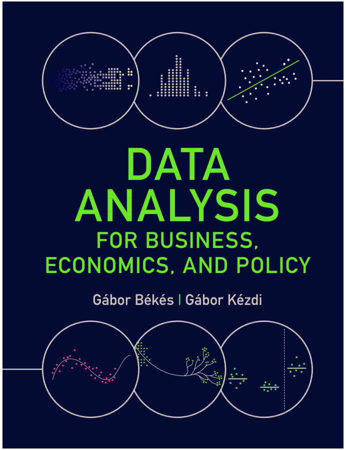
23. Methods for Panel Data

Gábor Békés

Data Analysis 4: Causality

2020

Slideshow for the Békés-Kézdi Data Analysis textbook



- Cambridge University Press, 2021
- **gabors-data-analysis.com**
  - Download all data and code: gabors-data-analysis.com/data-and-code/
- This slideshow is for **Chapter 23**

Multiple Time Periods Can Be Helpful

- Diff-in-diffs estimates the effect at a single point in time.
- Issue 1: Immediate effect, in one period, impact is steady.
- Most real-life situations: delayed effect, variation of impact over time
  - Having a single endline time period is not enough to tell the full story.
- To estimate how an effect plays out in time, need more time periods.
- Issue 2: subjects may be treated at various points in time
- Need method(s) that generalize diff-in-diffs for multiple periods.

Estimating Effects Using Observational Time Series

- Generalization: multiple periods
- Estimating an effect from a single time series: within subject comparisons only.
- An average effect across time for the same subject.
  - we care about a single country / shop; the intervention happens at one place.

- Time series regressions
- specified in levels as well as changes.
  - y_t variable is measured at which t time period. Could have lags.
  - Δ denotes change: $\Delta y_t = y_t - y_{t-1}$

Estimating Effects Using Observational Time Series

- Time series regression specified in levels:

$$y_t^E = \alpha + \beta x_t \tag{1}$$

  - α is the average y when $x = 0$;
  - β shows how much larger y is, on average, when x is larger by one unit.

Estimating Effects Using Observational Time Series

- Time series regression specified in terms of changes in y and changes in x :

$$\Delta y_t^E = \alpha + \beta \Delta x_t \tag{2}$$

  - α: estimates the trend: the average change in y when x doesn't change.
  - β: how much y changes, on average when x increases (or decreases), by one unit; *in addition* to the trend.
    - as y_t changes by the trend anyway, so how much more, is the question.
- Difference: avoid estimating spurious effects due to trends and random walks
  - Applied when x is binary / quantitative, same interpret.

Estimating Effects Using Observational Time Series

- Causal effect? Yes, if variation in Δx_t is exogenous.
- time periods with different changes in x would have experienced the same change in y, had x changed the same way for them.
  - Yes, units are the time periods, as we have a single subject
- Whatever makes x change at time t should be independent of all other things that would make y change at time t.
  - Within-subject criterion: changes in x and y are for the same subject.
  - A version of PTA. In time periods when the treatment status changed ($\Delta x_t \neq 0$), y would have changed the same way, had the treatment status remained the same, as it changed in time periods when the treatment status did remain the same.

Lags to Estimate the Time Path of Effects

- Advantage of multiple time periods: estimate the time path of effects,
  - immediate effects,
  - effects in the near future,
  - long-run effects.
- Include appropriate lags of Δx_t.
  - Application of what we covered earlier
- Causal effect condition the same: when variation in Δx is exogenous.

Lags to Estimate the Time Path of Effects

- With lags, we can estimate effects within the same time period (β_0 below), effects one time period later (β_1),etc.
- Time series regression that can estimate effects for up to K time periods has K lags of Δx:

$$\Delta y_t^E = \alpha + \beta_0 \Delta x_t + \beta_1 \Delta x_{t-1} + ... + \beta_K \Delta x_{t-K} \tag{3}$$

Lags to Estimate the Time Path of Effects

- Long-run effect on y = adding up the coefficients on all lags
- Or apply trick to get cumulative effect:

$$\Delta y_t^E = \alpha + \beta_{cumul} \Delta x_{t-K} + \delta_0 \Delta(\Delta x_t) + ... + \delta_{K-1} \Delta(\Delta x_{t-(K-1)}) \tag{4}$$

  - $\beta_{cumul} = \beta_0 + \beta_1 + ... + \beta_K$ above
  - β_{cumul} shows the total change in y within K time periods after a unit change in x, on average.

