Dr. Christian Czymara

EINFÜHRUNG IN DIE PANELREGRESSION

Day 2

DeZIM Summer School 2023

AGENDA

- Decomposition of variance into within and between part
- The logic of Fixed Effects (FE) models
- Benefits and limitations of FE
- Random Effects (Within-Between) models
- Logistic FE and RE models / linear probability FE

THE POPULARITY OF FIXED EFFECTS IN SOCIOLOGY

Source: Hill, Davis, Roos & French (2020). [Limitations of fixed-effects models for panel data](). Sociological Perspectives, 63(3): 359

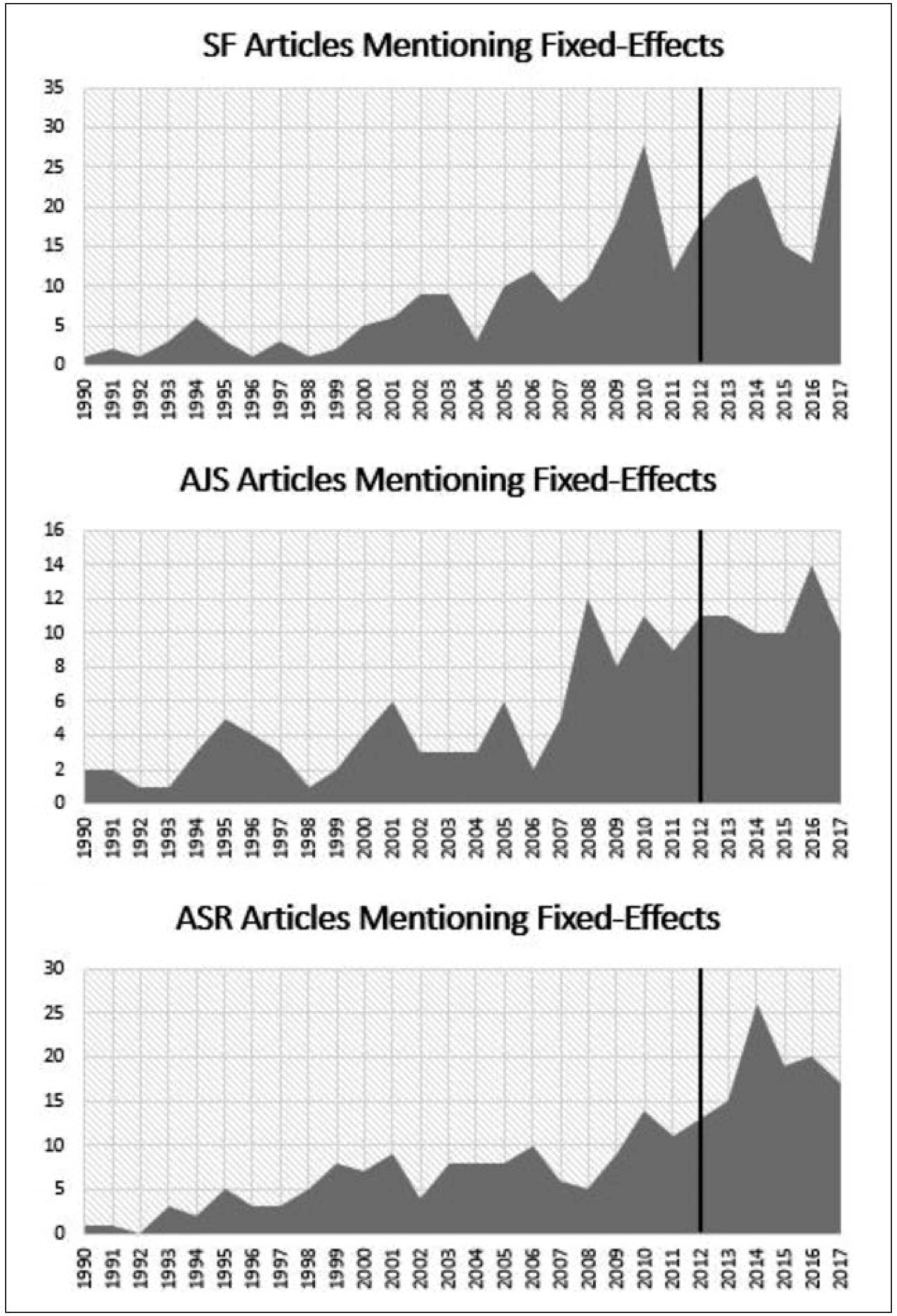
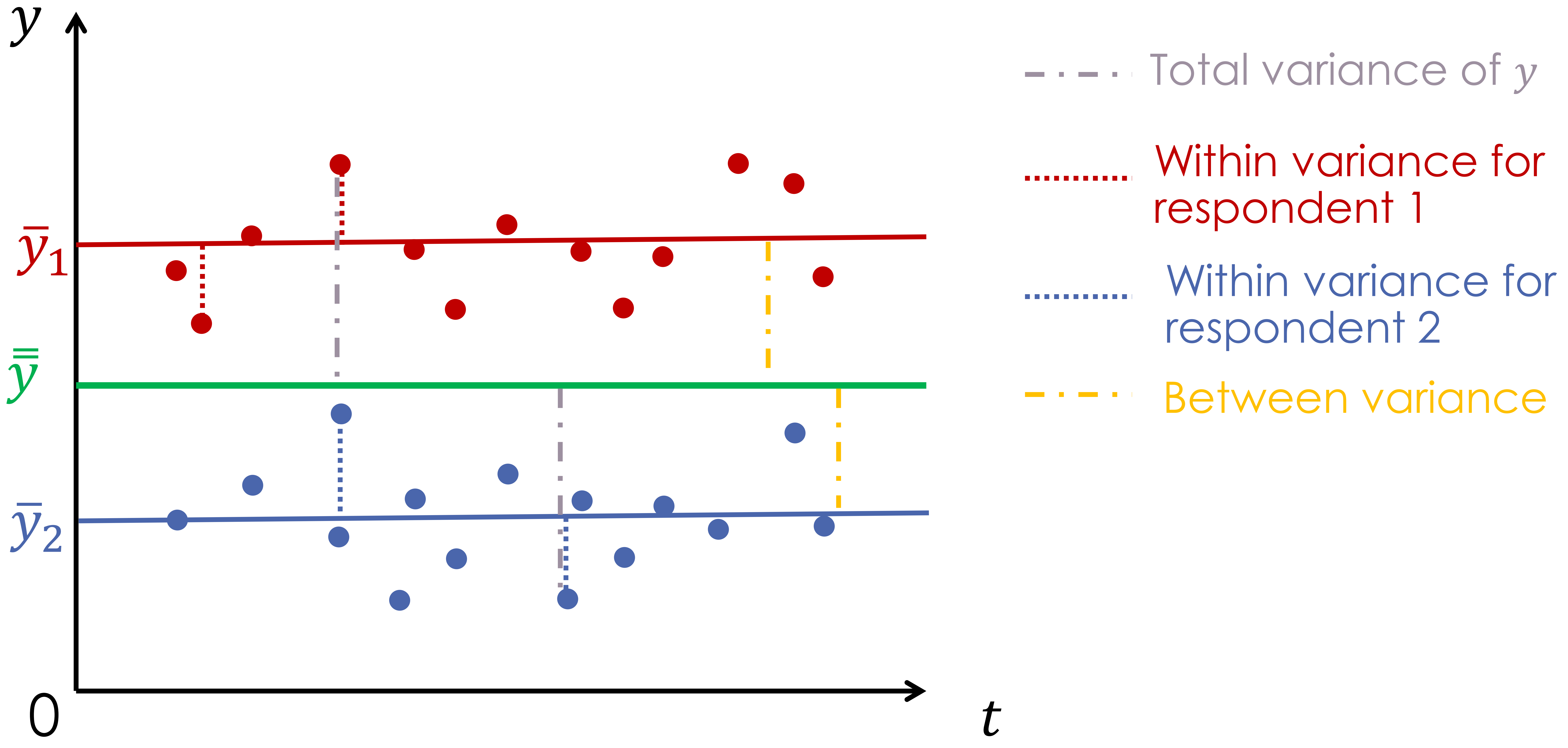


Figure 1. The growing number of articles mentioning fixed-effects (1990–2017) with demarcation for the current study period (2012–2017) by *Social Forces* (SF), *American Journal of Sociology* (AJS), and *American Sociological Review* (ASR).

WITHIN AND BETWEEN VARIANCE IN PANEL DATA

WITHIN AND BETWEEN VARIANCE



WITHIN AND BETWEEN VARIANCE

- Within variance: within one individual over time → $(y_{it} - \bar{y}_i)$
- Between variance: between individuals → $(\bar{y}_i - \bar{\bar{y}})$

ID	Year	y_{it}
1	2009	0.58
1	2010	0.88
1	2011	0.04
2	2009	0.66
2	2010	0.22
2	2011	0.5
3	2009	0.3
3	2010	0.3
3	2011	0.3

Individual-specific mean $(\bar{y}_i)$

0.50

0.46

- 0.30

Overall mean $(\bar{\bar{y}})$

0.42

WITHIN AND BETWEEN VARIANCE

- $\bar{\bar{y}} = 0.42$
- $\bar{y}_1 = 0.50$
- $\bar{y}_2 = 0.46$
- $\bar{y}_3 = 0.30$

ID	Year	y_{it}
1	2009	0.58
1	2010	0.88
1	2011	0.04
2	2009	0.66
2	2010	0.22
2	2011	0.5
3	2009	0.3
3	2010	0.3
3	2011	0.3

Overall variance: $(y_{it} - \bar{\bar{y}})$	Within variance: $(y_{it} - \bar{y}_i)$	Between variance: $(\bar{y}_i - \bar{\bar{y}})$
0.16	0.08	0.08
0.46	0.38	0.08
-0.38	-0.46	0.08
0.24	0.2	0.04
-0.2	-0.24	0.04
0.08	0.04	0.04
-0.12	0	-0.12
-0.12	0	-0.12
-0.12	0	-0.12

UNOBSERVED HETEROGENEITY

RECAP: OMITTED VARIABLE BIAS

- OLS yields biased effects if confounding variables are omitted
- Omitted variables ≙ *unobserved heterogeneity*
- So… How can we use panel data to estimate unbiased effects if there is correlated unobserved heterogeneity?

