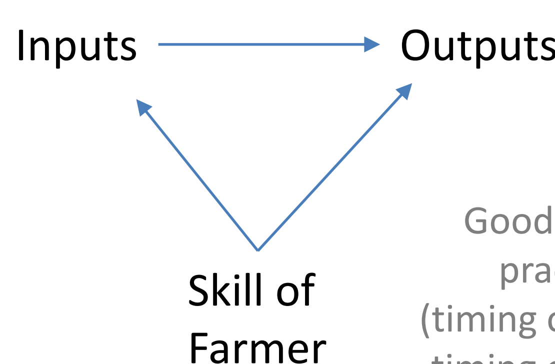


MANAGEMENT BIAS (MUNDLAK, 1965)



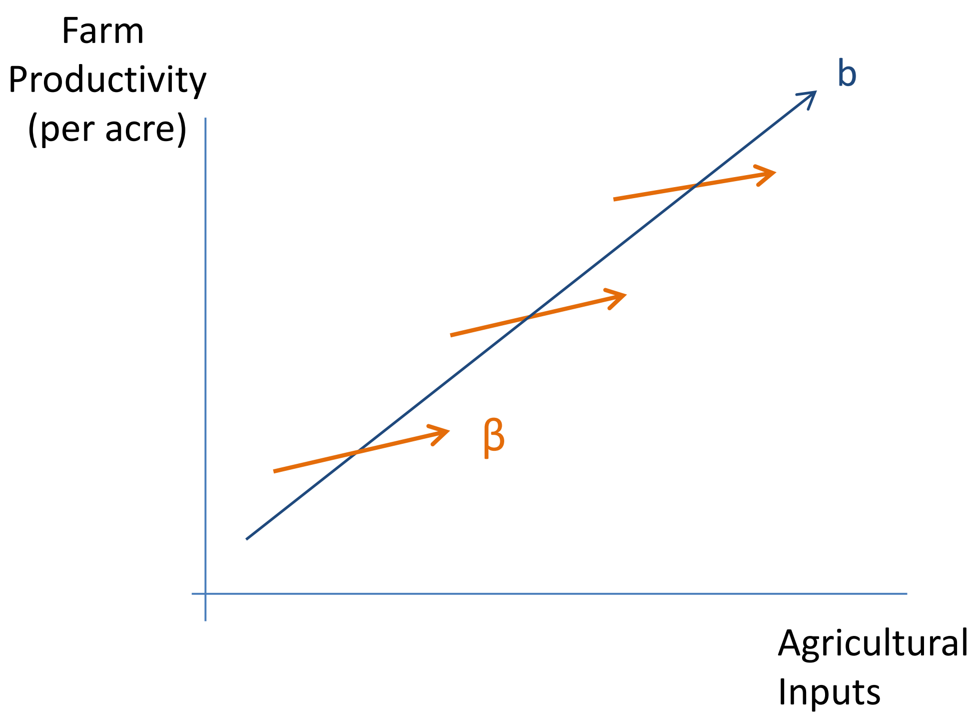
Contribution of inputs over-stated because high levels of input are correlated with other best practices





Good farming practices (timing of planting, timing of harvest, care of soil, quality of inputs)

MANAGEMENT BIAS (MUNDLAK, 1965)



How can we hold skill constant?

What if we instead focus on how the output changes on **each farm** when they change input?

PANEL DATA

PANEL DATA

Cross Section

X	Y	Gender
1	8	Male
3	6	Female
2	5	Female
2	3	Male

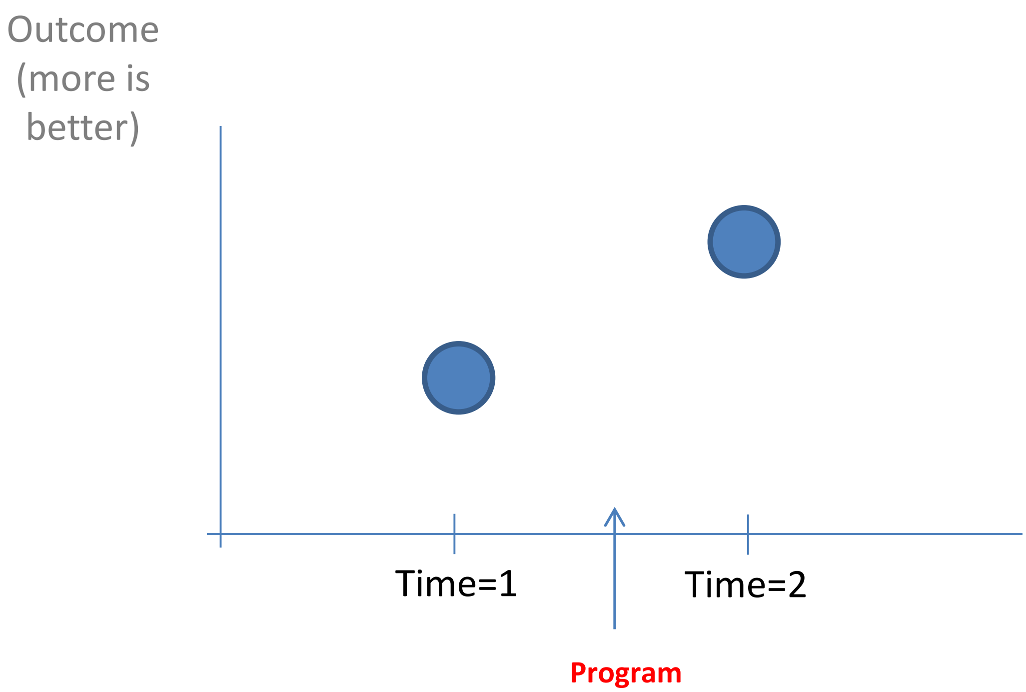
Time Series

Time	Y
1	8
2	6
3	5
4	3

PANEL DATA

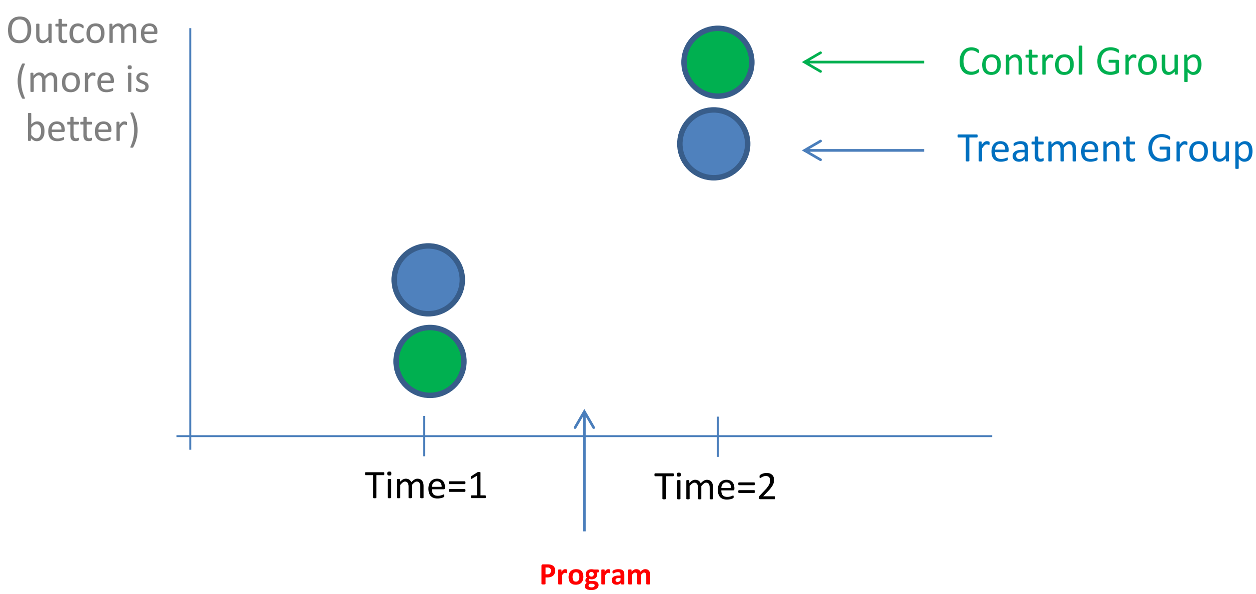
ID	Time	X	Y
A	1999	1	8
A	2000	3	6
B	1999	2	5
B	2000	2	3

IS THIS PROGRAM EFFECTIVE?



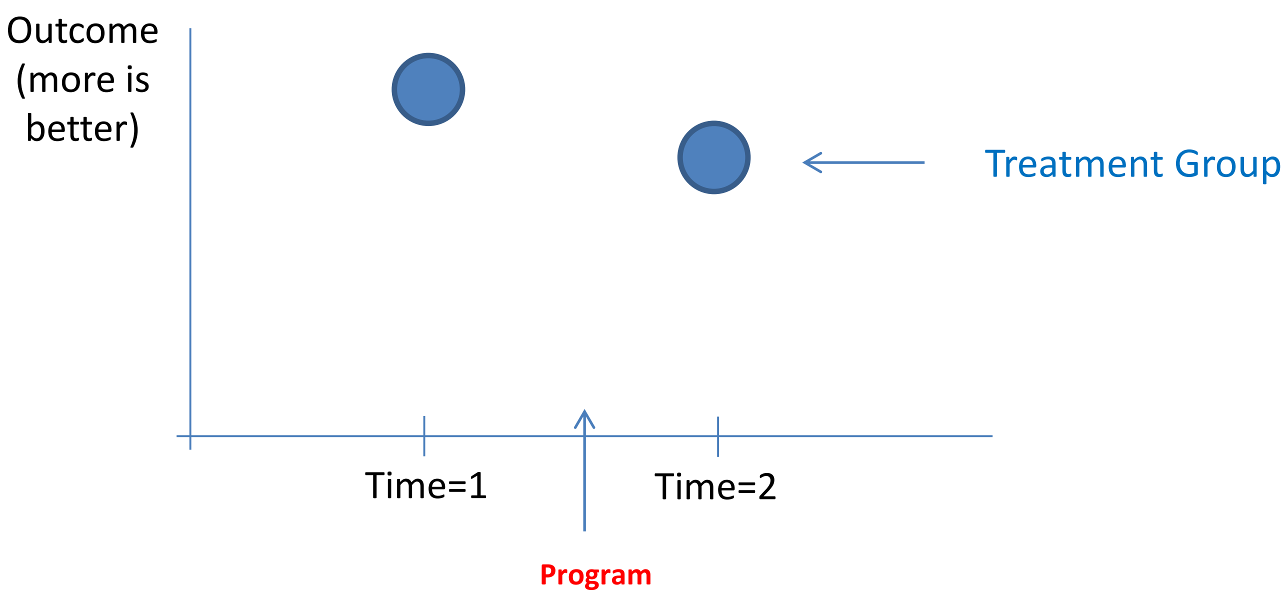
$O_1 \times O_2$

WHAT ABOUT NOW?



Treatment: O_1 x O_2
Control: O_1 O_2

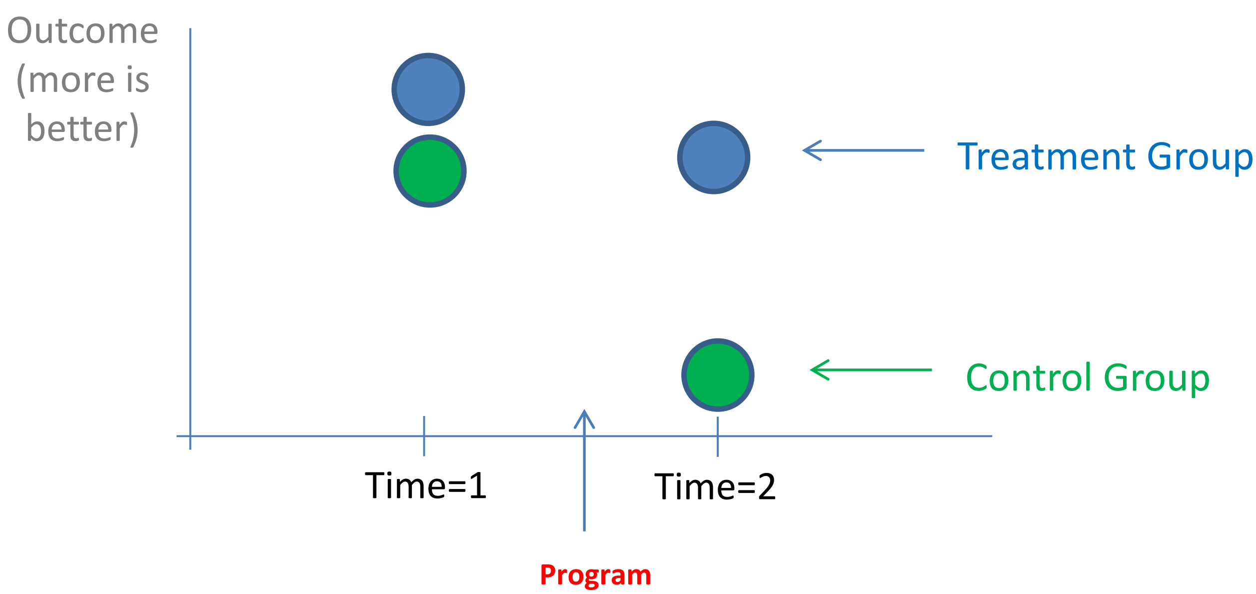
WHAT ABOUT NOW?