Leads to Examine Pre-trends and Reverse Effects

- Another aspect is related to exogeneity of Δx_t
  - impossible to assess directly
  - how y would have changed if x had changed
- Instead: we can examine how y did change in the previous time period(s)
- We need to include **lead terms** of Δx in the regression.
- This is, in fact, the parallel trends assumption we need here: it's analogous to pre-trends in diff-in-diffs regressions

Leads to Examine Pre-trends and Reverse Effects

- Include **lead terms** of Δx in the regression. With L leads:

$$\Delta y_t^E = \alpha + \beta \Delta x_t + \gamma_1 \Delta x_{t+1} + ... + \gamma_L \Delta x_{t+L} \tag{5}$$

  - The lead terms are Δx_{t+1} through Δx_{t+L}.
  - γ_1 shows how y tends to change one time periods before x changes.
  - γ_L shows how y tends to change L time periods before x changes.
  - They show that because Δy_t is one time period **before** Δx_{t+1}, two time periods before Δx_{t+2}, etc.
  - $\gamma_1 = ...\gamma_L = 0$ would show that, regardless of how x changes, y tends to change the same way one through L time periods earlier.

Leads to Examine Pre-trends and Reverse Effects

- Specific case of endogenous change in x – reverse causality effect: y affecting x.
- With observations from multiple time periods - capture this reverse effect.
- IF it takes time.
- Result of reverse effect: a change in x would tend to follow a change in y.
- One time period, Δy_t is associated with Δx_{t+1},
  - coefficient capture that reverse effect

Leads to Examine Pre-trends and Reverse Effects

- Causal model with a single series: combine leads and lags
- The lag terms help capture delayed effects.
- The lead terms help capture differences in pre-trends and reverse effects.
- A time series regression, in differences, with K lags and L leads, has the form

$$\Delta y_t^E = \alpha + \beta_0 \Delta x_t + \beta_1 \Delta x_{(t-1)} + ... + \beta_K \Delta x_{(t-K)} + \gamma_1 \Delta x_{(t+1)} + ... + \gamma_L \Delta x_{(t+L)} \quad (6)$$

Pooled Time Series to Estimate the Effect for One Unit

- Despite the advantages of estimating effects from time series, single time series are rarely used to estimate effects in practice.
- Time series are rarely long enough
- Even if long, are they relevant? Often, not.

- One solution: combine time series from several subjects i (cross-sectional units).
- Idea: time series of similar units are more representative than longer series of a single unit
- Use domain knowledge to select similar units

Pooled Time Series to Estimate the Effect for One Unit

- The simplest pooled time series regression estimates a single intercept and a single slope.
- Most often, though, we include separate intercepts for each i.
- Doing so allows for trends to be different across i.

$$\Delta y_{it}^{E} = \alpha_i + \beta \Delta x_{it} \tag{7}$$

  - Here β shows the average change in y, across time and units i, when x increases by one unit.
  - Conditional on i-specific trends: even if different subjects had different trends, this would not affect our estimate.

Pooled Time Series to Estimate the Effect for One Unit

- We had two ways to tackle serial correlation: Newey-West SE and adding lagged y_t. Here it's the lagged y_t
- Data table with pooled time series, N units, each with T_i observations.
  - There is no specific, ideal N, it's typically 5-20, depends on domain, could be more.
  - Ideally, each unit has same time series, but can work with them even if not —> end of lecture
- We can add leads and lags as before

Import Demand and Industrial Production

- Interested in understanding how external demand affects production
- Thai industrial production and US total imports: individual time series
  - Industrial production in Thailand, in logs, monthly time series
  - US total imports, in logs, monthly time series
- Source: asia-industry dataset. N=243.
  - Monthly data, seasonally adjusted, February 1998–April 2018.