PANEL DATA MODEL

- Adding an index for time: $y_{it} = \beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + \varepsilon_{it}$
- Differentiating between time-constant and time-variant variables:
$$y_{it} = \beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + \gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i + e_{it}$$
with
  - $i=1,\ldots,n$ units
  - $t=1,\ldots T$ observations
  - k time-varying variables x
  - l time-constant variables z
- Decomposition of the error term: $\varepsilon_{it} = u_i + e_{it}$

UNOBSERVED EFFECTS MODEL

- $y_{it} = \beta_0 + \underbrace{\beta_1 x_{1it} + \cdots + \beta_k x_{kit} + e_{it}}_{} + \underbrace{\gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i}_{}$

- Time varying characteristics
- Variables (x), for example: grades, attitudes, income, …
- Idiosyncratic error (e_{it}): All sources of time-varying variance not captured by x, treated similar to error term in OLS

- Time constant characteristics
- Variables (z), for example: country of birth, date of graduation, … (?)
- Unobserved heterogeneity (u_i): All time-constant sources of variation not captured by z

UNOBSERVED HETEROGENEITY: EXAMPLE

- True model:

$$mortality_{it} = \beta_0 + \beta_1 migration_{it} + \beta_2 health_i + u_i + e_{it}$$

- $health$ not observed:

$$mortality_{it} = \beta_0 + \beta_1 migration_{it} + u_i + e_{it}$$

- The error term is now correlated with the variables in the model (remember session ii)
- u_i includes $health$, a confounding variable that correlates with $migration$ and $mortality$(in this example)

CORRELATED UNOBSERVED HETEROGENEITY

- Analogous to OLS , unobserved effects model yields biased estimates if error terms (u_i or e_{it}) correlate with variables in the model
- Solution: Control everything that is time-constant of each unit (here: person)
- $mortality_{it} = \beta_0 + \beta_1 migration_{it} + u_i + e_{it}$
- u_i as something "typical" for person i
- Part of u_i might be observed, but other parts might not
- How can we control for such stable idiosyncrasies?

FIXED EFFECTS

LEAST SQUARE DUMMY VARIABLES

- Let say person i is a person which has been interviewed several times
- One solution: Control for individual i

→ Add a dummy for individual i (1: interviews of individual i, 0: interviews of all other respondents)

- Because we observe every person multiple times, we could add dummies for *all persons* without exhausting degrees of freedom

LEAST SQUARE DUMMY VARIABLES

- Include a dummy variable for each person (not person-year!)
- $mortality_{it} = \beta_0 + \beta_1 migration_{it} + \gamma_1\delta_1 + \cdots + \gamma_n\delta_n + u_i + e_{it}$
- Such a model is called a Least Square Dummy Variables (LSDV) regression
- Model yields so-called *fixed effects estimates*
- *Fixed* Effects because each unit has a specific fixed effect on the dependent variable

FIXED EFFECTS-TRANSFORMATION

- Including dummy variables for each unit might not always be feasible
- Another way to obtain results: Fixed Effects-Transformation
- Instead of controlling u_i, we eliminate it from the regression function

MEANS

- $t = 1$: $mortality_{i1} = \beta_0 + \beta migration_{i1} + u_i + e_{i1}$
- $t = 2$: $mortality_{i2} = \beta_0 + \beta migration_{i2} + u_i + e_{i2}$
- Mean: $\overline{mortality}_{i.} = \beta_0 + \beta\overline{migration}_{i.} + \bar{u}_i + \bar{e}_{i.}$

→ $\overline{mortality}_{i.} = \beta_0 + \beta\overline{migration}_{i.} + u_i + \bar{e}_{i.}$

TIME-DEMEANING AT T=1

- $(t = 1)\ - mean$:
- $\left(mortality_{i1} - \overline{mortality}_{i.}\right)$

$$= (\beta_0 + \beta migration_{i1} + u_i + e_{i1}) - (\beta_0 + \beta \overline{migration}_{i.} + u_i + \bar{e}_{i.})$$

$$= \beta(migration_{i1} - \overline{migration}_{i.}) + (u_i - u_i) + (\bar{e}_{i1} - \bar{e}_{i.})$$

$$= \beta(migration_{i1} - \overline{migration}_{i.}) + (\bar{e}_{i1} - \bar{e}_{i.})$$

TIME-DEMEANING AT T=2

- $(t = 2)\ - mean$:
- $\left(mortality_{i2} - \overline{mortality}_{i.}\right)$

$$= (\beta_0 + \beta migration_{i2} + u_i + e_{i2}) - (\beta_0 + \beta \overline{migration}_{i.} + u_i + \bar{e}_{i.})$$

$$= \beta(migration_{i2} - \overline{migration}_{i.}) + (u_i - u_i) + (\bar{e}_{i2} - \bar{e}_{i.})$$

$$= \beta(migration_{i2} - \overline{migration}_{i.}) + (\bar{e}_{i2} - \bar{e}_{i.})$$

TIME-DEMEANING

- Time-demeaning of panel data
- $\left(mortality_{it} - \overline{mortality_{i.}}\right) = \beta(migration_{it} - \overline{migration_{i.}}) + (\bar{e}_{it} - \bar{e}_{i.})$
- All estimates are based on within-unit variation over time
- All between-unit variance (time stable difference between persons) is removed from the data

FIXED EFFECTS TRANSFORMATION

- $t = 1$: $y_{i1} = \beta_0 + \beta_1 x_{1i1} + \cdots + \beta_k x_{ki1} + \gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i + e_{i1}$
- $t = 2$: $y_{i2} = \beta_0 + \beta_1 x_{1i2} + \cdots + \beta_k x_{ki2} + \gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i + e_{i2}$
- $t = T$: $y_{iT} = \beta_0 + \beta_1 x_{1iT} + \cdots + \beta_k x_{kiT} + \gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i + e_{iT}$
- Mean:

$$\bar{y}_{i.} = \beta_0 + \beta_1 \bar{x}_{1i.} + \cdots + \beta_k \bar{x}_{ki.} + \gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i + \bar{e}_{i.}$$

FIXED EFFECTS TRANSFORMATION

- Mean: $\bar{y}_{i.} = \beta_0 + \beta_1\bar{x}_{1i.} + \cdots + \beta_k\bar{x}_{ki.} + \gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i + \bar{e}_{i.}$
- $t - mean$: $(y_{it} - \bar{y}_{i.}) =$

$$(\beta_0 + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + \gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i + e_{it}) -$$
$$(\beta_0 + \beta_1\bar{x}_{1i.} + \cdots + \beta_k\bar{x}_{ki.} + \gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i + \bar{e}_{i.})$$

$\rightarrow (y_{it} - \bar{y}_{i.}) = \beta_1(x_{1it} - \bar{x}_{1i.}) + \cdots + \beta_k(x_{kit} - \bar{x}_{ki.}) + \gamma_1(z_{1i} - z_{1i}) + \cdots + \gamma_l(z_{li} - z_{li}) + (u_i - u_i) + (e_{it} - \bar{e}_{i.})$

$\rightarrow (y_{it} - \bar{y}_{i.}) = \beta_1(x_{1it} - \bar{x}_{1i.}) + \cdots + \beta_k(x_{kit} - \bar{x}_{ki.}) + (e_{it} - \bar{e}_{i.})$

$\rightarrow \ddot{y}_{it} = \beta_1\ddot{x}_{1it} + \cdots + \beta_k\ddot{x}_{kit} + \ddot{e}_{it}$

RECAP: TRANSFORMING THE DATA

ID	Year	y_{it}
1	2009	0.58
1	2010	0.88
1	2011	0.04
2	2009	0.66
2	2010	0.22
2	2011	0.5
3	2009	0.3
3	2010	0.3
3	2011	0.3

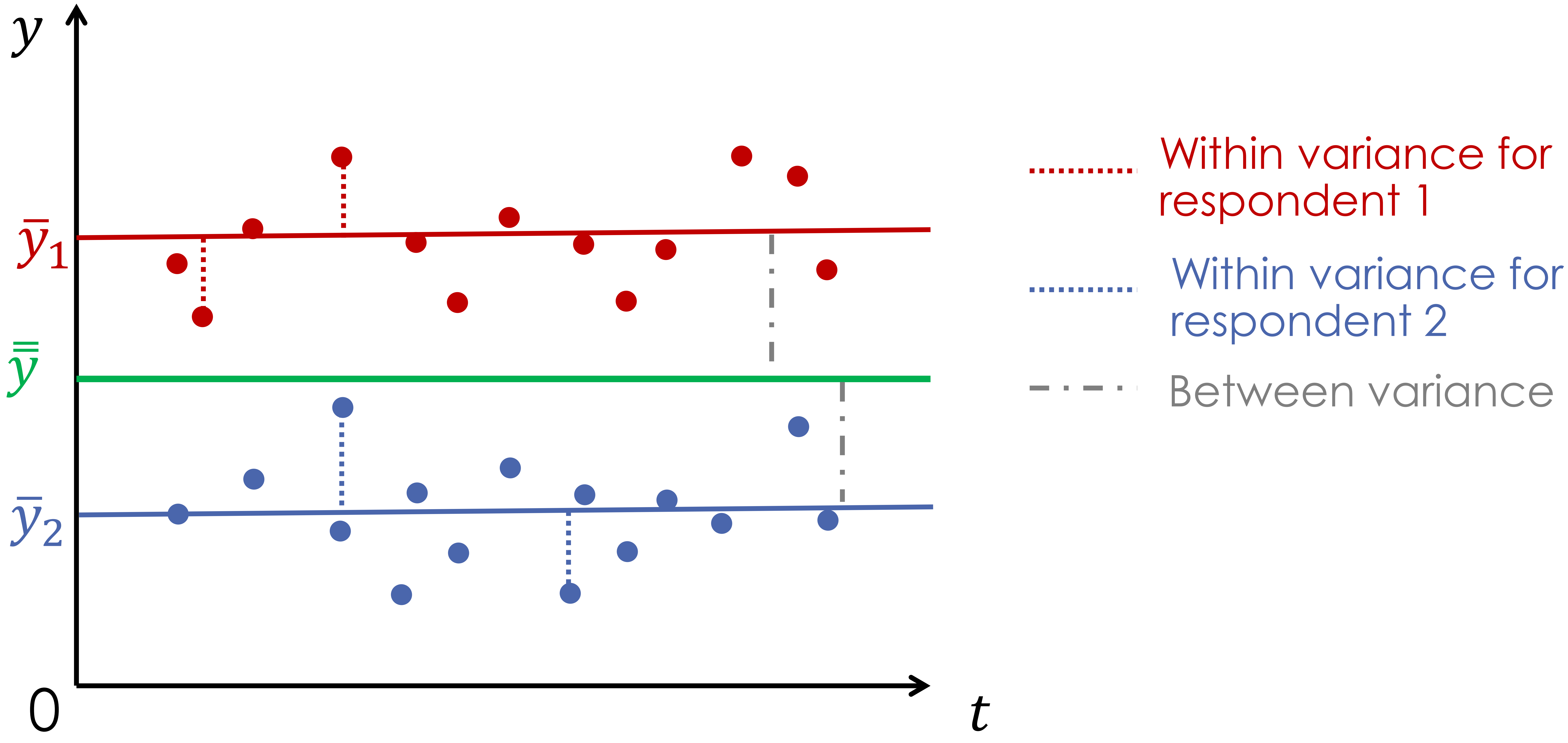
Overall variance: $(y_{it} - \bar{\bar{y}})$	Within variance: $(y_{it} - \bar{y}_i)$	Between variance: $(\bar{y}_i - \bar{\bar{y}})$
0.16	0.08	0.08
0.46	0.38	0.08
-0.38	-0.46	0.08
0.24	0.2	0.04
-0.2	-0.24	0.04
0.08	0.04	0.04
-0.12	0	-0.12
-0.12	0	-0.12
-0.12	0	-0.12