Treatment: O_1 x O_2

Control: O_1 O_2

WHAT ABOUT NOW?



Treatment: $O_1 \times O_2$

Control: $O_1 \quad O_2$

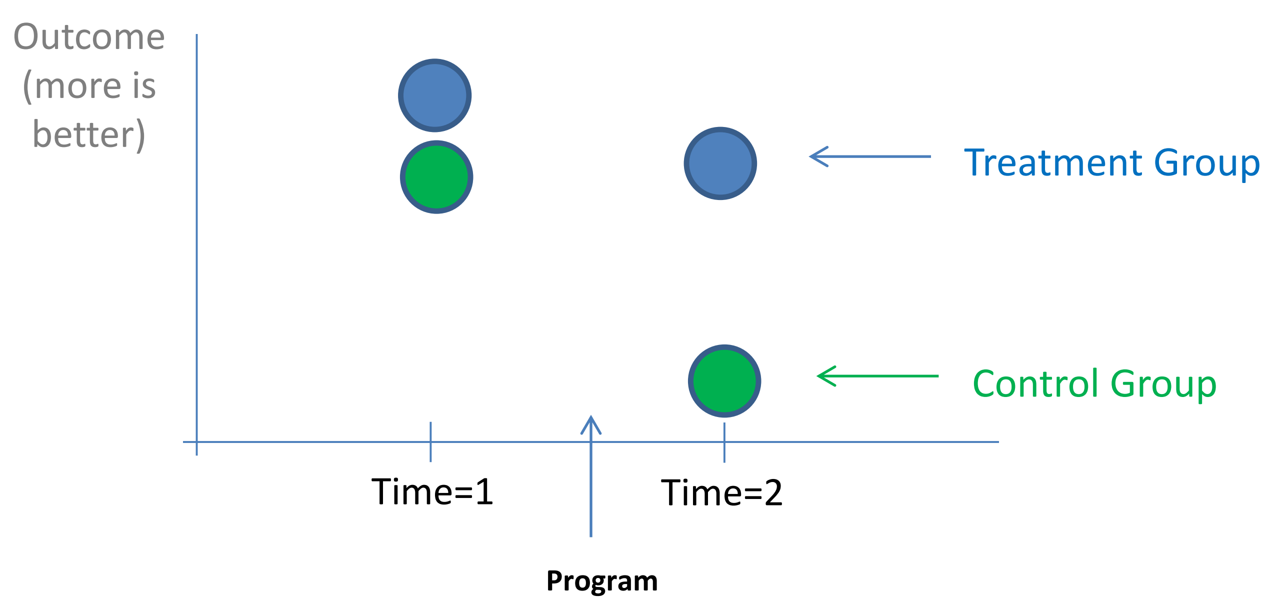
HOW ARE WE REASONING ABOUT IMPACT?

Is the treated group better off than the control?

T2 – T1 = C2 – C1 ?

Change over time

Change over time



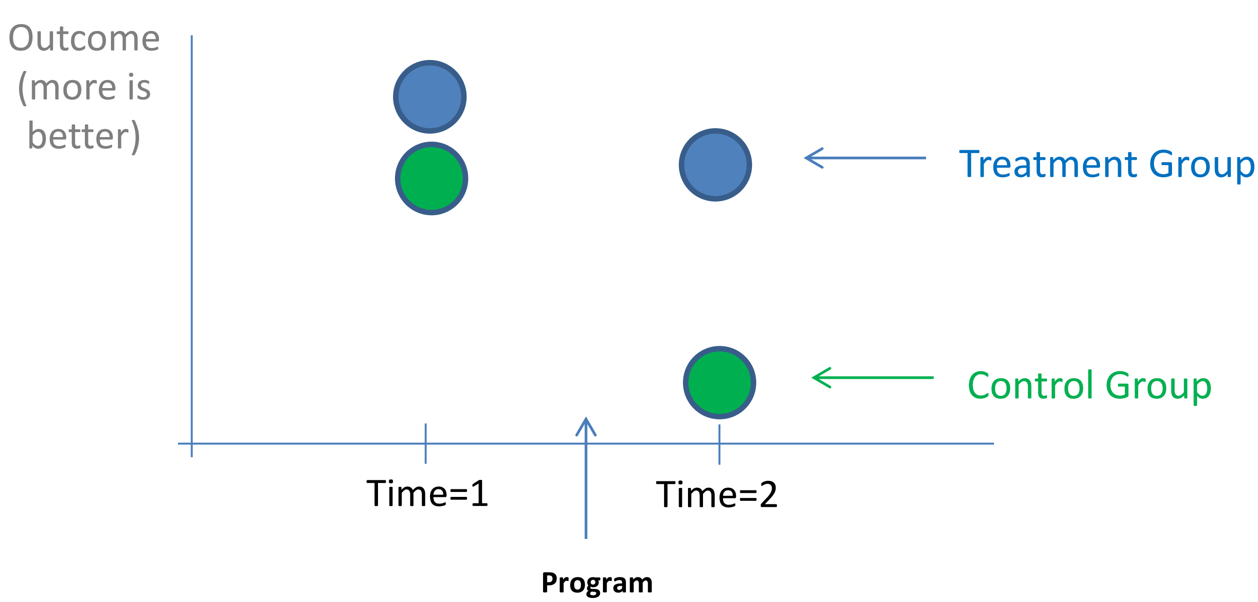
HOW ARE WE REASONING ABOUT IMPACT?

Diff-in-diff models are appropriate when:

- Treatment is discrete (yes or no you received it)
- The outcome is continuous
- We have data from before and after the intervention

Limitations:

- Many treatments are better represented as levels (how many hours at the gym is better than did you go to the gym)
- We often have many groups, not just treatment and control



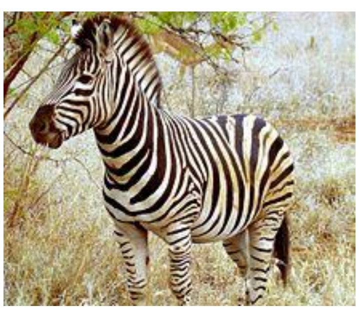
The context –

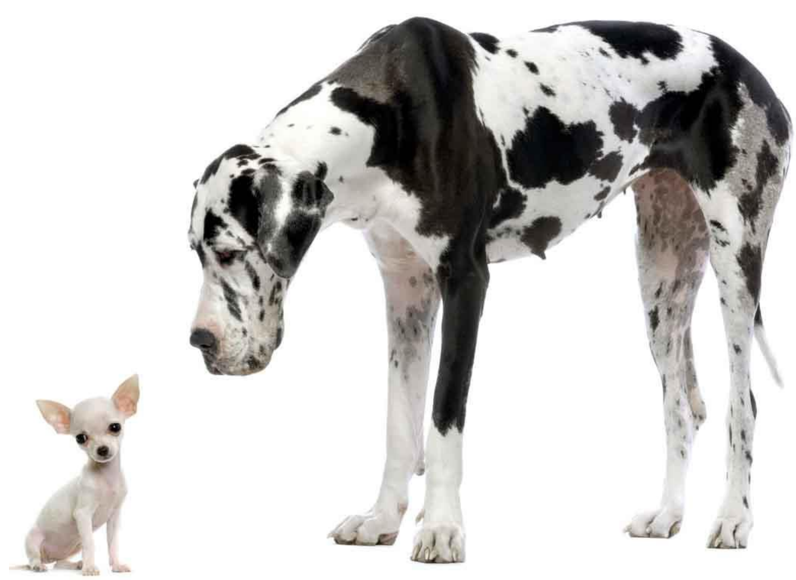
the initial level of outcomes for treatment and control groups

– is important !

GROUP STRUCTURE MATTERS

When individuals are isolated by geography or institution they develop differently. This geographic / group structure matters a great deal in social sciences.





HETEROGENEITY BIAS

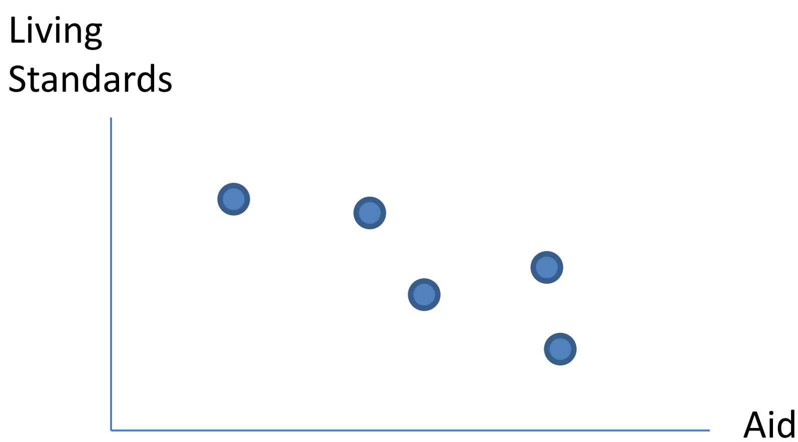
Bias that results when you try to determine the impact of a program or policy and you don't take into account group structure, i.e. different initial starting conditions of the outcome.

Would we treat all ungulates the same in a study?

GROUP STRUCTURE IN REGRESSION TERMS?

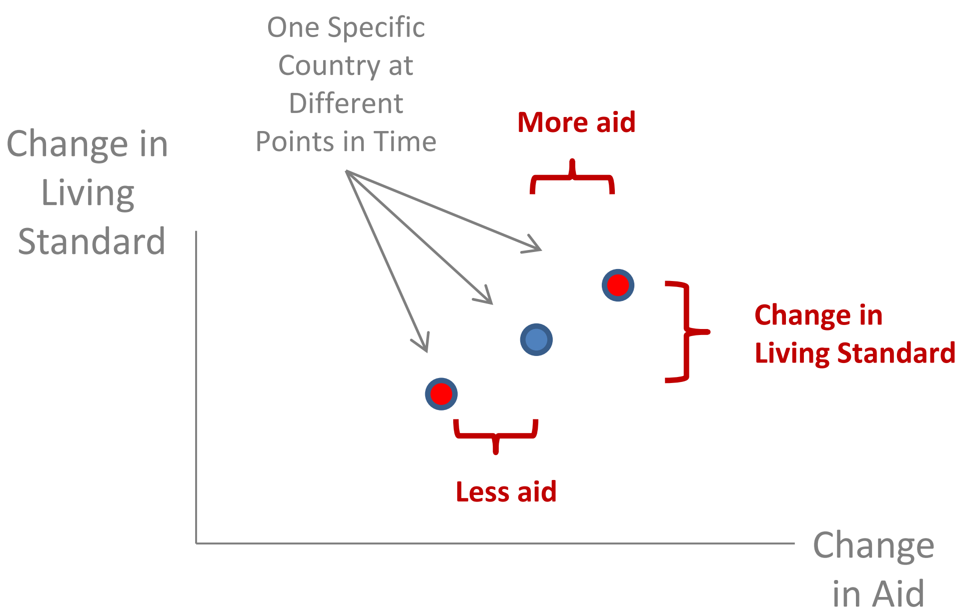
FIXED EFFECTS MODELS

LEVELS

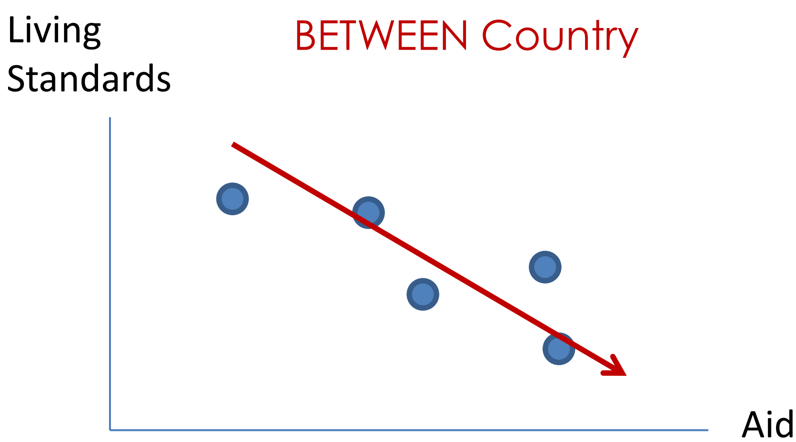


This is what the "levels" data will look like. Note that poor countries NEED the most aid, so the direction of causality is opposite of what this regression suggests. We cannot interpret the slope as a program impact here.