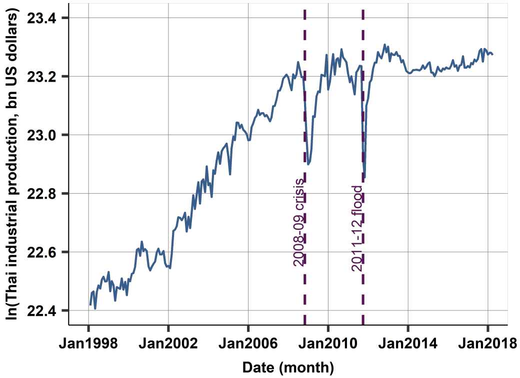
Thai industrial production and US total imports

- Question: how the import demand of the USA affects industrial production in Thailand.
- Causal question, but no explicit intervention.
- what happens in a mid-sized open economy when something changes externally - major trading partner.
- Mechanism: global supply chains, Thailand sells to USA directly, and indirectly (often through China).
- We care about coefficient not just if there is an effect – policy

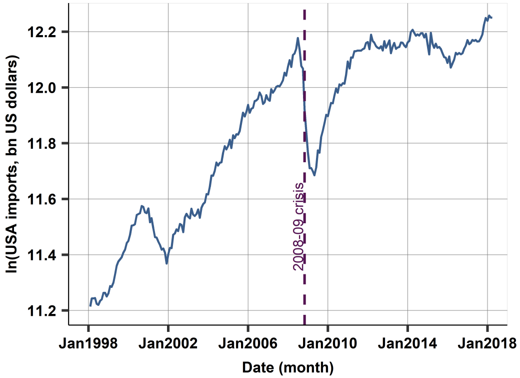
Thai industrial production and US total imports

- Thai industrial production and US total imports: individual time series
  - Industrial production in Thailand, in logs, monthly time series
  - US total imports, in logs, monthly time series
- Source: World Bank WDI – asia-industry dataset. N=243.
  - Monthly data, seasonally adjusted, February 1998–April 2018.

Thai industrial production and US total imports



Thailand IP, in logs, Feb 1998–April 2018, monthly



US total imports, in logs, monthly

Thai industrial production and US total imports [REV]

- There is a trend, an extreme event (2009 great crisis), care about relative change
- First difference. Log values.
- Lags=4 -a one-time change in U.S. imports can have an effect on how Thai industrial production changes through four months.
- No leads - expect no reverse causality
- TS regression estimate the effect of U.S. import demand on Thai industrial production (IP):

$$\Delta(\ln(ipTHA)_t) = \alpha + \beta_0\Delta(\ln(impUSA)_t) + \beta_1\Delta(\ln(impUSA)_{t-1}) + ... + \beta_4\Delta(\ln(impUSA)_{t-4}) + \phi\Delta(\ln(ipTHA)_{t-1}) \tag{8}$$

Import Demand and Industrial Production

- US imports and industrial production in Thailand and three other countries
- Dependent variable is change of log industrial production in each country;
- Explanatory variable cumulative effect of the change in log US imports, four lags.
- Add lagged dependent variable to capture serial correlation
- Monthly time series, seasonally adjusted, February 1998–April 2018. N=243 - for all units

US imports and IP in Thailand + 3 other countries

Variables	(1) Thailand	(2) Malaysia	(3) Philippines	(4) Singapore	(5) Pooled
USA imports log change, cumulative coeff.	0.400*	0.358**	0.556**	0.367	0.437**
	(0.190)	(0.112)	(0.185)	(0.289)	(0.103)
Industrial production log change, lag	-0.119	-0.460**	-0.242**	-0.376**	-0.315**
	(0.065)	(0.059)	(0.064)	(0.061)	(0.031)
Malaysia					0.000
					(0.004)
Philippines					-0.001
					(0.004)
Singapore					0.002
					(0.004)
Constant	0.002	0.004*	0.001	0.005	0.003
	(0.003)	(0.002)	(0.003)	(0.004)	(0.003)
Observations	238	238	238	238	952
R-squared	0.070	0.231	0.140	0.183	0.123

TS regression; dep.var= change of log industrial production in country; log US imports change: 4 lags. Monthly, SA, Feb 1998–April 2018. N=243. Standard error estimates in parentheses. ** $p < 0.01$, * $p < 0.05$.