RECAP: TRANSFORMING THE DATA

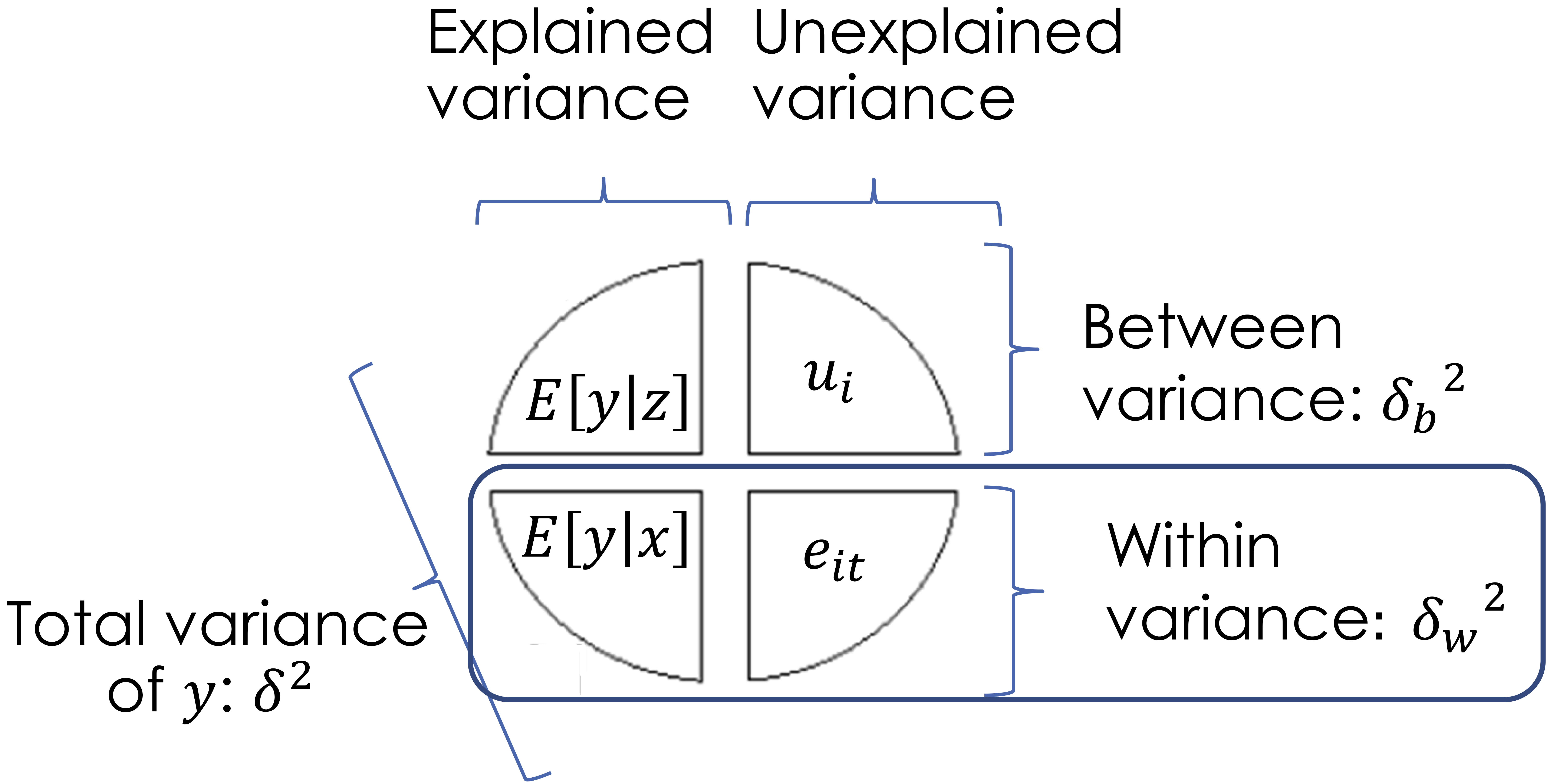
ID	Year	y_{it}
1	2009	0.58
1	2010	0.88
1	2011	0.04
2	2009	0.66
2	2010	0.22
2	2011	0.5
3	2009	0.3
3	2010	0.3
3	2011	0.3

Overall variance: $(y_{it} - \bar{\bar{y}})$	Within variance: $(y_{it} - \bar{y}_i)$	Between variance: $(\bar{y}_i - \bar{\bar{y}})$
0.16	0.08	0.08
0.46	0.38	0.08
-0.38	-0.46	0.08
0.24	0.2	0.04
-0.2	-0.24	0.04
0.08	0.04	0.04
-0.12	0	-0.12
-0.12	0	-0.12
-0.12	0	-0.12

RECAP: WITHIN AND BETWEEN VARIANCE



COMPOSITION OF y



WHY DO DEMEANED FE AND LSDV YIELD THE SAME EFFECTS?

- Both are practically linear models using the OLS estimator
- Including dummies for units partials out the effects of individuals
- What is left is independent of all differences between individuals or in other words: rid of between variance
- Dummy variables capture all (also unmeasured) time-constant characteristics of individuals
- Thus, you also get the FE estimates when you control for the unit-specific means of each variable

WHY DO FE AND LSDV *NOT* YIELD THE SAME STANDARD ERRORS?

- *N* seems to be the same (data points)
- But time-demeaning actually costs degrees of freedom because it uses information from the data (the unit-specific means)
- … Or LSDV: each dummy costs one degree of freedom
- Running a linear model with manually demeaned data does not account for this
- Hence, significance tests need to be corrected manually
- If they are not, OLS with manually demeaned variables yields underestimated standard errors

MODELLING TIME TRENDS IN FE

ONE- VS. TWO-WAY FE

- Person FE: Average change in y if x increases by one unit *over time*
- Time FE: Average change in y if x increases by one unit *between cases*
- Two-way FE: Average difference in within-person changes in y at time point t for each one unit increase in x at t, averaged over all t

→"*two-way FE model unhelpfully combines within-unit and cross-sectional variation in a way that produces un-interpretable answers.*" (Kropko & Kubinec 2020: 1)

FIXED EFFECTS INDIVIDUAL SLOPES

- FE assume parallel trends between treated and untreated
- I.e.: Both groups would follow the same over-time trend in y if x wouldn't change
- For example: Does marriage increase hourly wage for men? → Men who eventually get married show steeper wage growth even before marriage
- See Rüttenauer & Ludwig (2020)

LIMITS OF FIXED EFFECTS

See Hill, Davis, Roos & French (2020). Limitations of fixed-effects models for panel data. *Sociological Perspectives* 63 (3) 357 - 369.

LOW STATISTICAL POWER

- Observations without temporal variation do not contribute to FE estimator by design

→Reduced sample size

→Low statistical power (high standard errors)

- Observations with little temporal variation contribute little to FE estimator

→Coefficients are based on small number of observations

→Limited reliability ("Silly estimators", Beck & Katz 2001: 494)

→Increased Type II error rate (false negative)

- *Statistically significant FE estimate likely robust, but non-significant FE may be due to low power*

EXTERNAL VALIDITY

- Observations with little temporal variation contribute little to FE estimator

→Model of temporal changes only apply to a specific subgroup of observations

→Subgroup might differ from broader population (i.e. sample might no longer be representative)

→P-values might have less statistical meaning

- FE are treatment effects on treated (only units with change are observed), OLS (theoretically) are average treatment effects

OTHER ISSUES OF FE

- Estimates less reliable with less time periods
- Repeated measurement error (overly conservative estimates)
- FE useless for estimating time stable differences
- Unclear which variables are time-stable or varying
- FE only control *time-stable* effects of time-stable variables

LIMITS TO CAUSAL INFERENCE

1. Time-varying confounders (erogeneity assumption)
2. Reverse causality (y affecting x)
3. Lagged effects (past x affecting current y)

- All would be solved by including all time-varying confounders, but how realistic is that?