CHANGE IN LEVELS

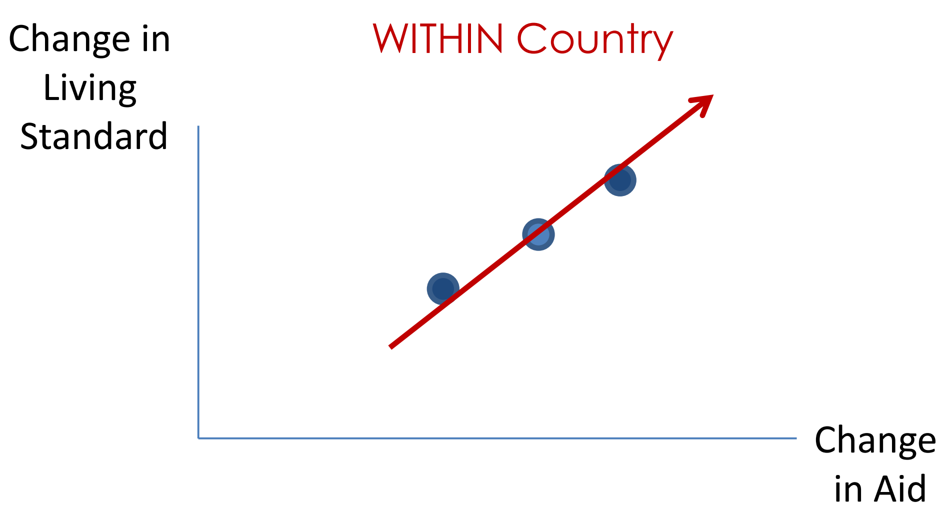


If we look at changes in the levels of aid for one specific country, we can look at how that change either increases or decreases the standard of living in the country.



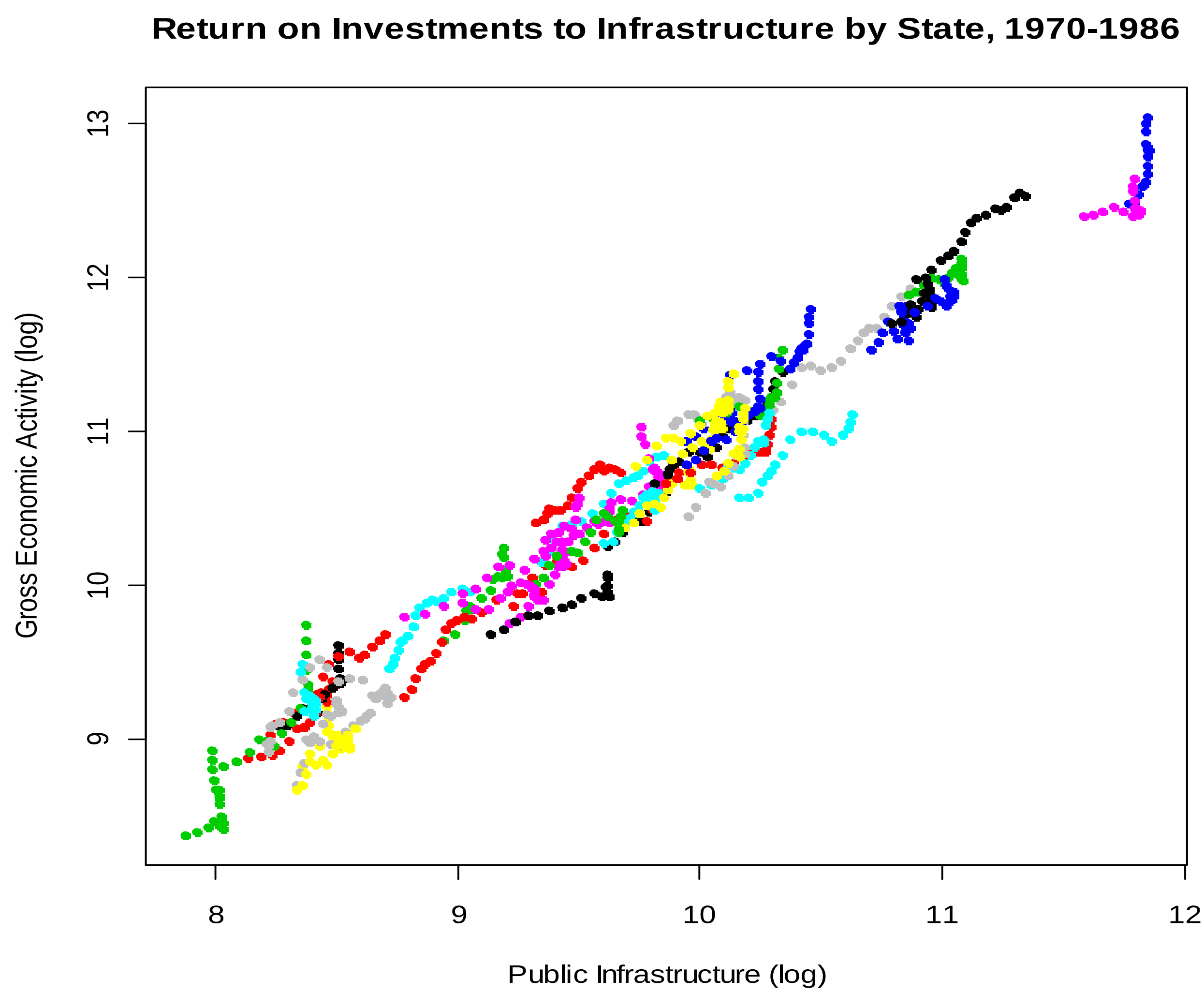
One Point in Time
(cross-section)

What the relationship looks like when you estimate the policy impact using all of the data from many countries.



Change Over Time
(panel)

The relationship focusing on one country over time.



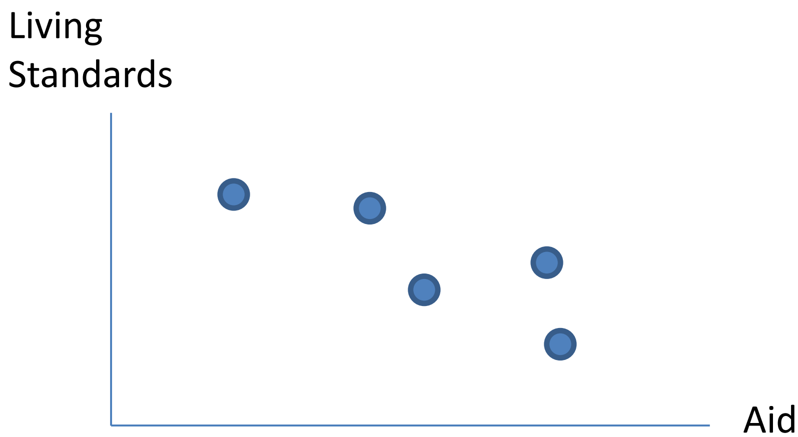
Each color represents data from a state over the 17-year study period.

Levels of spending and levels of economic development represent the specific characteristics of a group (a state in this case).

What is our inference about the policy impact of increased infrastructure spending when we use all of the data together, and does it change when we take into account group structure?

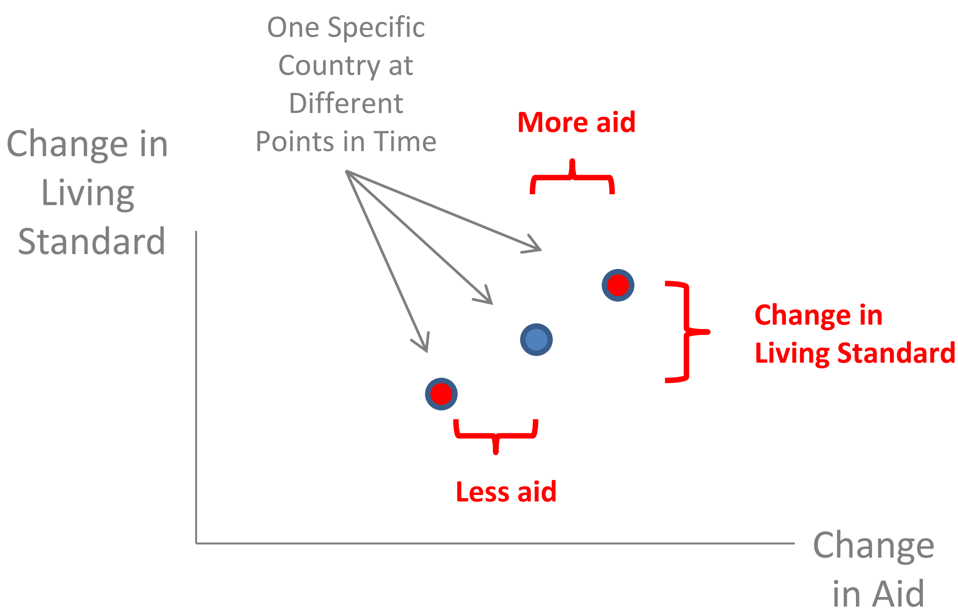
(data from Baltagi & Pinnoi 1995)

LEVELS

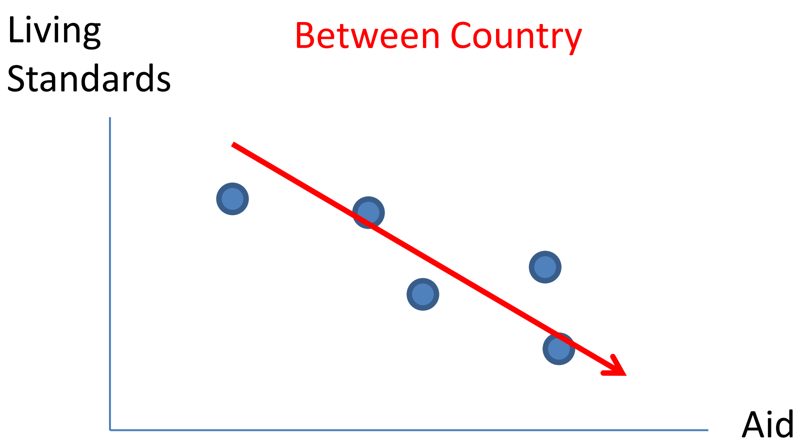


This is what the “levels” data will look like. Note that poor countries NEED the most aid, so the direction of causality is opposite of what this regression suggests. We cannot interpret the slope as a program impact here.

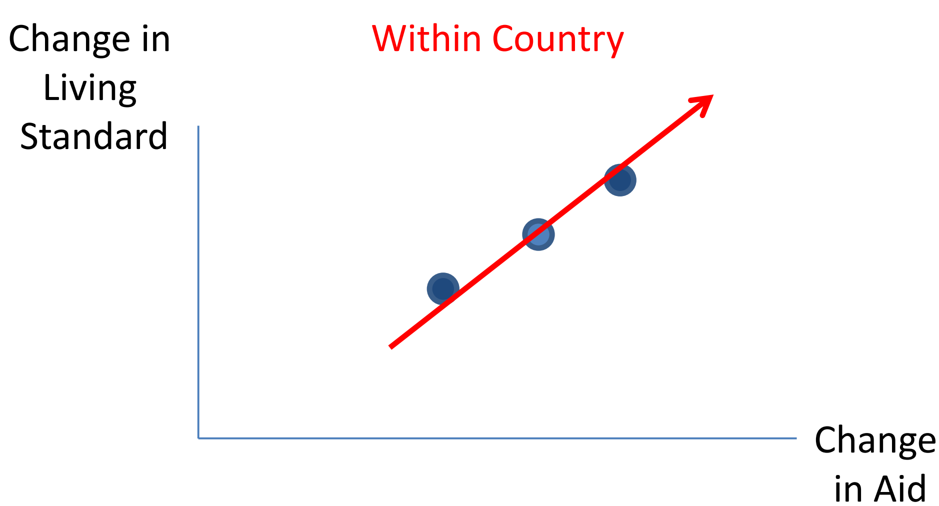
CHANGE IN LEVELS



If we look at changes in the levels of aid for one specific country, we can look at how that change either increases or decreases the standard of living in the country.