US imports and IP in Thailand + 3 other countries

- Estimate is 0.44, 95% confidence interval is [0.24,0.64].
- Causality: we have good reasons to take estimate as causal effect
  - First difference takes care of level, trend.
  - Unlikely reverse causality (but may add leads)
- What can go wrong?
- A confounder affecting the **change** in output and demand
- Examples?

Panel Regression

- Pooled time series from a **few** subjects to estimate the expected effect of a causal variable x on outcome y. Policy question was for **one of the subjects**.

- Change of question: the average effect of x on y across many subjects.
- Same kind of question to diff-in-diffs, but multiple periods
- So we'll have: N units, over T periods
  - Typically N is large, T is relatively small
- Will look at different models, approaches

Panel Regression with Fixed Effects

- First model is the **fixed-effects regression** (**FE regression**).
- In FE regressions we have y and x (in levels), panel (xt) data
- Fixed effects are separate intercepts for different cross-sectional units.
- We look for average relationship
- The simplest linear panel regression with cross-section fixed effects:

$$y_{it}^{E} = \alpha_i + \beta x_{it} \tag{9}$$

- The fixed effects are denoted by α_i.
- Intercept varies for different cross-sectional units.

Panel Regression with Fixed Effects

- Why do we include the fixed effects?
  - Separate intercepts for each xsec unit instead of a common intercept?
- IF subjects tend to have higher y on average due to some unobserved confounder that affects x or y in the same way at all times.
- THEN, fixed effects help avoid/mitigate bias.
- Including fixed effects = conditioning on all variables that don't change through time.
  - = Fixed effects condition on time invariant confounders
- Model fit: within R-squared - based on the transformed model, ie comparing mean differenced y and x

Panel Regression with Fixed Effects

- In the FE regression, β shows how much larger y is, on average, compared to its mean within the cross-sectional unit, where and when x is higher by one unit compared to its mean within the cross-sectional unit.
- That's a within-subject comparison, and it's not affected by whether one subject has larger average y.
- That's why it's not affected by whether an unobserved confounder affects the average y values of the different subjects.

Aggregate Trend in panel data

- Aggregate trend is a global trend that affects all unit the same way
- such as global business cycle
- varies across time periods but not units
- With xt panel data, we can can condition on an aggregate trend, whatever form it has, including nonlinear trends or even ups and downs.

Aggregate Trend in panel data

- To condition on aggregate trends, we need to include **time dummies**: binary variables for each time period.
- Sometimes called **time fixed effects**

$$y_{it}^{E} = \alpha_i + \theta_t + \beta x_{it} \tag{10}$$

- β shows how much larger y is, on average, compared to its mean within the cross-sectional units and its mean within the time period, where and when x is higher by one unit compared to its mean within the cross-sectional unit and its mean within the time period.

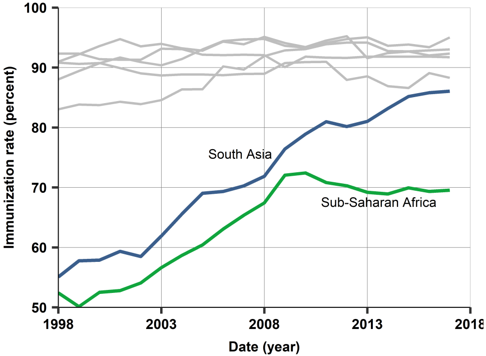
Clustered Standard Errors

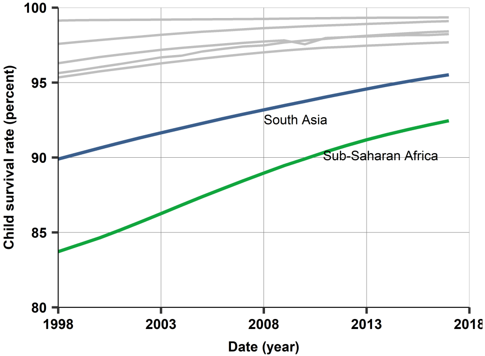
- Instead of heteroskedasticity robust SE (cross-section) or Newey West SE (time series), we'll use a new type called clustered standard error.
- **Standard errors clustered at the level of cross-sectional units**
- Clustered standard errors are robust in two aspects.
  - They are fine in the presence of any kind of serial correlation, and they are also fine without any serial correlation.
  - They are also fine in the presence of heteroskedasticity as well as homoskedasticity
- Thus, with panel models, we always use clustered SE.
  - .... we need a not small (>30) number of units