SUMMING UP

SUMMARY

- FE eliminate any between-unit variance from the data

→Estimates only based on within-unit variation

→Automatically control for unobserved heterogeneity (everything time-constant)

- … Which is a huge step forward for estimating unbiased effects in many cases

SUMMARY

- However, time-constant variables drop out (effects of constants cannot be estimated, just like OLS)
- … But interactions between time-constant and time-varying variables can still be estimated → Does the effect of x depend on z?
- Often more crucial: Many aspects might not be totally constant but empirically vary only little over time
- FE "*may kill some of the omitted variables bias bathwater, but they also remove much of the useful information in the baby, the variable of interest.*" (Angriest & Pischke 2009: 225)

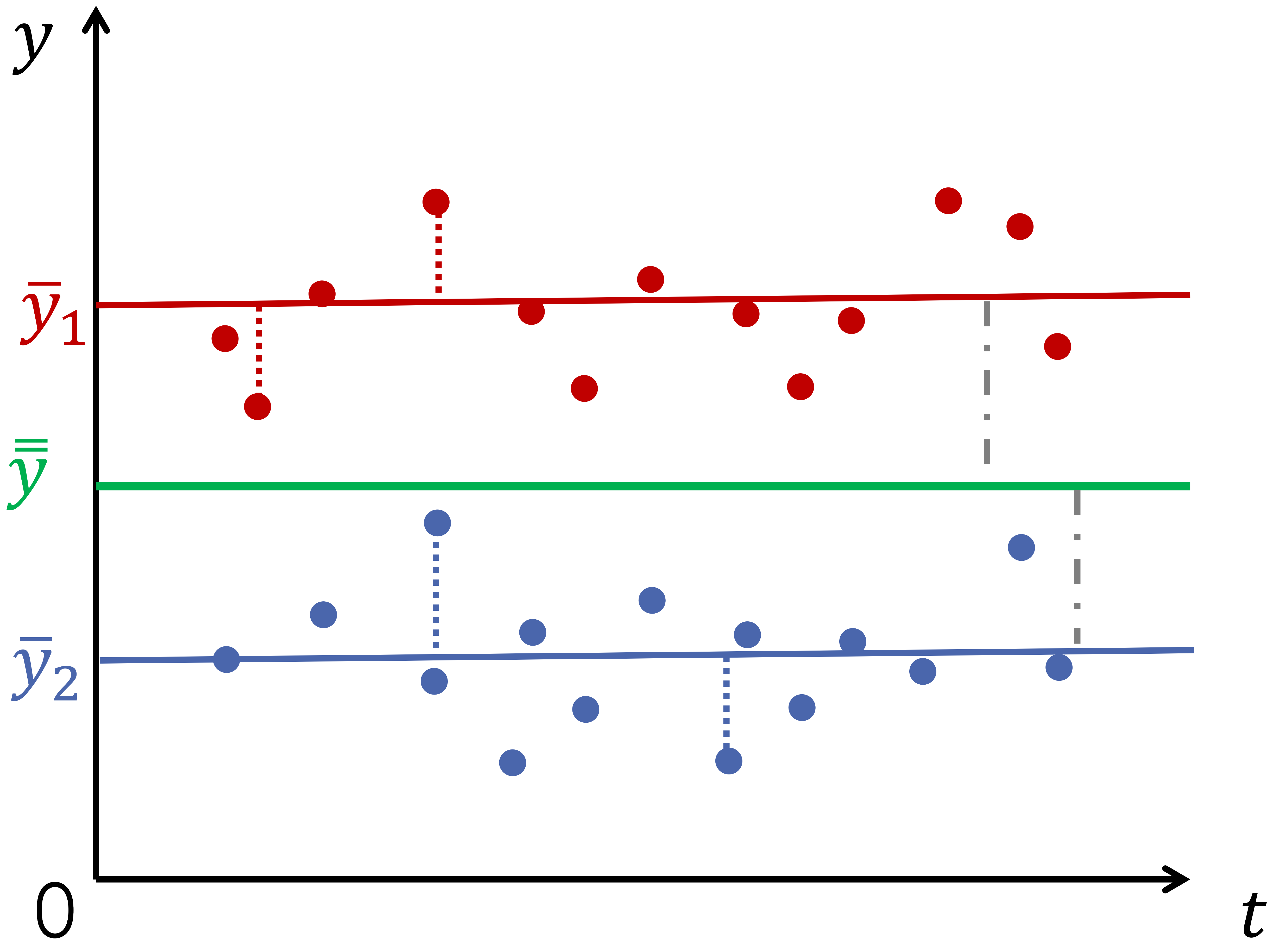
QUESTIONS OR COMMENTS?

EXERCISE 3: FIXED EFFECTS

But first, 15 minutes break

WITHIN & BETWEEN RELATIONSHIPS

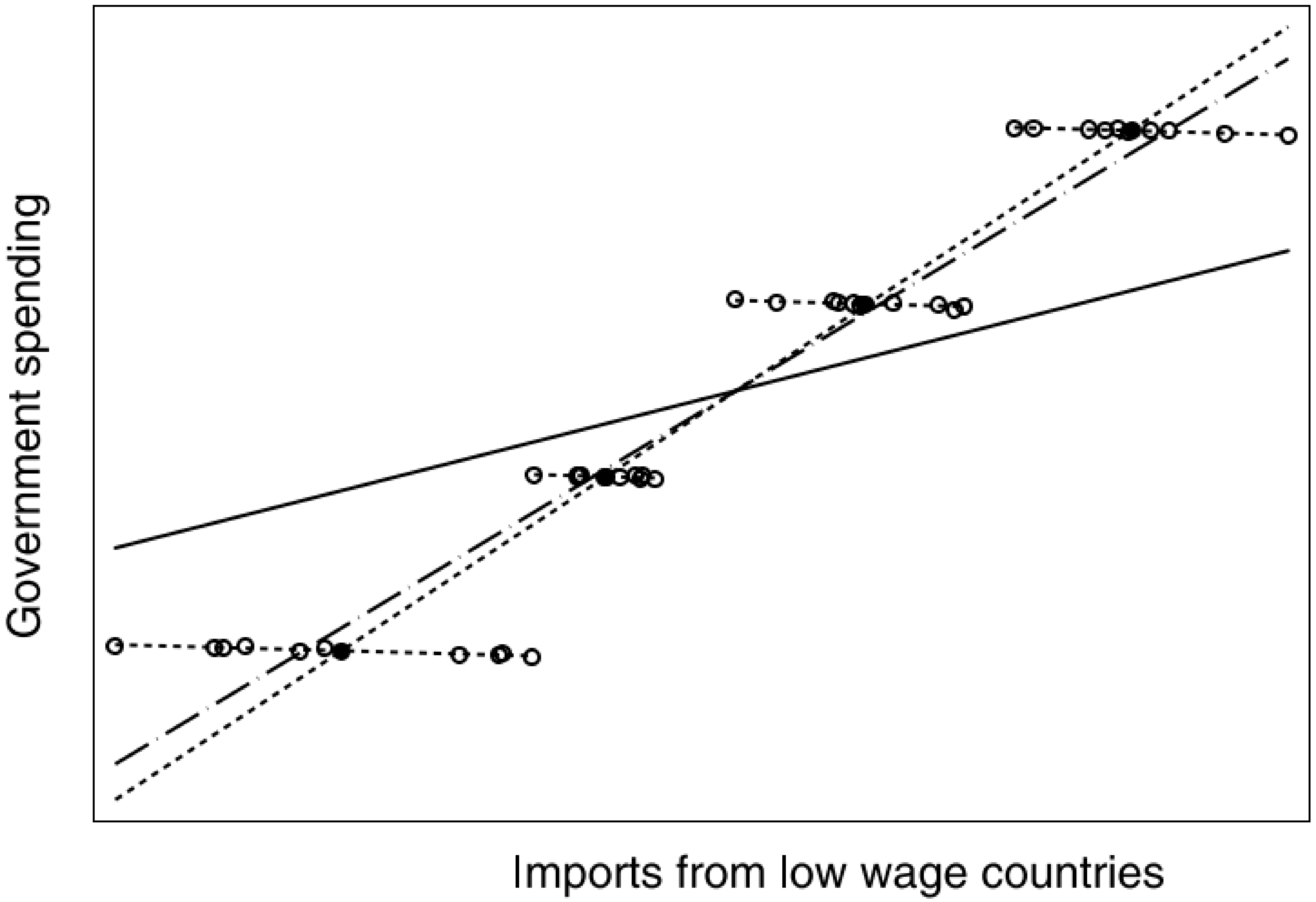
RECAP: WITHIN AND BETWEEN VARIANCE



Overall variance: $(y_{it} - \bar{\bar{y}})$	Within variance: $(y_{it} - \bar{y}_i)$	Between variance: $(\bar{y}_i - \bar{\bar{y}})$
0.16	0.08	0.08
0.46	0.38	0.08
-0.38	-0.46	0.08
0.24	0.2	0.04
-0.2	-0.24	0.04
0.08	0.04	0.04
-0.12	0	-0.12
-0.12	0	-0.12
-0.12	0	-0.12

EXAMPLE

- RQ: Government spending and import (simulated data)
- Countries importing more have higher spending (dashed lines)
- But *within* each country, government spending is associated with lower import rates (dotted lines)
- The "total" relationship is a mix of within and between (solid line)

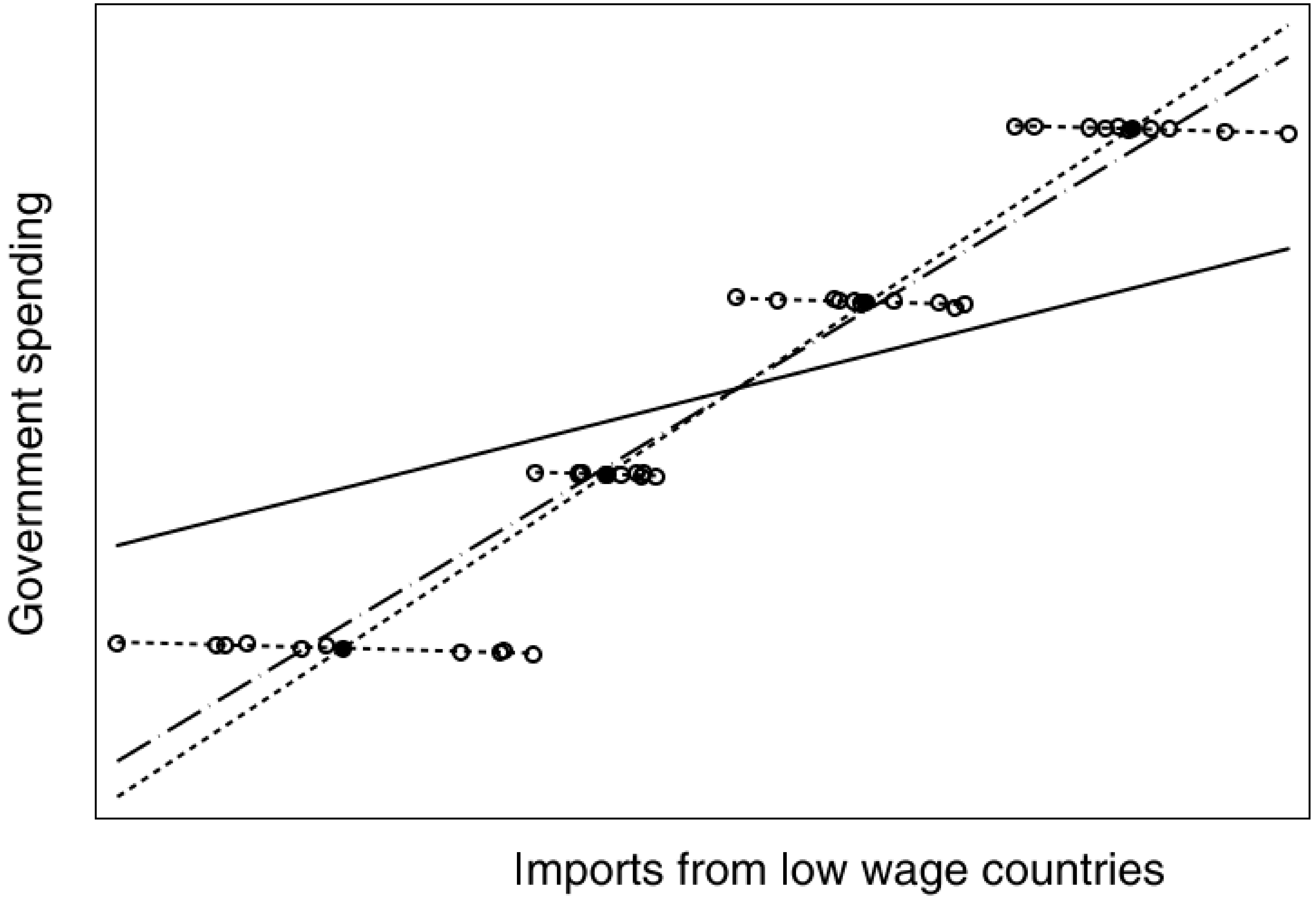


Source: efficiency2 data (see Example 4.3)