What the relationship looks like when you estimate the policy impact using all of the data from many countries.



The relationship focusing on one country over time.

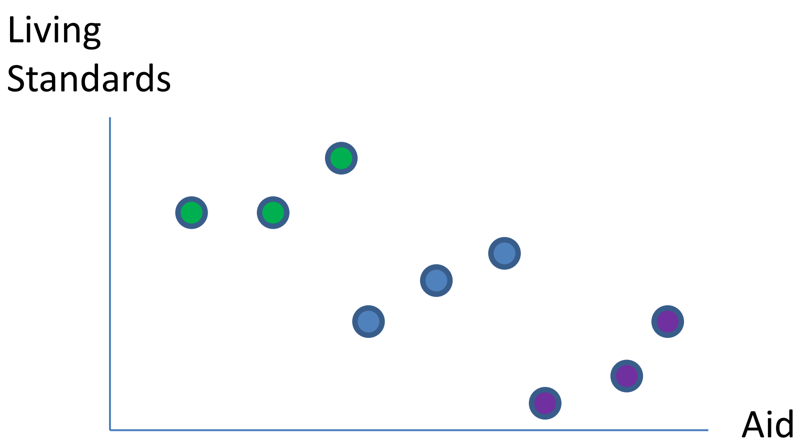
DIFFERENCES BETWEEN "LEVELS" AND CHANGES IN LEVELS

- If you are looking at the relationship between foreign aid and the standard of living within a country, there is a large potential for bias because the countries that receive the most aid are going to be the poorest countries.

- As a result, if you regress the standard of living onto aid – simply looking at the levels of living standards and the levels of foreign aid – you will incorrectly conclude that foreign aid actually hurts nations.

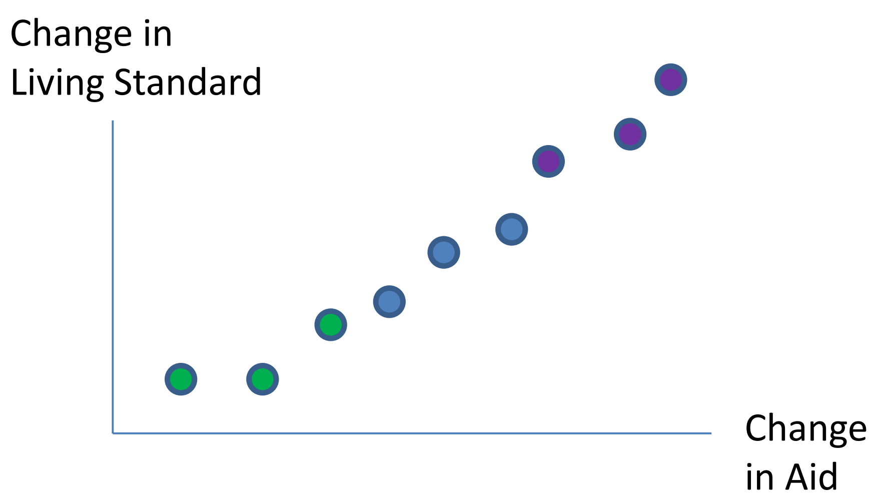
- If we think about it more like an experiment, we would want to randomly assign levels of aid to different countries to see if it helps. That is not politically feasible, so instead we might think of a different experiment. What if we took all of the countries that are receiving aid, and we randomly assign them a one-unit increase or a one-unit decrease in aid. So they still receive close to the same levels of aid as the previous year, but they get a small boost or a small shock.

- We are moving from an examination of the "levels" of foreign aid to changes in foreign aid.

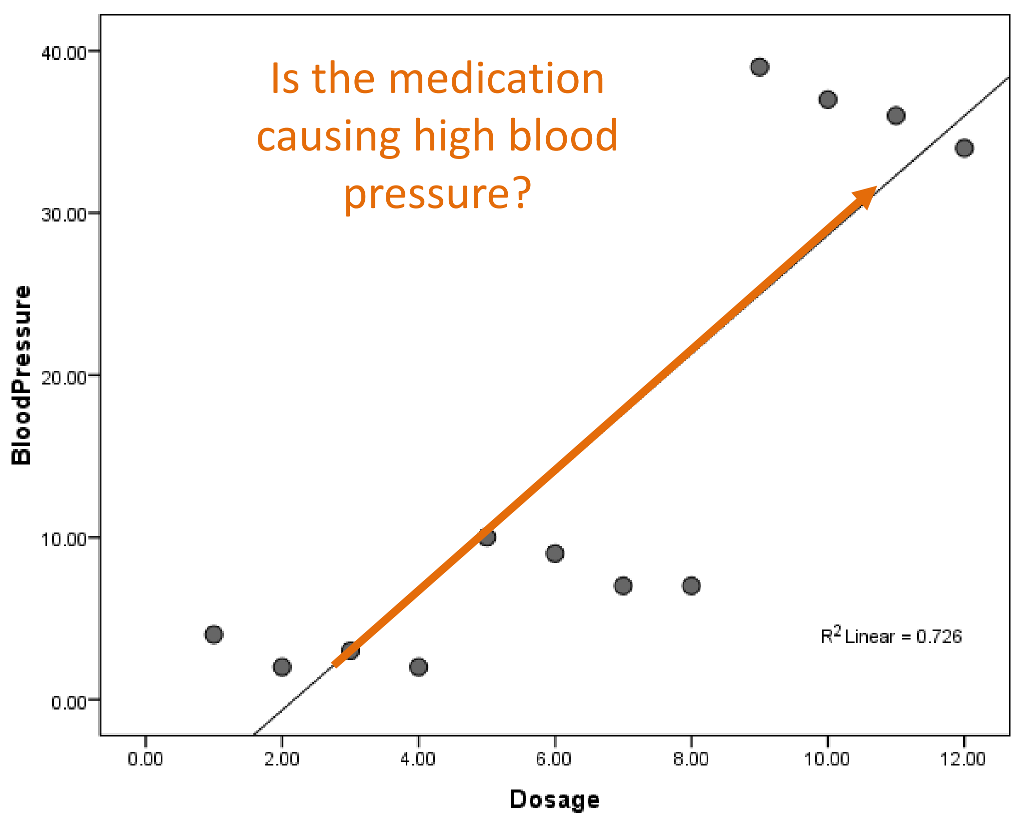


The fixed effect model does this by examine changes that occur within-group over time.

By adding an intercept for each group, you are putting them on the same axis so now you can estimate a common slope while taking into account the effects of the unique history/culture of the groups.

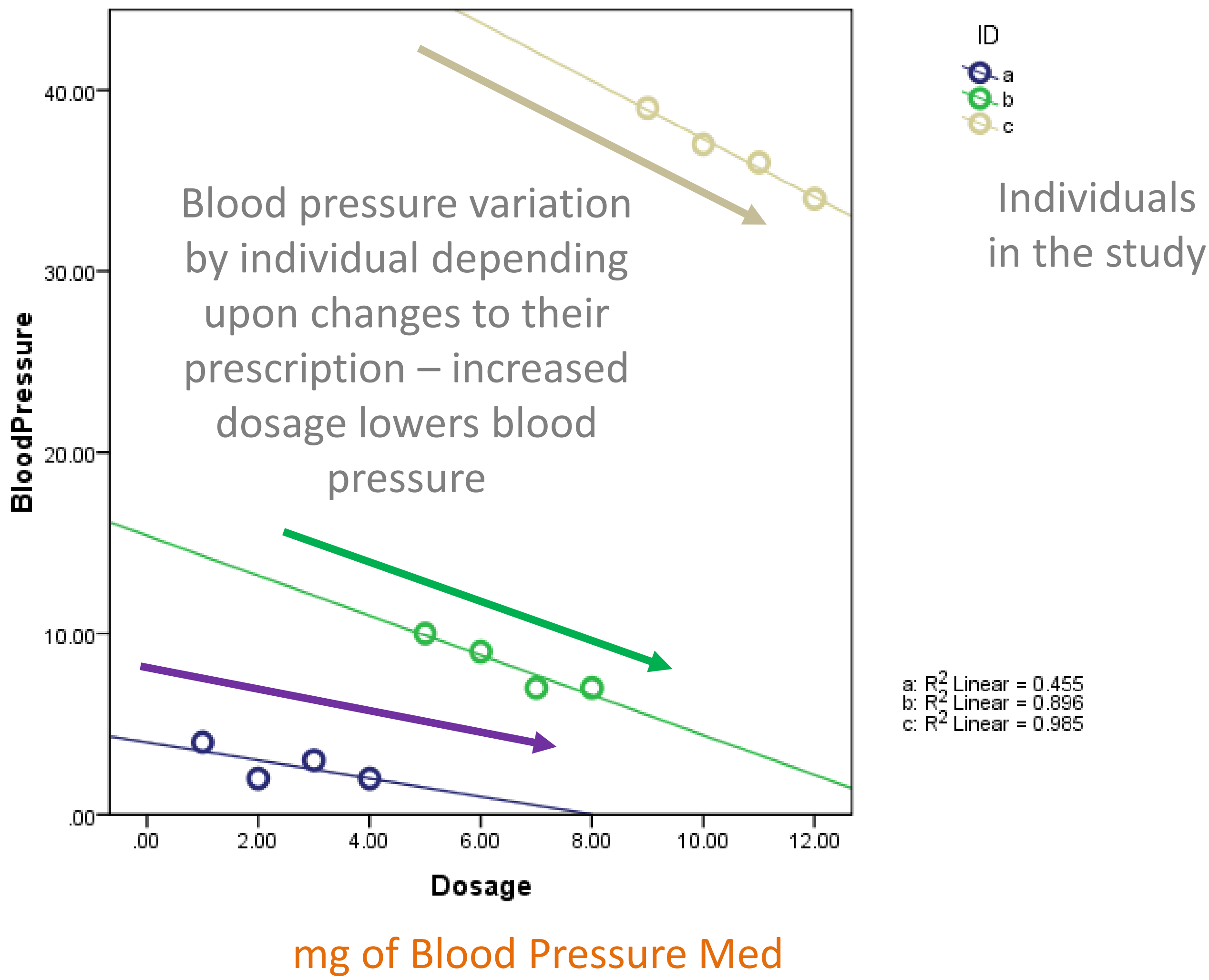


CROSS-SECTION VARIATION (LEVELS)

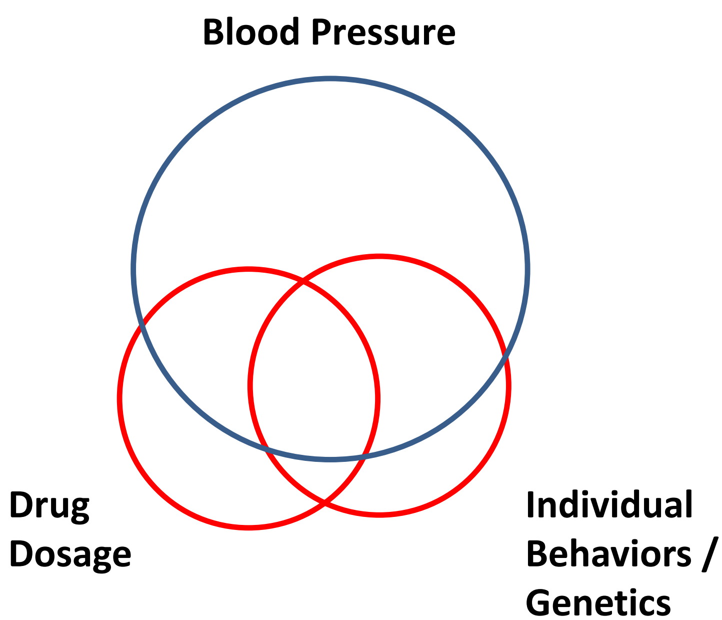


mg of Blood Pressure Med

WITHIN GROUP VARIATION



WHY DO THE SLOPES CHANGE?