Immunization against Measles and Saving Children

- Immunization against measles and child survival rate in seven regions of the world
  - Immunization rate
  - Child survival rate
  - Immunization rate: percentage of children of age 12 to 23 months who received vaccination against measles.
  - Child survival rate: 100% minus the percentage of children of age 0 to 5 years who died in the given year.
- Source: worldbank-immunization dataset.
- Annual data, 1998–2017, aggregated to seven geographical regions.
- Many, but not all countries, N=172

Immunization against measles and child survival rate in seven regions of the world





Immunization rate Child survival rate
Source: worldbank-immunization dataset. Annual data, 1998–2017, aggregated to seven geographical regions. N=140

The effect of measles immunization on child survival. FE regressions

- The effect of measles immunization on child survival.
- FE regressions
- Within R-squared presented for FE regressions.
- Source: worldbank-immunization dataset;
- balanced yearly panel, years 1998–2017 in 172 countries.

The effect of measles immunization on child survival. FE regressions

Variables	(1) Survival rate	(2) Survival rate
Immunization rate	0.077**	0.038**
	(0.010)	(0.011)
ln GDP per capita		1.593**
		(0.399)
ln population		12.049**
		(1.648)
Year dummies	Yes	Yes
Observations	3,440	3,440
R-squared	0.717	0.848
Number of countries	172	172

Within R-squared presented for FE regressions. Appropriate standard error estimates in parentheses. ** p <0.01, * p <0.05. Source: `worldbank-immunization dataset`; balanced yearly panel, years 1998–2017 in 172 countries.

The effect of measles immunization on child survival. FE regressions

- The slope parameter estimate on immunization is 0.077 without conditioning on any confounders
- drops to 0.038 when we condition on GDP per capita and population
- **When we compare years with the same GDP and population, in years when the immunization rate is higher by 10 percentage points than its average rate within a country, child survival tends to be 0.38 percentage points higher than its average within the country, conditional on aggregate trends in the world.**
- We can expect it to be 0.16 to 0.6 percentage points higher in the general pattern represented by our data.
  - 100 percent - 3.8 percent, but not a realistic improvement, 10% makes more sense

The effect of measles immunization on child survival.

FE regressions with different Simple and Clustered SE estimates.

Variables	(1) Clustered SE	(2) Simple SE
Immunization rate	0.038**	0.038**
	(0.011)	(0.002)
ln GDP per capita	1.593**	1.593**
	(0.399)	(0.071)
ln population	12.049**	12.049**
	(1.648)	(0.227)
Observations	3,440	3,440
R-squared	0.848	0.848
Number of countries	172	172

Within R-squared presented for FE regressions. Standard error estimates in parentheses. ** $p<0.01$, * $p<0.05$.
Source: `worldbank-immunization dataset`; balanced yearly panel, years 1998–2017 in 172 countries.

Panel Regression in First Differences

- We can also specify xt panel regressions in changes.
- Different approach, alternative to FE model in modeling
- **panel regression in first differences** or **FD regression**.
- FD = changes -> $\Delta y_{it} = y_{it} - y_{i(t-1)}$.
- FD panel regression with a common intercept across all i.

$$\Delta y^E_{it} = \alpha + \beta \Delta x_{it} \tag{11}$$

- Looks like a pooled a cross-section with first difference.
- But here, we have a single intercept, α

Panel Regression in First Differences

- β shows the difference in the average change of y for units that experience a change in x during the same period.
- Comparing different cross-sectional units for the same time, or comparing different time periods for the same unit, β shows how much more y changes, on average, where and when x increases by one unit.