Andreß, Golsch & Schmidt (2014): 161

EXAMPLE

- Other country factors seem to distort the relationship *between* government spending and imports
- Unobserved heterogeneity on the country level
- Models only based on cross-sectional comparisons will lead to wrong conclusion that spending increases imports (*between*), when it actually seems to decrease it (*within*)



Source: `efficiency2` data (see Example 4.3)

Andreß, Golsch & Schmidt (2014): 161

RANDOM EFFECTS

WHAT IF UNOBSERVED HETEROGENEITY IS NOT PRESENT?

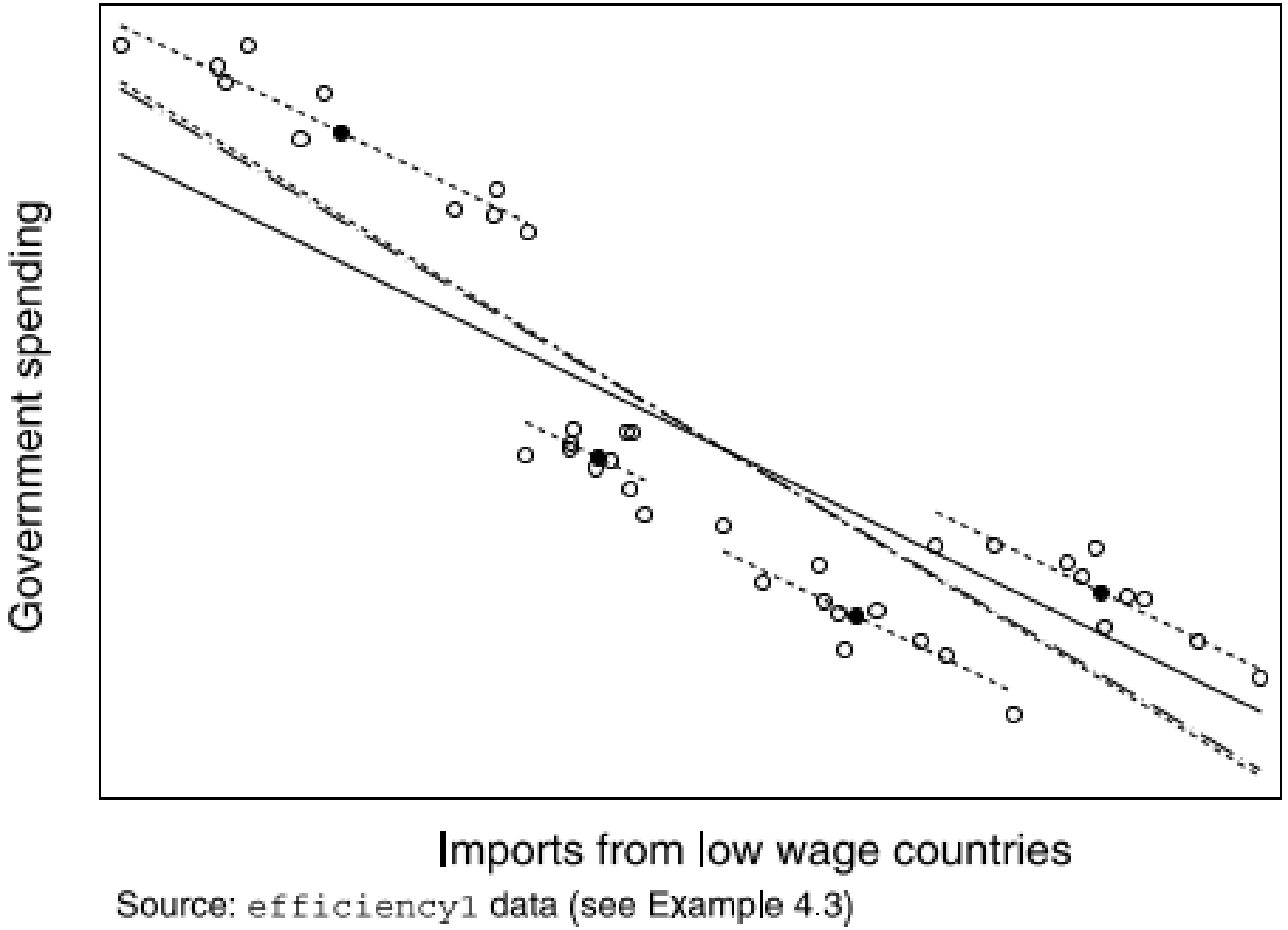
- FE automatically control u_i
- What if u_i does not correlate with x?
- What if there are no time-constant confounders?
- OLS?
  - Biased coefficients?
  - Biased standard errors?

EXAMPLE

- Same RQ as before but different data
- In this scenario, government spending is associated with lower import rates within each country (dotted lines)
- And the same holds also between countries (dashed lines)
- Countries importing more have less spending

→Slope of all lines rather similar

→Analysis of within-, between variance, or a combination of both all tell the same story



Andreß, Golsch & Schmidt (2014): 161

STATISTICAL (IN)DEPENDENCE OF PANEL DATA

- If u_i is not important why not use POLS?
- POLS assumes statistical independence of observations (every data points carries new information)
- With panel data, each unit contributes several data points (repeated measurements)

→Inflated sample size

→Underestimation of standard errors ("too significant" effects)

STATISTICAL (IN)DEPENDENCE OF PANEL DATA

- Information of one individual still (assumed) independent of information of other individuals
  - Between individuals
  - Just like cross-sectional data
- But: Information of one individual at time point t not independent from the information of this individual at $t-1$
  - Within individuals
  - Between time points

SERIAL CORRELATION

- Correlation of a variable with itself over time
- For example: unemployment rate in Germany 2016 probably not independent from its rate in 2015
- Put differently, the value of the unemployment rate in 2015 does not carry completely new information when you know the value of 2016
- Similarly, error term at t likely to correlate with error term at $t-1$
- Also called *autocorrelation* or *serial dependence*

SERIAL CORRELATION

- First order case: Pearson's correlation of y with $t-1$ lag of same variable

→Example: correlation of y with $L1.y$

→ $r = 0.78; n = 6$

ID	Year	y	$L1.y$
1	2009	0.04	-
1	2010	0.58	0.04
1	2011	0.88	0.58
2	2009	0.22	-
2	2010	0.51	0.22
2	2011	0.66	0.51
3	2009	0.08	-
3	2010	0.43	0.08
3	2011	0.92	0.43

SOLUTION TO SERIAL CORRELATION DUE TO U_i

- Eliminate between variance completely → demeaning / Fixed Effects-transformation → Fixed Effects
- Eliminate only share of between variance related to serial correlation → quasi-demeaning → Random Effects
- Both solve serial correlation due to u_i
- Neither solve serial correlation due to e_{it}

QUASI-DEMEANING

- Fixed Effects-Transformation (demeaning): $(y_{it} - \bar{y}_{i.}) = \beta(x_{it} - \bar{x}_{i.}) + (e_{it} - \bar{e}_{i.})$
- Completely eliminates time constant part
- Random effects transformation (quasi-demeaning): only subtract a part of the unit-specific mean

→Which part? That which produces serial correlation

→$(y_{it} - \theta\bar{y}_{i.}) = \beta(x_{it} - \theta\bar{x}_{i.}) + (e_{it} - \theta\bar{e}_{i.}) + \gamma(z_i - \theta z_i) + (u_i - \theta u_i)$

- θ: Demeaning parameter

ASSUMPTIONS

	POLS	RE	FE
Omitted Variable Bias			
Not in e_{it} (strict exogeneity): $cov(e_{it}, x) = 0$	✓	✓	✓
Not in u_i (RE assumption): $cov(u_i, x) = 0$	✓	✓	✗
Serial correlation			
Not in e_{it}: $cov(e_{it}, e_{is}) = 0$	✓	✓	✓
Not in ε_{it}: $corr(\varepsilon_{it}, \varepsilon_{is}) = var(u_i) = 0$	✓	✗	✗

BETWEEN EFFECTS

BETWEEN EFFECTS

- Variance of time-varying y_{it} can be decomposed in
  1. Within variance → $y_{it} - \bar{y}_i$
  2. Between variance → $\bar{y}_i - \bar{\bar{y}}$
- Fixed Effects eliminate all between variance (2.)
- The opposite: Eliminate all within variance (1.)
- … And estimate effects solely based on between variance (2.)