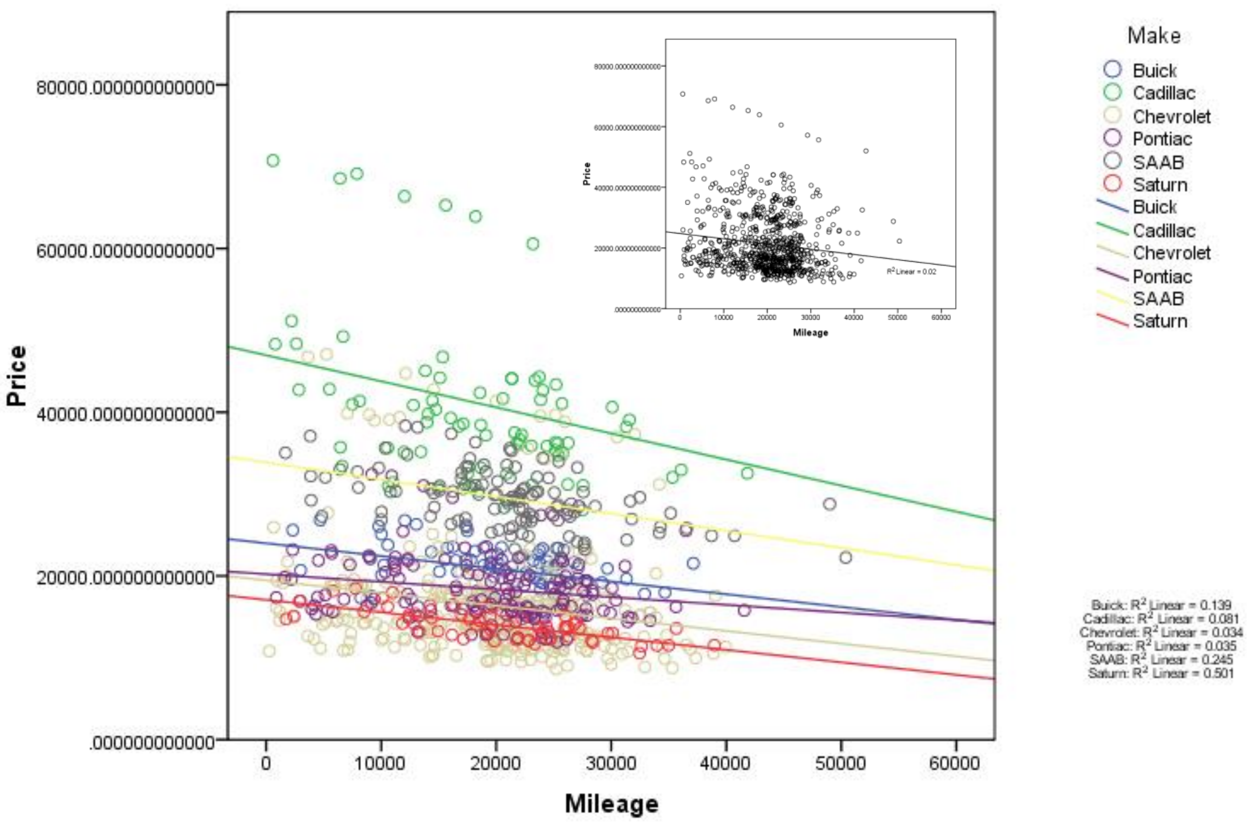


CORRELATION

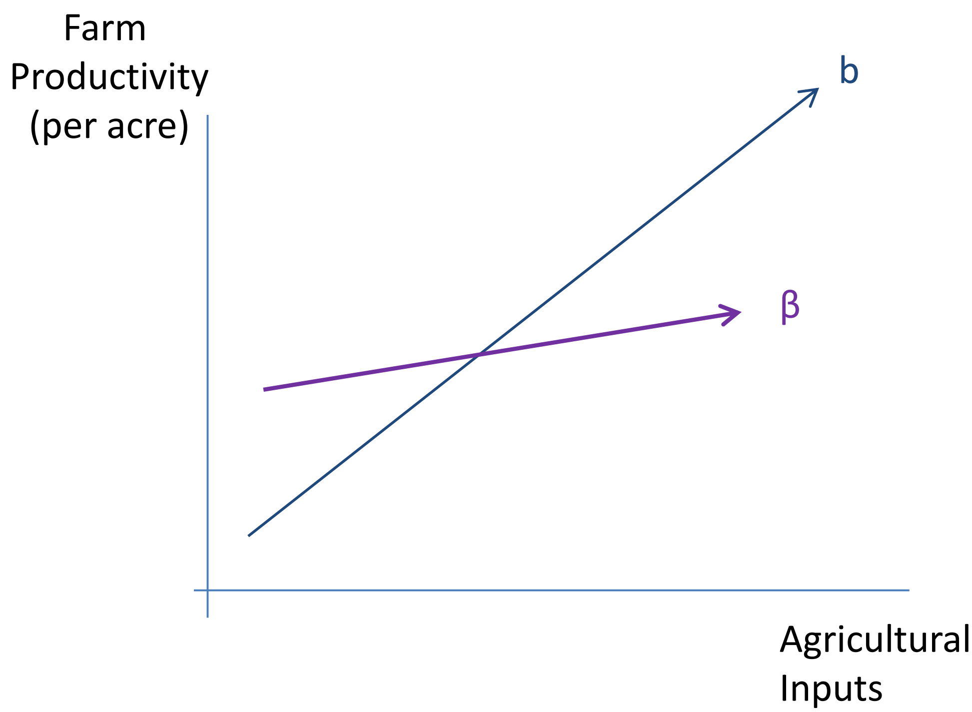
Dosage is set based upon the baseline levels

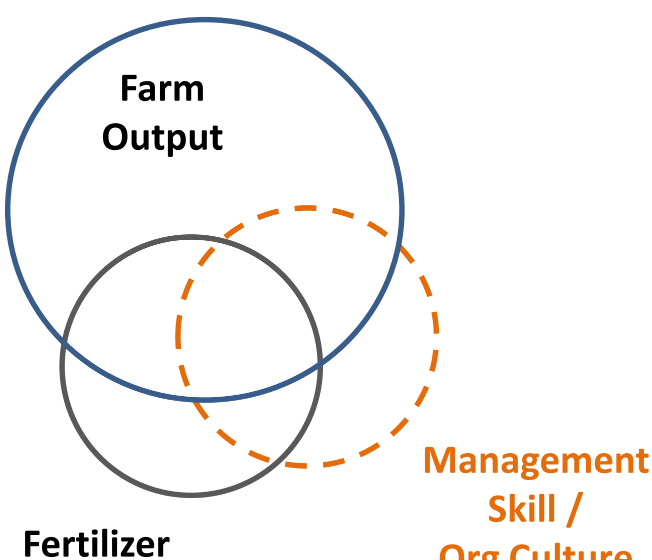
Fixed traits that shape the baseline blood pressure

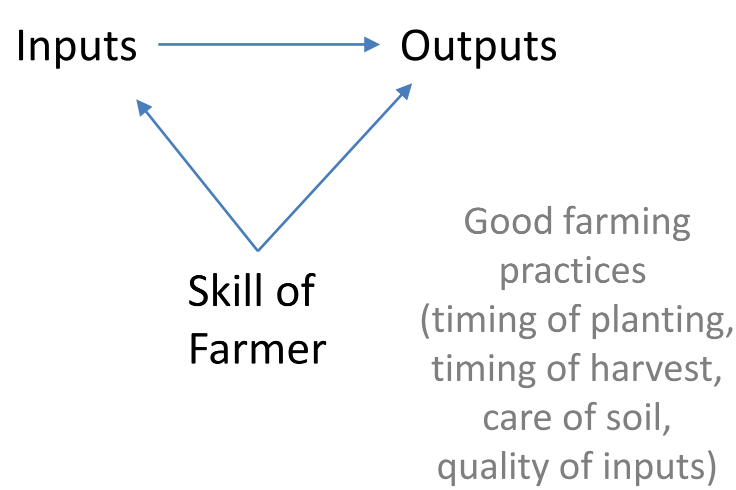
ANOTHER EXAMPLE



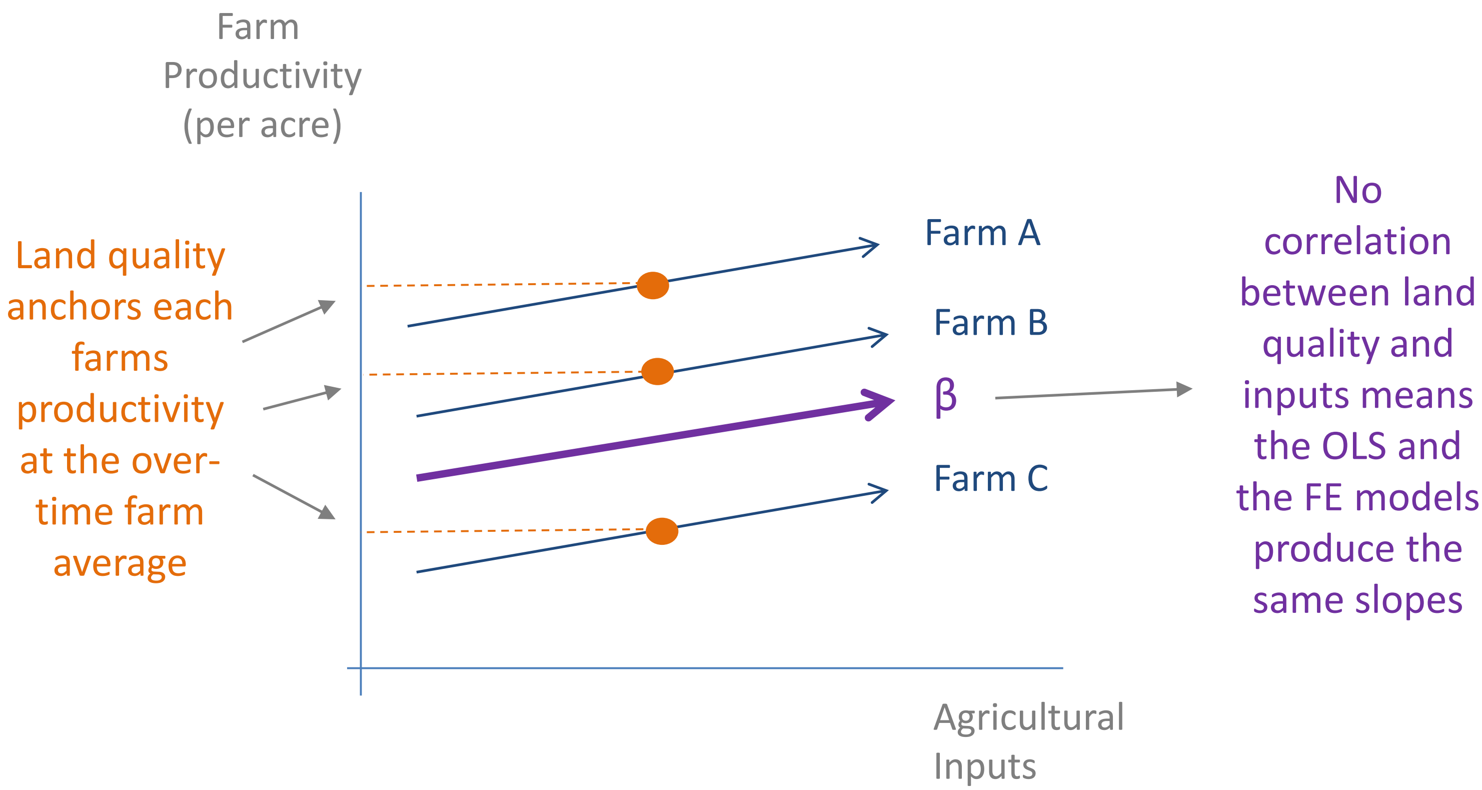
WHY ARE THE SLOPES THE SAME?