Lags and Leads in FD Panel Regressions

- Often, we want to estimate not only immediate effects but longer run effects, too.
- Multiple time periods allow us to capture the time path of the effects by including lags of Δx in the regression.
- Same idea as with pooled time series

- Regression in FD with K lags:

$$\Delta y_{it}^{E} = \alpha + \beta_0 \Delta x_{it} + \beta_1 \Delta x_{i(t-1)} + \ldots + \beta_K \Delta x_{i(t-K)} \tag{12}$$

- **cumulative effect** or long-run effect of the change of x = sum of the immediate effect and all lagged effects.

Lags and Leads in FD Panel Regressions

- We can also add lead terms to an FD regression to examine pre-trends and capture reverse effects, just like with single time series.
  - Better than inspecting pre-trends, but PTA remains an assumption.
- An FD panel regression with K lags and L leads looks like this:

$$\Delta y_{it}^{E} = \alpha + \beta_0 \Delta x_{it} + \beta_1 \Delta x_{i(t-1)} + \ldots + \beta_K \Delta x_{i(t-K)} + \gamma_1 \Delta x_{i(t+1)} + \ldots + \gamma_L \Delta x_{i(t+L)} \quad (13)$$

  - The γ coefficients on the lead terms are zero if, prior to time periods when x may change, y tends to change the same way regardless of whether and how much x actually changes.

Aggregate Trend in FD Models

- As for FE models, we can add time dummies to capture non-linear trend
- FD regression with K lags and time dummies (time FE) is the following:

$$\Delta y_{it}^{E} = \theta_t + \beta_0 \Delta x_{it} + \beta_1 \Delta x_{i(t-1)} + ... + \beta_K \Delta x_{i(t-K)} \tag{14}$$

- θ_t = coefficients of the time dummies
  - = time-specific intercepts = time fixed effects.

Individual Trends in FD Models

- Time dummies capture an aggregate trend in a completely flexible way
- Cross-sectional units in the data may have their own trends, too.
  - Here we don't have the opportunity to estimate flexible trends, because we have only one observation for each time period for each unit.

- Can capture **individual linear trends**: allow the intercept to be different across cross-sectional units.
  - trend = average change per unit
  - as with pooled time series

Individual Trends in FD Models

- FD regression with K lags, time dummies, and individual-specific intercepts:

$$\Delta y_{it}^{E} = \alpha_i + \theta_t + \beta_0 \Delta x_{it} + \beta_1 \Delta x_{i(t-1)} + \ldots + \beta_K \Delta x_{i(t-K)} \tag{15}$$

- α_i: the average change in y in cross-sectional unit i across all time periods
  - measured as a deviation from the flexibly estimated aggregate trend θ_t,
  - and when x does not change (and didn't change for the past K time periods).

Immunization against Measles and Saving Children

- The immediate and lagged effect of measles immunization on child survival
- FD panel regression estimates
- Cumulative effect estimates calculated via transformation.
- Clustered standard error
- balanced yearly panel, years 1998–2017 in 172 countries.

Variables	(1) $\Delta surv$	(2) $\Delta surv$	(3) $\Delta surv$	(4) $\Delta surv$
Δimm	0.009**	0.010**		
	(0.002)	(0.002)		
Δimm lag 1		0.010**		
		(0.002)		
Δimm lag 2		0.011**		
		(0.002)		
Δimm lag 3		0.009**		
		(0.002)		
Δimm lag 4		0.007**		
		(0.002)		
Δimm lag 5		0.006**		
		(0.002)		
Δimm lead 1				0.008**
				(0.002)
Δimm lead 2				0.007**
				(0.002)
Δimm lead 3				0.005
				(0.003)
Δimm cumul			0.053**	0.054**
			(0.010)	(0.008)
Constant	0.188**	0.136**	0.136**	0.125**
	(0.024)	(0.018)	(0.018)	(0.018)
R-squared	0.013	0.078	0.078	0.093
Observations	3,268	2,408	2,408	1,892

Immunization against Measles and Saving Children

- The effect of measles immunization on child survival. FD panel regression estimates with year dummies, confounders, and country-specific trends
- FD panel regressions with 5 lags of all right-hand-side variables.
  - Cumulative coefficient on the change of immunization over the 5 lags.
  - Clustered standard error estimates in parentheses.
- Adding leads - 3 periods