BETWEEN EFFECTS

- Cross-sectional analysis usually only use between unit variance
- How do you model between variation with panel data?
- Remember that time-stable differences between units are captured by unit-specific means ($\bar{y}_{i.}$, $\bar{x}_{i.}$)

→ $\bar{y}_{i.} = \beta_0 + \beta_1 \bar{x}_{1i.} + \cdots + \beta_k \bar{x}_{ki.} + \gamma_1 z_{1i} + \cdots + \gamma_l z_{li} + u_i$

- $\bar{e}_{i.} = u_i$

BE VS. RE VS. FE

- BE ≠ RE
- BE focus on *between* variation only
- FE focus on *within* variation only
- RE and POLS are a *mixture* of BE and FE
- BE > POLS > RE > FE; or BE < POLS < RE < FE (if FE ≠ BE)
- RE and POLS have lowest standard errors because they draw upon the most information (within and between)
- With longer panels, BE have highest standard errors because they draw upon least information (one observation per unit)

USING RE TO COMBINE FE AND BE

FE VS. RE

- FE are *unbiased* if model is correctly specified with respect to *time-varying* characteristics (time-constant aspects automatically controlled)
- FE estimates *less efficient* because between variance not used
- RE are *unbiased* if model is correctly specified with respect to *time-varying* <u>*and*</u> *time-constant* characteristics
- … But more efficient because they draw upon within and between variation
- First and foremost ensure that the model is correctly specified
- … then worry about standard errors

FE AND BE IN ONE MODEL

- RE may include x and z
- z variables yield BE
- RE of x are mixture of FE and BE
- Idea: Decompose the total effect of x into within part and between part

FE AND BE IN ONE MODEL

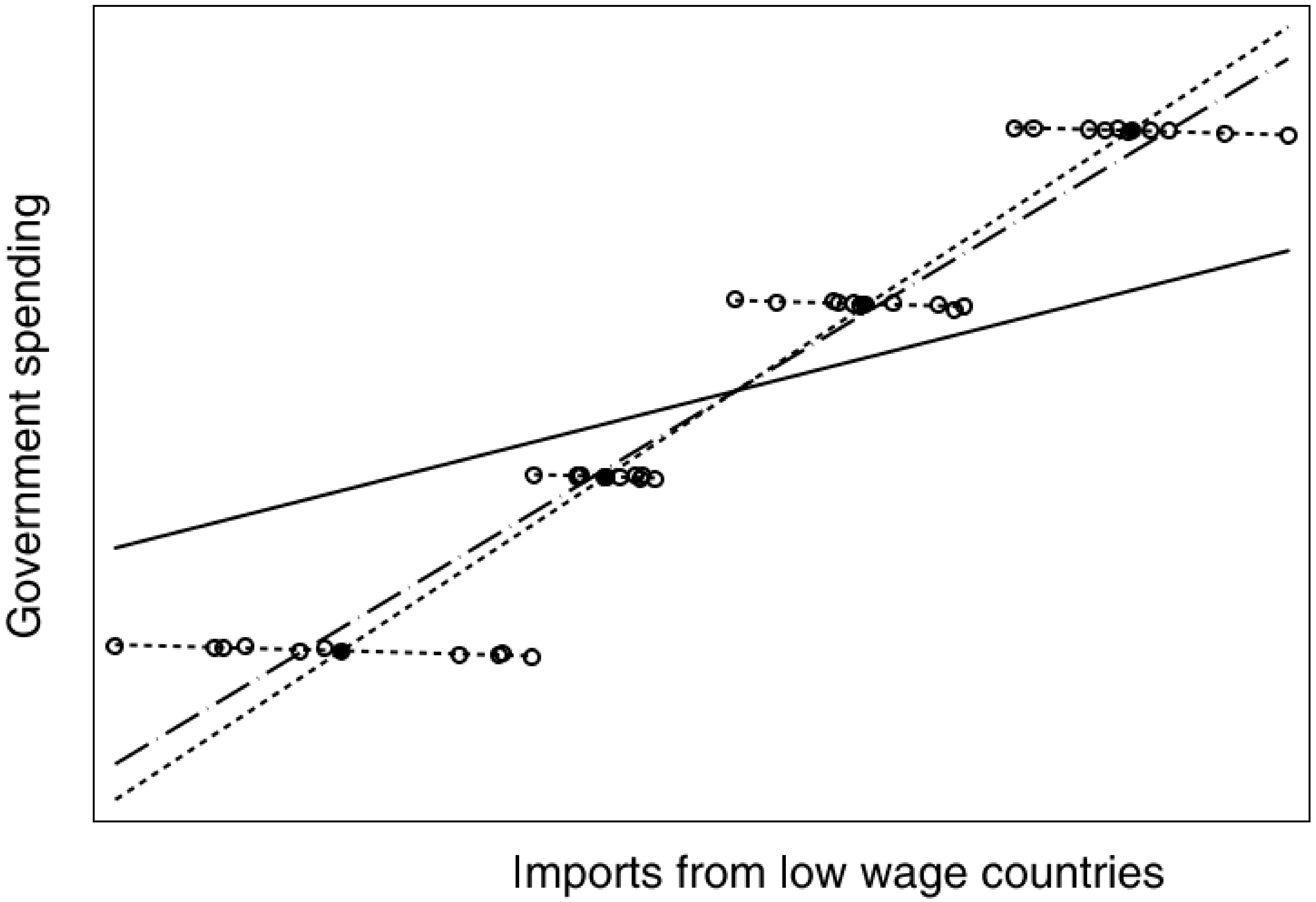
- BE for x variables are captured by their unit-specific means ($\bar{x}_i$)

→Include x and $\bar{x}_i$ to RE model (to be able to include time-constant variables)

- This will control between-unit differences
- The effects of original x variables are thus FE

EXAMPLE

- Simulated data
- Short dashed lines: FE = -0.19
- Solid line: RE = 1.74
- Long dashed line: POLS = 4.14
- Dotted line: BE = 4.53
- Circles: observations
- Black dots: Unit-specific means
- Government spending: $spend$
- Imports from low wage countries: $lowwage$



Source: `efficiency2` data (see Example 4.3)

Andreß, Golsch & Schmidt (2014): 161

INCLUDING $\bar{x}_i$ INTO RE MODEL

- $spend_{it} = 24.66 - 0.19 * lowwage_{it} + 4.72 * \overline{lowwage_{i.}}$
- Model replicates FE estimate (-0.19) for $lowwage$ when $\overline{lowwage}$ included
- $\overline{lowwage}$ nets out time stable differences in import between countries
- $\overline{lowwage}$ yields the *difference between the BE and FE* of $lowwage$ (4.72)
- *Hybrid model type 1* (Andreß et al. 2014)
- Or: *Mundlak model* (e. g.: Bell et al. 2019)

INCLUDING $\bar{x}_i$ AND $\ddot{x}_{it}$ INTO RE MODEL

- $\ddot{x}_{it} = x_{it} - \bar{x}_i$
- $spend_{it} = 24.66 - 0.19 * low\ddot{w}age_{it} + 4.53 * \overline{lowwage}_{i.}$
- $low\ddot{w}age_{it}$ is the demeaned variable
- Demeaned variables yield FE (-0.19) because, remember, this is the Fixed Effects-Transformation
- $\overline{lowwage}_{i.}$ now yields BE (4.53)
- *Hybrid model type 2* (Andreß et al. 2014)
- Or: *Random Effect Within-Between model (REWB)* (Bell et al. 2019)

INCLUDING $\bar{x}_i$ AND $\ddot{x}_{it}$ INTO RE MODEL

- $spend_{it} = 24.66 - 0.19 * \ddot{lowwage}_{it} + 4.53 * \overline{lowwage}_{i.}$
- $\bar{x}_i$ and $\ddot{x}_{it}$ are orthogonal (uncorrelated), so are $\ddot{x}_{it}$ and u_i
- Effects of $\ddot{x}_{it}$ thus not biased due to u_i
- Effects of $\bar{x}_i$ might be correlated with u_i and therefore biased (cross-sectional effects)

BENEFITS OF HYBRID MODELING

- FE effects for x and BE for z as well as for x all within one RE model
- Control for u_i but still estimate effects of time-constant variables
- Test for differences between BE and FE estimates

EXAMPLE: CZYMARA, C. S., & MITCHELL, J. (2023). ALL COPS ARE TRUSTED? HOW CONTEXT AND TIME SHAPE IMMIGRANTS' TRUST IN THE POLICE IN EUROPE. ETHNIC AND RACIAL STUDIES, 46(1), 72-96.

https://www.tandfonline.com/doi/full/10.1080/01419870.2022.2060711

All code available at: https://osf.io/bwxea/

RESEARCH QUESTION

- Does the size of the police force influence immigrants' trust in the police?
- Prior research reports higher levels of police per capita are associated with lower levels of trust in the police as more police does not seem to lower crime rates

→Cross-sectional variation of police size and trust *between* countries

- But increasing police between t_1 and t_2 may also change trust

→Longitudinal variation of police size and trust distress *within* countries

→But countries with more police are likely to differ from those with fewer police also in confounders (e.g., corruption)

- Are all confounders measured in the data / can we control them?
- If not: Problem of unobserved heterogeneity

WITHIN AND BETWEEN EFFECTS

- Cross-sectional differences between countries
  - Based on between variation
  - Between Effects (BE)
  - Likely to be plagued by unobserved heterogeneity
- Longitudinal differences within countries
  - Based on within variation
  - Fixed Effects (FE) / Within Effects (WE)
  - Automatically controlling unobserved heterogeneity
  - Countries are “their own controls”
  - Only possible with panel data

DATA

- 19,147 immigrants from 22 destination countries included in ESS waves 3 to 9
- Here, we ignore the individual-level structure and are just interested in the country differences
- Outcome: Trust in the Police (0 to 10)
- Predictor: Police size per 100,000 inhabitants (ranging from 147 in Finland to 624 in Cyprus)

MODEL

```
m3_all_mig <- lmer(trstplc ~ numPolPer100k_within + numPolPer100k_between
               + gdp_within + gdp_between
               + ruleoflaw_within + ruleoflaw_between
               + CrimePer100k_within + CrimePer100k_between
               + immigration_per100k_within + immigration_per100k_between
               + discrim + minority
               + ruleoflaw_diff + livecnty_comb1 + educ
               + hincfel + gndr + agea + age2
               + (1 | cntry/essround),
               data = data_combined_migr_50)
```

RESULTS

Table 4. Police force size and immigrants' trust in the police.