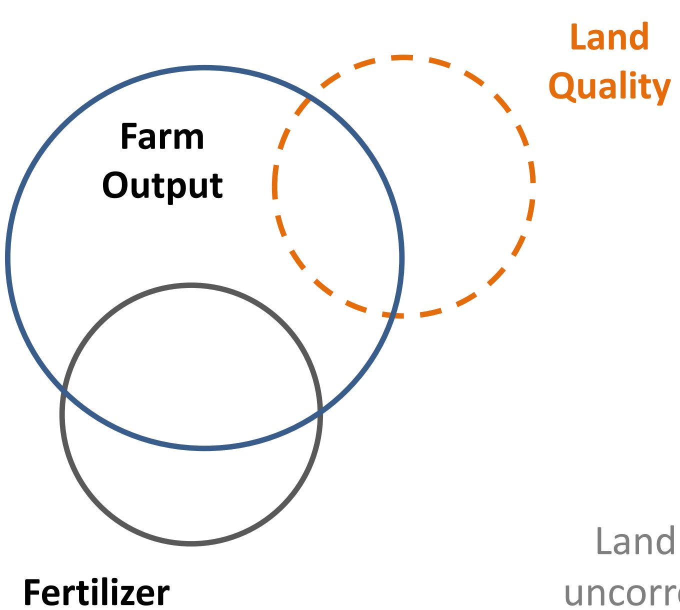


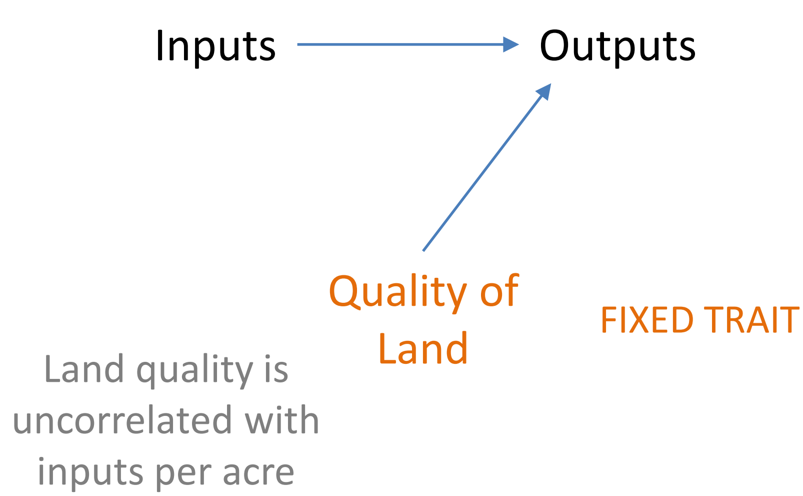




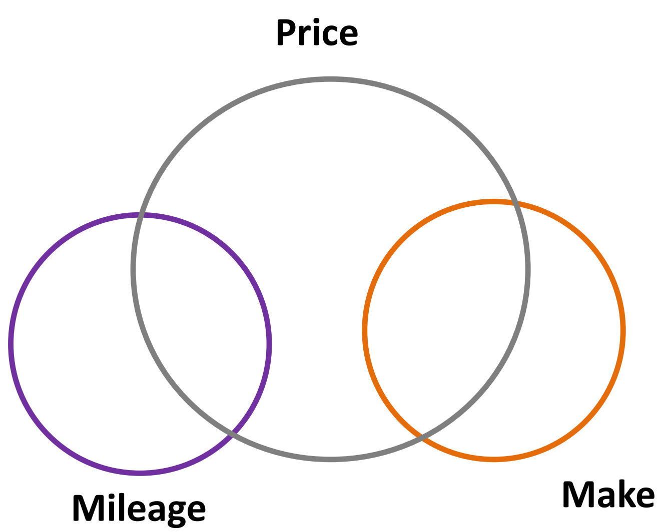
WHY ARE THE SLOPES THE SAME?

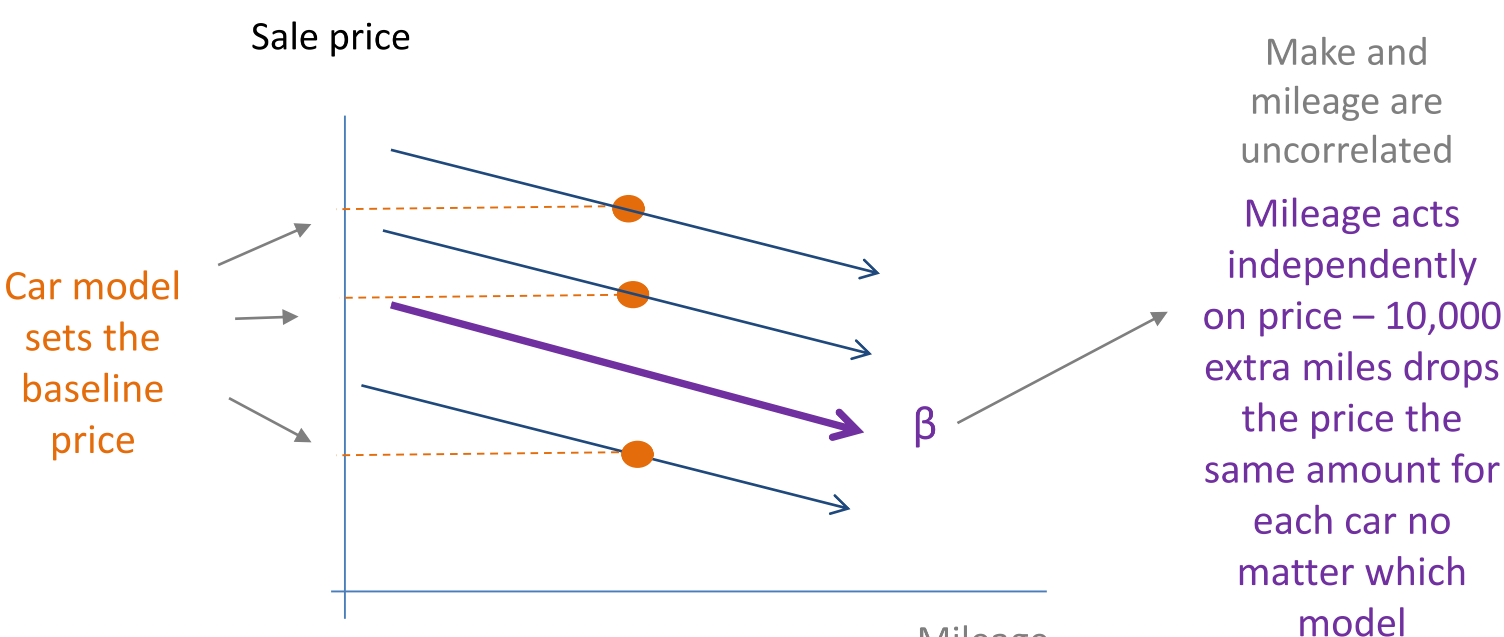




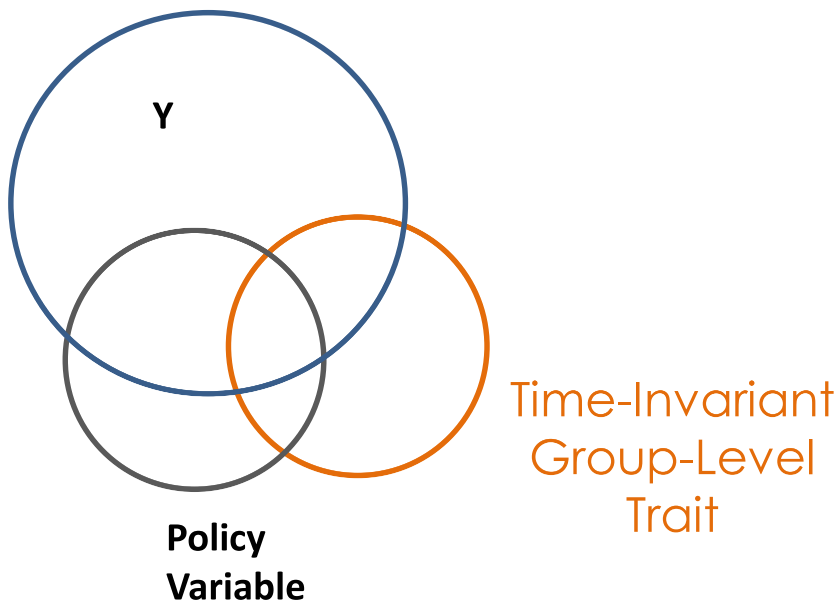


WHY ARE THE SLOPES THE SAME?



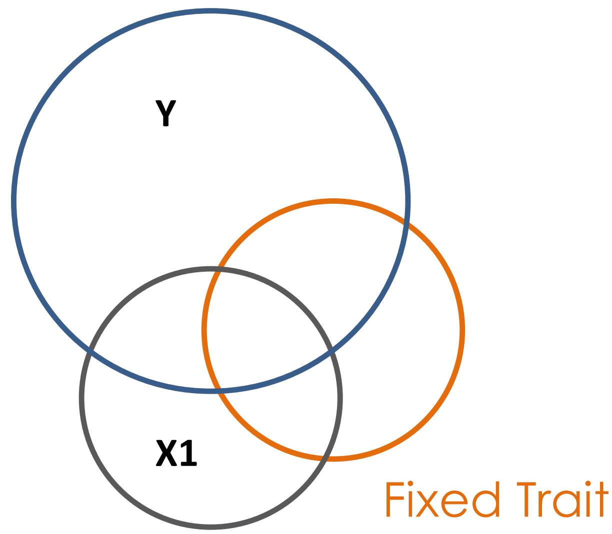


WHEN TO USE FIXED EFFECTS?



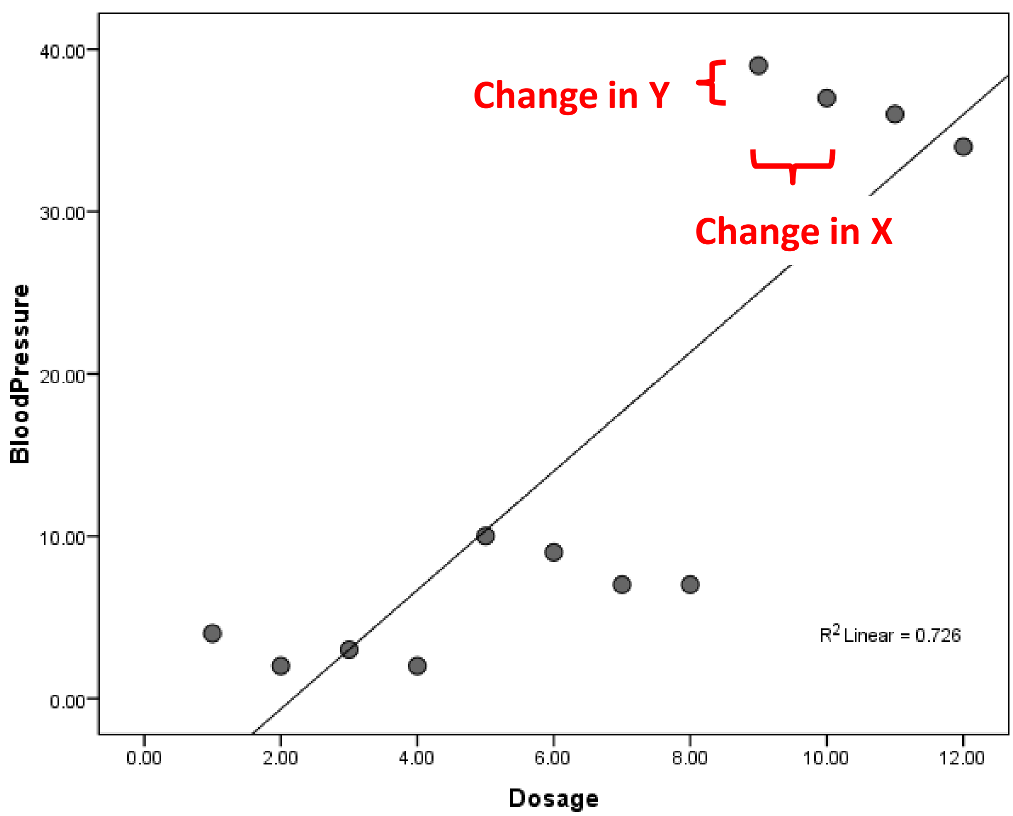
THE "FIXED EFFECT" IS A CATCH-ALL VARIABLE FOR ANYTHING ABOUT THE GROUP THAT IS "FIXED" ACROSS TIME

- "Culture"
- Race
- Gender
- IQ
- Management or Productivity
- Ability



Even if we cannot measure it,
we can get rid of it!