The effect of measles immunization on child survival

The effect of measles immunization on child survival - FD model estimates

Variables	(1) $\Delta surv$	(2) $\Delta surv$	(3) $\Delta surv$
Δimm cumulative ,	0.052**	0.030**	0.011**
	(0.010)	(0.009)	(0.003)
Year dummies	Yes	Yes	Yes
Confounder variables	No	Yes	Yes
Country-specific trends	No	No	Yes
Observations	2,408	2,408	2,408
R-squared	0.088	0.212	0.331

FD panel regressions with 5 lags of all right-hand-side variables. Confounders: GDP per cap, population. Cumulative coefficient w 5 lags. Clustered SE estimates in parentheses. ** p<0.01, * p<0.05. Source: `worldbank-immunization dataset`; balanced yearly panel, years 1998–2017 in 172 countries.

The effect of measles immunization on child survival

- Baseline result 0.05
- Year dummies + confounders: 0.030 - confounders clearly important
- Adding individual linear time trend: 0.011 - small but precisely measured

- Causal effect?

- We can't be certain. It's observational data.

- We did a great deal of efforts to condition on all kinds of confounders.
  - FD model with lags - takes out level differences and accounts for dynamics
  - Key confounders added: GDP per capita and population + individual linear trends
  - PTA - make a very good effort: Adding leads or confounders like population, gdp makes no difference.

- Good approximation to what the true effect: A 10 percent increase in the immunization rate **leads to** a 0.1 percentage point increase in the child survival rate within five years.

Panel Regressions and Causality

- FE regressions and FD regressions can estimate the effect of x on y without the bias due to confounders that don't change over time.
- Confounders that change through time need to be observed and included in the FE or FD regression.
- Conditioning on individual trends is feasible with FD regressions
  - Can do something similar in FE, but (even more) complicated

- Panel model allow us conditioning on a great deal of confounding factors
- But, as always, there can be omitted variables - so never certain.

First Differences or Fixed Effects?

- Have seen many models, which one to choose?
- FE and FD regressions are similar because both condition on confounders that affect the level of y and x and don't change through time.
  - FE regressions do that by comparing values of y and x to their cross-sectional means.
  - FD regressions do something similar by comparing values of y and x to their values in the previous time period.
- Confounders that affect the change in y or x still matter for both FE and FD regressions, whether the confounders themselves change through time or not

First Differences or Fixed Effects?

- FD main advantage 1: capture serial correlation by first differencing
  - important if time series properties key
- FD main advantage 2: capture transparent dynamics
- As long as we keep adding lags. But that means smaller and smaller panel for estimation.
  - FD takes care of linear trend automatically, but as we add anyway, no big deal
- FD main advantage 3: can easily capture individual linear trends

First Differences or Fixed Effects?

- FE main advantage 1: simple method of estimating longer run effects, easier to use
  - estimate of the average of short-term and long term effects.
  - When the long-term effects kick in fast, that's a good approximation of the long-term effects themselves
- FE main advantage 2: Works when missing values in panel (see next bit)
- In many cases, both FD and FE can work.
- Key consideration is if time path to effect matters

Dealing with Unbalanced Panels

- Missing observations: missing at random or not
- If missing at random - okay to keep. Maybe FE models will be better.
- If not
  - Reduce T - focus only on more recent years when coverage is high
  - Reduce N - drop unit (countries) where coverage is low
- Sample design (filtering out observation) means we have a different sample, and may not be representative to what we started with.
- Many analytical choice, but must make notes

Summary: Panel Regression

- Data with multiple time periods can help uncover short- and long-run effects and examine pretrends.
- When interested in the effects on a single cross-sectional unit, we may analyze a single time series or pool several time series of similar units.
- With panel data having multiple time periods, several modeling options
- use an FD regression to uncover the development of the effect over time, and an FD or an FE regression to uncover the long-run effect
- Watch out for interpretation - hard

- Overall big picture: using panel data methods can take us much closer to a causal interpretation.