Variables	Police size *Estimates*	Country controls *Estimates*	Individual controls *Estimates*	Interaction *Estimates*
Police size (WE)	0.0009 (0.0014)	0.0007 (0.0012)	0.0006 (0.0012)	−0.0093 (0.0087)
Police size (BE)	−0.0048*** (0.0010)	−0.0032* (0.0016)	−0.0028 (0.0016)	−0.0036* (0.0015)
Time since migration			−0.2743*** (0.0195)	−0.1719 (0.1179)
Police size (WE) × Time since migration				0.0006 (0.0007)
Police size (BE) × Time since migration				0.0025 (0.0028)
Police size (WE) × Time since migration (BE)				−0.0004 (0.0004)
Country-level controls		✓	✓	✓
Individual-level controls			✓	✓
Random Effects				
σ^2	5.61	5.61	5.43	5.40
τ_{00}	0.15 $_{\text{essround:cntry}}$ 0.26 $_{\text{cntry}}$	0.08 $_{\text{essround:cntry}}$ 0.12 $_{\text{cntry}}$	0.08 $_{\text{essround:cntry}}$ 0.12 $_{\text{cntry}}$	0.12 $_{\text{essround:cntry}}$ 0.08 $_{\text{cntry}}$
τ_{11}				0.01 $_{\text{essround:cntry.livecnty\_comb1}}$ 0.02 $_{\text{cntry.livecnty\_comb1}}$
ρ_{01}				−0.71 $_{\text{essround:cntry}}$ −0.58 $_{\text{cntry}}$
ICC	0.07	0.03	0.04	0.04
N	7 $_{\text{essround}}$ 22 $_{\text{cntry}}$	7 $_{\text{essround}}$ 22 $_{\text{cntry}}$	7 $_{\text{essround}}$ 22 $_{\text{cntry}}$	7 $_{\text{essround}}$ 22 $_{\text{cntry}}$
N (countries)	22	22	22	22
N (country-waves)	110	110	110	110
N (respondents)	19,147	19,147	19,147	19,147
AIC	87.623.494	87.663.511	87.106.530	87.123.478

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

RESULTS

- Significant raw between "effect", but no within effect
- Raw BE plagued by omitted variable bias
- Controlling for country variables and compositional differences renders BE insignificant
- However, also little variation in police force size over time

TESTING DIFFERENCES BETWEEN WE AND BE

DIFFERENCES IN WE AND BE

- RE are unbiased if BE and WE are essentially the same
- Mundlak model automatically tests for differences between WE and BE (effects of $\bar{x}_i$)
- REWB
  - Effects of $\bar{x}_i$ =BE
  - Effects of $\ddot{x}_{it}$ =WE
  - Test whether $BE = WE$, or $BE{-}WE = 0$
- Both types of tests numerically equivalent

TESTING FOR DIFFERENCES IN BE AND WE

- Test whether $BE - WE = 0$
- Either for single parameters
  - Mundlak: Test if coefficient of $\bar{x}_i = 0$ (automatically done in regression output)
  - REWB: Test whether difference in coefficients of $\ddot{x}_{it}$ and $\bar{x}_i$ is zero
- … or the model in total
  - Mundlak: Test if coefficient of *all* $\bar{x}_i = 0$
  - REWB: Test whether differences in coefficients of *all* $\ddot{x}_{it}$and $\bar{x}_i$ are zero
- Overall tests and Hausman test are asymptotically equivalent

DIFFERENCES OF BE AND FE

- If BE and WE do not differ, there are no time-constant confounders in the BE model
- If BE and WE differ, BE are plagued by unobserved heterogeneity
- However, BE might still be interesting, as they might be proxies for "a range of unmeasured social processes, which might include those omitted variables themselves" (Bell et al. 2019: 1059 f.)
- E. g.: "Effect" of ethnicity (time-constant)
  - Not direct causal effect of particular genes
  - Rather, effects of unmeasured social and cultural factors that are related to ethnicity
  - BE can help understanding patterns in the world, but needs theoretical knowledge

DIFFERENCES OF BE AND FE

- There is some relevant but unmeasured time-constant characteristic that is not in the model
- However, can be seen as something of substantive interest
- Opens opportunity for theoretical speculation: *What* is it that might be different between units that is relevant? How did prior (cross-sectional) research deal with this issue?
- Depending on time span of panel: BE may indicate "historical" differences which are not captured by rather short-term over time variation

SUMMING UP

SUMMARY

- Causal claims can more easily be defended with effects based purely on within variation → FE
- In some cases, differences between FE and RE are only marginal
  - Long panel: Relative share of within variation tends to increase
  - “Sluggish” data
  - Good controls: If all confounders are measured, all relevant differences between units can be statistically controlled, no need to turn to within variance

SUMMARY

- Hybrid models combine virtues of FE and RE models
- WE and BE all in one model
- Estimates of z
- Allows to test for differences between WE and BE
  - Are there differences between the two?
  - Test whether these differences are statistically significant (t-test in linear case)
  - Quantify how large the difference is

A CAUTIONARY TAIL

- WE are not *necessary* for causal inference
  - See randomized control trials
- WE are not *sufficient* for causal inference
  - Unobserved time-varying confounders are important
  - … and most things vary over time
- But: WE are very helpful for causal inference
  - No time-constant or slow confounders
  - Better compare apples and apples (WE) than apples and oranges (BE)
  - WE might be "less wrong" than BE
- But it all depends on your theoretical model and its assumptions!
- See Rohrer & Murayama (2023). These are not the effects you are looking for: causality and the within-/between-persons distinction in longitudinal data analysis. Advances in methods and practices in psychological science 6 (1).

QUESTIONS OR COMMENTS?

EXERCISE 4: REWB MODELS

But first, 15 minutes break

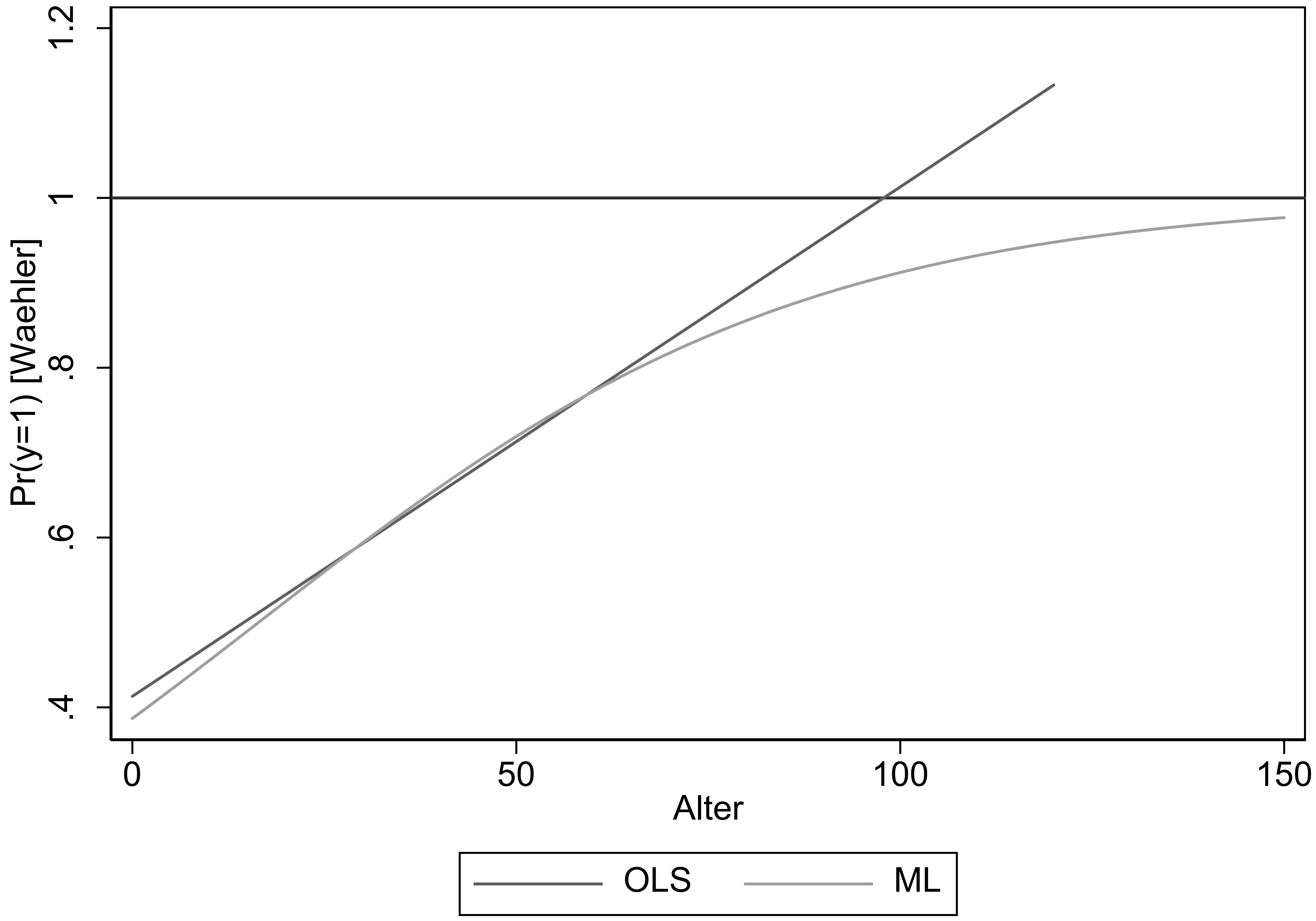
PROBABILITY MODELS

BINARY OUTCOMES

- So far, we treated outcomes as continuous
- Sometimes this assumption is at stake
- Binary y?
- OLS → linear probability model (LPM)
  → `lm(y ~ x, data = data)`
- Maximum Likelihood (ML) → logistic regression
  → `glm(y ~ x, data = data, family="binomial")`
- Both part of base R

	LPM	Logistic Regression
Estimator	Ordinary Least Squares (OLS)	Maximum Likelihood (ML)
Optimization function	Minimize sum of squared residuals	Maximize likelihood function
Solution	Analytical (calculated)	Numerical (iterative "trying")
Effects of explanatory variables	<u>Linear</u>: effect of x is the same for all values/at all levels of x <u>Additive</u>: effect of x is independent from other x	<u>Non-linear</u>: effect of x depends on its value/level <u>Multiplicative</u>: effect of x depends on the values of other x
Predicted values	Can be below 0 or above 1	$[0,1]$
Interpretation of coefficients	Percentage point change in probability of $y = 1$	Absolute change in logged odds (logits) of $y = 1$ *or* relative change in odds ratios of $y = 1$

LPM VS LOGISTIC REGRESSION



LOGISTIC POOLED MODEL

LINEAR VS LOGISTIC POOLED

	Linear	Logistic
Properties	Unbiased (if assumptions met)	Consistent (if assumptions met)
Estimator	OLS	ML
u_i	Ignored	Ignored
e_{it}	Distribution assumed $\sim N(0, \sigma_e^2)$	Distribution assumed $\sim logistic(0, \frac{\pi^2}{3})$