ISOLATING CHANGE OVER TIME



HOW IT WORKS MATHEMATICALLY:

$$
\begin{aligned}
& Income_{t=2} = a_0 + a_1 Edu_{t=2} + a_2 Gender_{t=2} + \varepsilon_{t=2} \\
-&\left(Income_{t=1} = a_0 + a_1 Edu_{t=1} + a_2 Gender_{t=1} + \varepsilon_{t=1}\right)
\end{aligned}
$$

$$
\Delta Income = \quad \Delta b_0 + b_1 \Delta Edu + b_2 \Delta Gender + \Delta\varepsilon
$$

If any covariate is **time-invariant**:

$$
\Delta Income = \quad \cancel{\Delta b_o} + b_1 \Delta Edu + b_2 \cancel{\Delta Gender} + \Delta\varepsilon
$$

"FIXED EFFECTS" REGRESSION MODEL

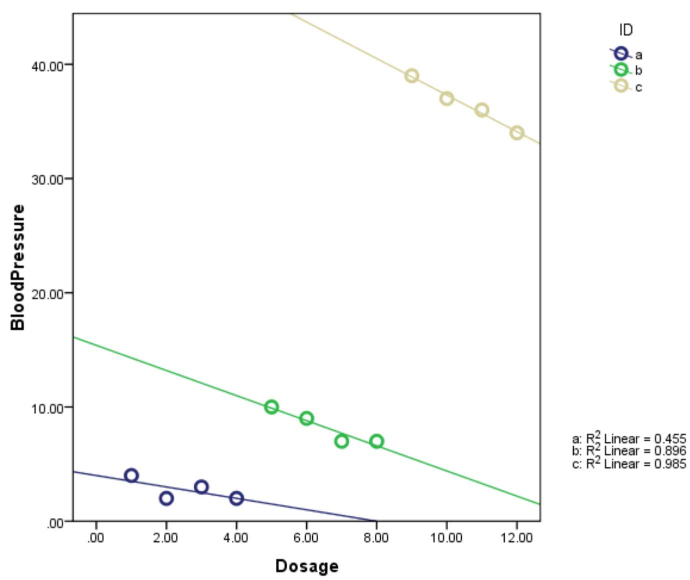
$$BP = b_1 Dosage + b_2 S1 + b_3 S2 + b_4 S3 + e$$

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	Dosage	-1.067	.213	-.362	-5.013	.001
	Subject1	5.417	.673	.144	8.050	.000
	Subject2	15.183	1.443	.404	10.521	.000
	Subject3	47.700	2.272	1.269	20.995	.000

a. Dependent Variable: BloodPressure

b. Linear Regression through the Origin



"FIXED EFFECTS" REGRESSION MODEL

Note – there is no intercept term. Why is this?

$$BP = b_1 Dosage + b_2 S1 + b_3 S2 + b_4 S3 + e$$

C	S1	S2	S3
1	1	0	0
1	1	0	0
1	0	1	0
1	0	1	0
1	0	0	1
1	0	0	1

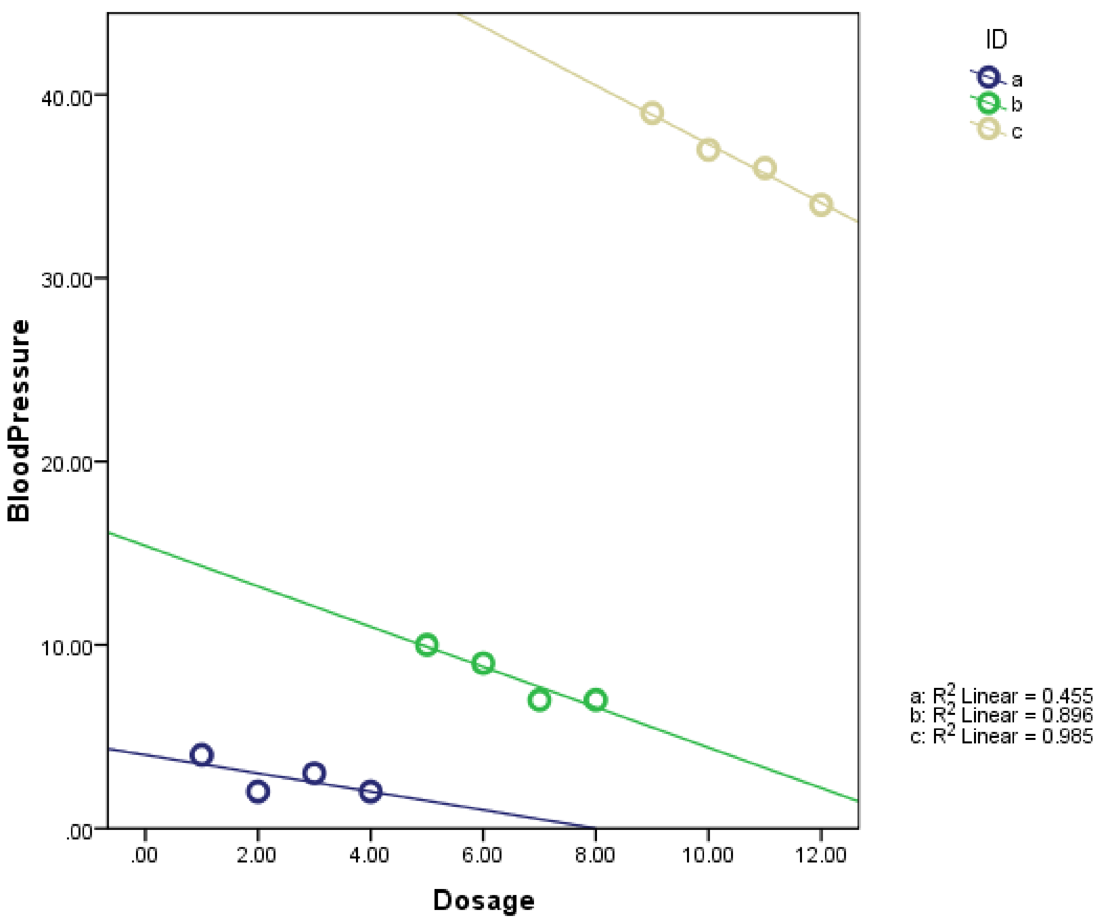
$$C = S1 + S2 + S3$$

Perfect Multicollinearity

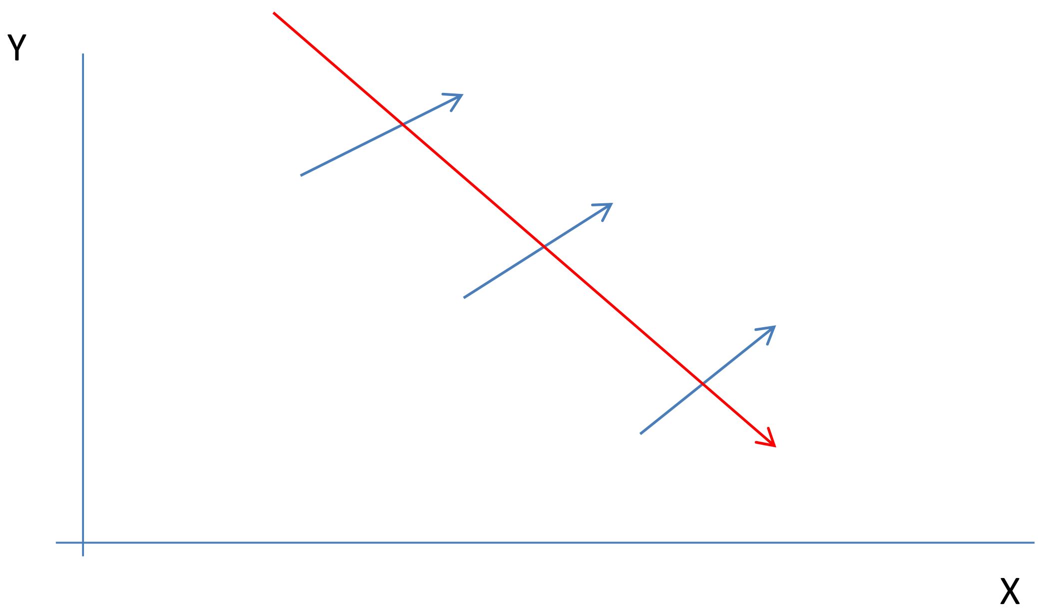
“FIXED EFFECTS” REGRESSION MODEL

Also note, there is only one slope in the model.

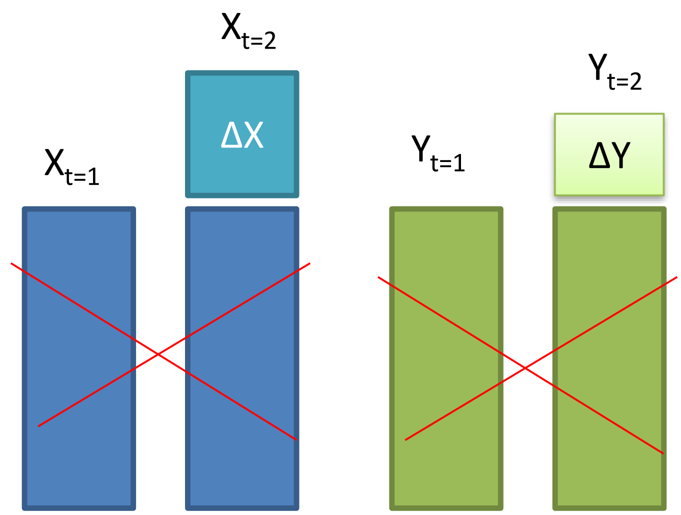
$$BP = b_1 Dosage + b_2 S1 + b_3 S2 + b_4 S3 + e$$



BIAS CAN GO IN THE OTHER DIRECTION



DIFFERENCE IN FIT



Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.852[a]	.726	.698	8.51795

a. Predictors: (Constant), Dosage

b. Dependent Variable: BloodPressure

Model Summary[c,d]

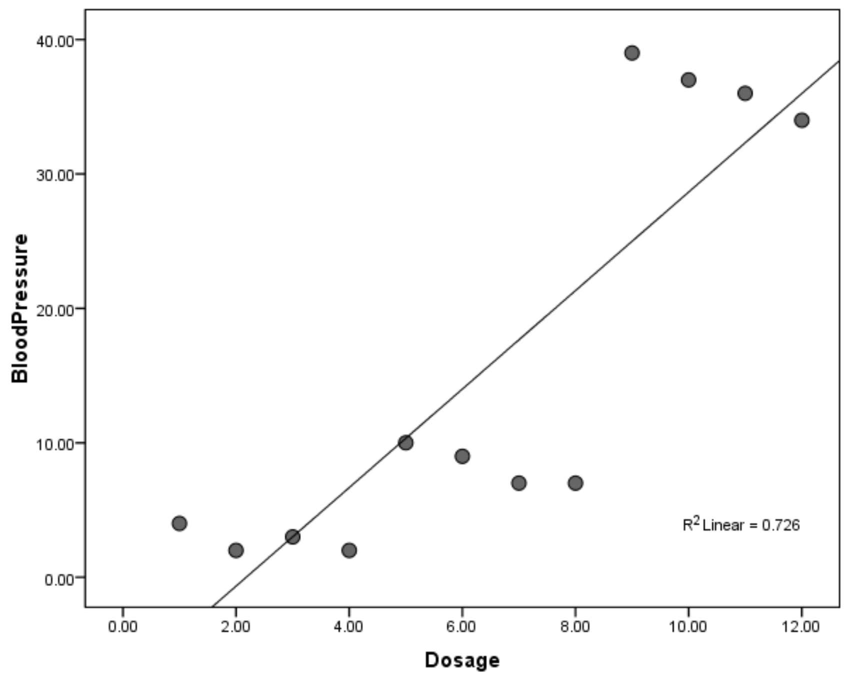
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	1.000[a]	.999	.999	.82412

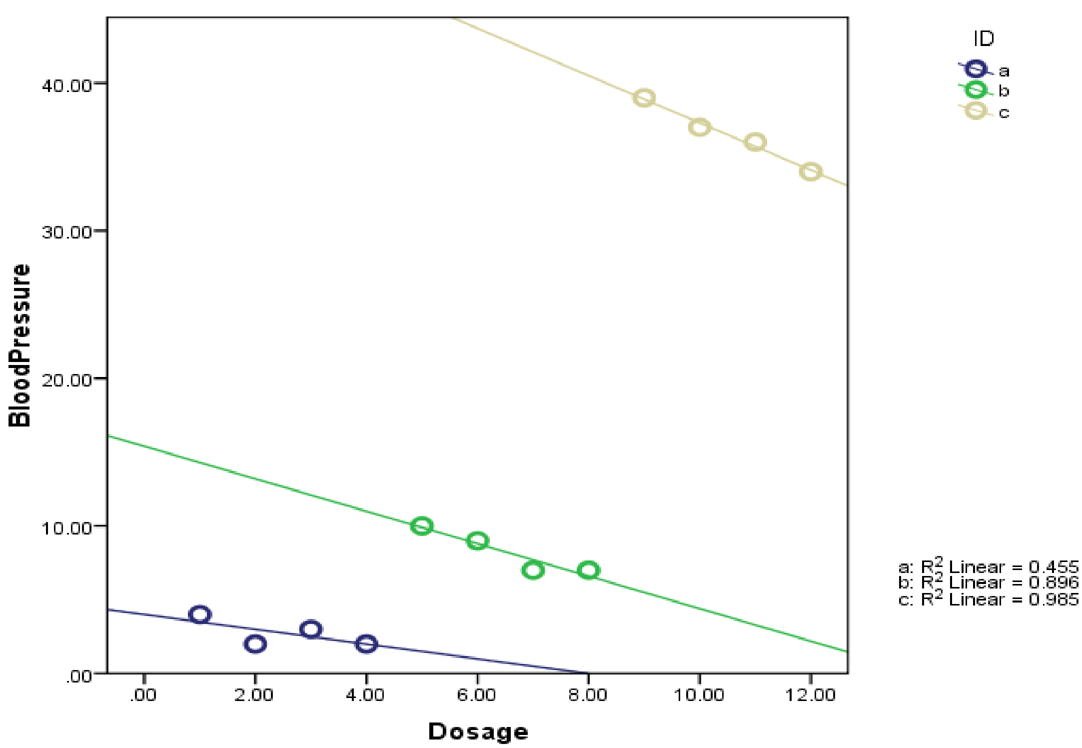
a. Predictors: Subject3, Subject2, Subject1, Dosage

c. Dependent Variable: BloodPressure

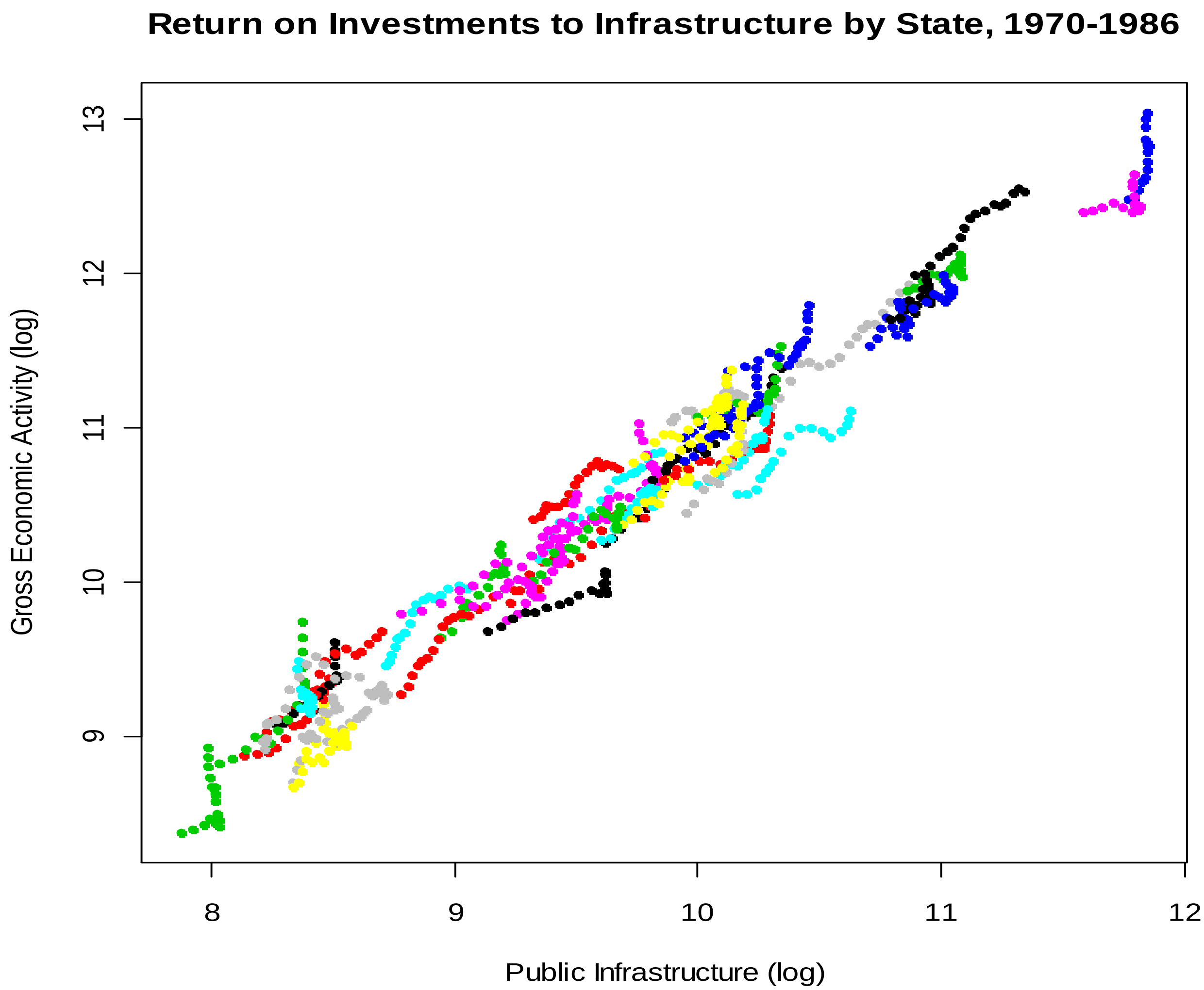
d. Linear Regression through the Origin

DIFFERENCE IN FIT





ECONOMIC DEVELOPMENT POLICY EXAMPLE:



Each color represents data from a state over the 17-year study period.

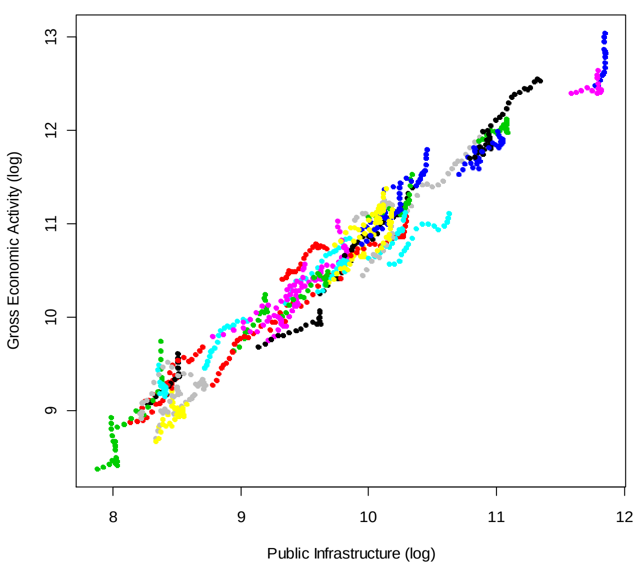
Levels of spending and levels of economic development represent the specific characteristics of a group (a state in this case).

What is our inference about the policy impact of increased infrastructure spending when we use all of the data together, and does it change when we take into account group structure?

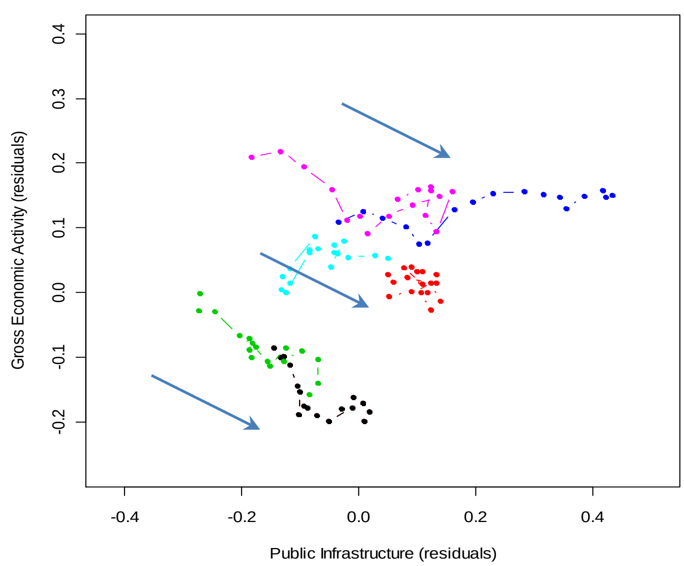
(data from Baltagi & Pinnoi 1995)

EXAMPLE

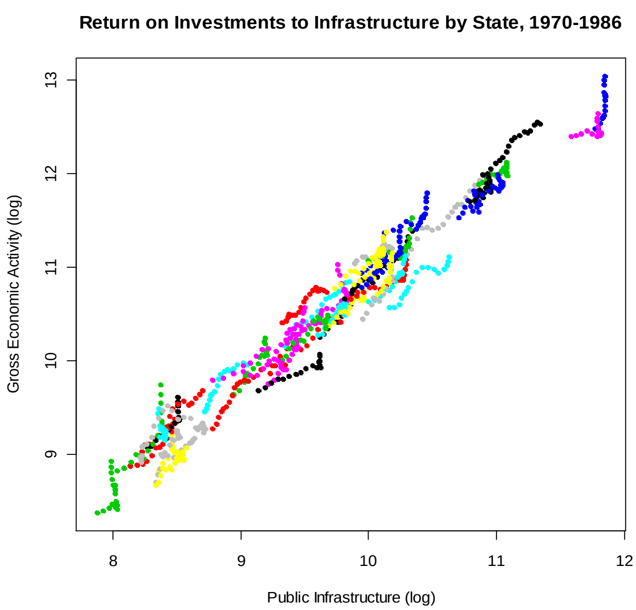
Return on Investments to Infrastructure by State, 1970-1986

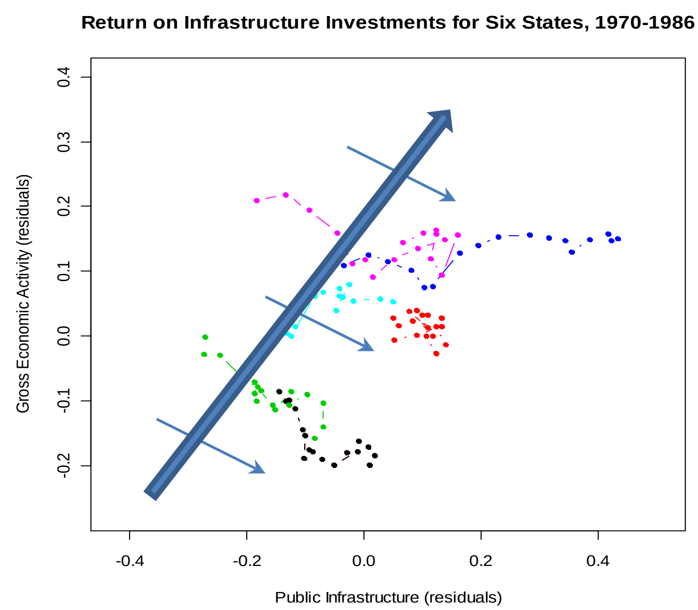


Return on Infrastructure Investments for Six States, 1970-1986



EXAMPLE


Return on Investments to Infrastructure by State, 1970-1986
Gross Economic Activity (log)
Public Infrastructure (log)
8
9
10
11
12
13


Return on Infrastructure Investments for Six States, 1970-1986
Gross Economic Activity (residuals)
Public Infrastructure (residuals)
-0.4
-0.2
0.0
0.2
0.4
0.3
0.1
-0.1

WHY DOES THIS HAPPEN?