- The same as linear & logistic regression for cross-sectional data → idea of pooled models is to ignore the panel structure

LOGISTIC FIXED EFFECTS

FIXED EFFECTS

- In linear case: eliminate everything time-constant (observed z and unobserved u_i)
  - Time-demeaning (Fixed Effects transformation): remove unit-specific means
  - Dummy for each unit
- Based on within-unit variation over time
- FE ≙ conditional estimates (OLS: given unit-specific means)

LOGISTIC REGRESSION FOR PANEL DATA

- $\ln\left(\frac{\text{pr}(y = 1|x)}{1-\text{pr}(y = 1|x)}\right) = u_i + \beta_1 x_1 + \cdots + \beta_k x_k$
- u_i: unobserved heterogeneity
- Time-constant variables z part of u_i
- Problem: u_i not easily eliminated from likelihood function
- Unit dummies no alternative (*incidental parameters problem*: ML consistent, based on $n \to \infty$)

CONDITIONAL MAXIMUM LIKELIHOOD

- In linear case: FE based on transforming data
- In logistic case: also conditional estimation

→Maximize probability of observing a specific sequence conditional on certain number of 1s

- Control frequency of 1s means to control for average level of y
- Sequences with no change in y do not contribute to conditional likelihood ($pr = 1$)

→Conditional Maximum Likelihood (CML) only based on within variance, similar to linear FE

EXAMPLE: UNION MEMBERSHIP

- Andreß et al. (2014): page 231
- $t = 2$
- Union membership = dummy (1: yes, 0: no)
- Possible sequences
  - No member in both years: 00 ($n = 386$)
  - Became member: 01 ($n = 22$)
  - Left union: 10 ($n = 37$)
  - Member in both years: 11 ($n = 100$)

EXAMPLE: UNION MEMBERSHIP

Member at $t = 1$	Member at $t = 2$		Total
	0	1	
0	386	22	408
1	37	100	137
Total	423	122	545

Sequences	Number of 1s	Frequency	Conditional probability	#outcomes/ #sample
00	0	386	$\Pr(00 \mid 1s = 0)$	$386/386 = 1$
11	2	100	$\Pr(11 \mid 1s = 2)$	$100/100 = 1$
01	1	22	$\Pr(01 \mid 1s = 1)$	$22/(22 + 37) = 0.37$
10	1	37	$\Pr(10 \mid 1s = 1)$	$37/(22 + 37) = 0.63$

EXAMPLE

- $Pr(01 \,|\, 1s = 1)$

$$\frac{pr(y_{i1} = 0)\, and\, pr(y_{i2} = 1)}{\big(pr(y_{i1} = 0)\ and\, pr(y_{i2} = 1)\big)\ or\ \big(pr(y_{i1} = 1)\ and\ \ pr(y_{i2} = 0)\big)}$$

$$= \frac{e^{u_i + \beta_1 x_1}}{1 + e^{u_i + \beta_1 x_1}}$$

- I spare you the math (see Andreß et al. (2014): 229 ff.)
- For $t = 2$ CML can use traditional ML
  - Outcome: sequence 01 vs. sequence 10
  - Explanatory variables: First differences of x
- When $t > 2$ traditional ML cannot be used

EXAMPLE: LANCEE, B., & SARRASIN, O. (2015). EDUCATED PREFERENCES OR SELECTION EFFECTS? A LONGITUDINAL ANALYSIS OF THE IMPACT OF EDUCATIONAL ATTAINMENT ON ATTITUDES TOWARDS IMMIGRANTS. EUROPEAN SOCIOLOGICAL REVIEW, 31(4), 490-501.

https://academic.oup.com/esr/article/31/4/490/496810

RESEARCH QUESTION

- More educated people are more pro-immigration, but why?
- Liberalization effect of education (promotion of tolerant values)
- More tolerant people are more likely to be successful in school (self-selection)

DATA

- Swiss Household Panel (SHP, 1999–2011)
- 4,339 individuals and 16,571 person-years
- Outcome: "Are you in favour of Switzerland offering foreigners the same opportunities as those offered to Swiss citizens, or in favour of Switzerland offering Swiss citizens better opportunities?" (1: in favour of better opportunities for Swiss citizens; 0: otherwise)
- Predictor: Highest level of education
- REWB model

RESULTS

Table 2. Logistic hybrid model predicting the favouring Swiss citizens, odds ratios

	OR	se
Within effects		
Educational attainment		
Primary	ref.	
Secondary without Matura	0.872	(0.322)
Secondary with Matura	0.808	(0.121)
Secondary vocational	0.942	(0.117)
Tertiary vocational	0.614*	(0.138)
University	0.610	(0.170)
Employed	1.050	(0.084)
Leaving parental home	1.381**	(0.161)
Life satisfaction	0.979	(0.030)
Satisfaction with financial situation	0.959*	(0.019)
Political interest	1.016	(0.018)
Age	0.962*	(0.016)
Percentage of immigrants	0.974	(0.021)
Unemployment rate	1.216***	(0.050)
Between effects		
Educational attainment		
Primary	ref.	
Secondary without Matura	0.684	(0.362)
Secondary with Matura	0.167***	(0.040)
Secondary vocational	1.352	(0.284)
Tertiary vocational	0.375**	(0.112)
University	0.083***	(0.029)

RESULTS

- Again, significant BE but no WE
- Authors' conclusion: "*differences between educational groups are mostly due to selection effects, and not to the alleged liberalizing effect of education*"

SUMMING UP

SUMMARY

- Logistic Fixed Effects estimated with CML
- Conditional probability of certain sequence given all sequences that include the same number of 1s
- When $t = 2$ similar to First Difference estimation for continuous y
  - $y = pr(01\ vs\ 10)$
  - $x =$ first differences

LINEAR VS LOGISTIC FIXED EFFECTS

	Linear	Logistic
Conditional estimation	… given unit-specific means	… given certain number of 1s
Units without change	Time-demeaning (Fixed Effects Transformation) → Units without change of y do not contribute to within variance	Sequences without change → Do not contribute to conditional likelihood function

LINEAR VS LOGISTIC FE

	Linear	Logistic
Properties	Unbiased (if (fewer) assumptions met)	Consistent (if (fewer) assumptions met)
Estimator	OLS (demeaned) or LSDV	Conditional ML
u_i	Estimated (LSDV) or controlled (through demeaning)	Controlled
e_{it}	Distribution assumed $\sim N(0, \sigma_e^2)$	Distribution assumed $\sim logistic(0, \frac{\pi^2}{3})$

QUESTIONS OR COMMENTS?

15 minutes break

LOGISTIC RANDOM EFFECTS

ASSUMPTIONS

	Pooled	RE	FE
Omitted Variable Bias			
Not in e_{it} (strict exogeneity): $cov(e_{it}, x) = 0$	✓	✓	✓
Not in u_i (RE assumption): $cov(u_i, x) = 0$	✓	✓	✗
Serial correlation			
Not in e_{it}: $cov(e_{it}, e_{is}) = 0$	✓	✓	✓
Not in ε_{it}: $corr(\varepsilon_{it}, \varepsilon_{is}) = var(u_i) = 0$	✓	✗	✗

RANDOM EFFECTS

- As in linear case: RE need extra assumption compared to FE
- Issue 1: u_i independent of x *and* z variables (model correctly specified)
- Issue 2: Potential serial correlation (biased standard errors)

SOLUTION TO MISSPECIFICATION DUE TO z IN LOGISTIC RE MODELS

- None (like linear case)
- If this is an issue, use FE

SOLUTION TO SERIAL CORRELATION DUE TO u_i IN LOGISTIC RE MODELS

- Make assumption about distribution of u_i
- E. g.: normally distributed, $u_i \sim N(0, \sigma_u^2)$
- Estimate serial correlation: $corr(\varepsilon_{it}, \varepsilon_{is}) = \rho$ (rho)
- ρ: Fraction of total error variance due to u_i (ICC)
- σ_e^2 in logistic models fixed to $\frac{\pi^2}{3}$

→ $\rho = \frac{\sigma_u^2}{\sigma_u^2 + \frac{\pi^2}{3}}$

TESTING RE ASSUMPTIONS

	Assumption	Linear	Logistic
FE vs RE	Omitted variable bias of z?	Hausman Test $H_0\colon FE = RE$	Hausman Test $H_0\colon FE = RE$
RE vs pooled	Serial correlation?	Breusch-Pagan Test $H_0\colon \sigma_u^2 = 0$	Likelihood Ratio Test $H_0\colon \rho = 0$

- Test for omitted variable bias follows same logic in both cases
- Test for serial correlation
  - Breusch Pagan Test
  - Likelihood Ratio Test: Compare RE with pooled model (pooled model nested in RE with restriction $\sigma_u^2 = 0$)

EXAMPLE: LANCEE, B., & SARRASIN, O. (2015). EDUCATED PREFERENCES OR SELECTION EFFECTS? A LONGITUDINAL ANALYSIS OF THE IMPACT OF EDUCATIONAL ATTAINMENT ON ATTITUDES TOWARDS IMMIGRANTS. EUROPEAN SOCIOLOGICAL REVIEW, 31(4), 490-501.

https://journals.sagepub.com/doi/abs/10.1111/imre.12015

RESEARCH QUESTION

- Does economic competition cause anti-immigration attitudes?
- Being economically vulnerable or competing for jobs might lead to more exclusionary immigration attitudes

DATA

- German Socio-Economic Panel Survey(GSOEP, 1999–2008)
- 15,694 individuals and 77,691 person-years
- Outcome: Concerns about immigration (1: very; 0: somewhat or not concerned)
- Predictor: Various SES variables
- Logistic RE and FE models

RESULTS (PART)

	M1 Random effects	M2 Fixed effects
Female	0.825*** (0.041)	
Age	1.005* (0.002)	
Marital status		
Married	Ref.	Ref.
Single	0.883 (0.089)	0.891 (0.101)
Divorced/separated/widowed	1.076 (0.068)	1.097 (0.128)
Educational attainment		
Inadequately/general elementary	1.541*** (0.181)	0.865 (0.234)
Basic vocational	1.309*** (0.076)	0.974 (0.176)
Intermediate vocational/general	Ref.	Ref.
General/vocational maturity	0.392*** (0.035)	0.946 (0.139)
Tertiary education	0.219*** (0.016)	0.763 (0.138)
Year and federal state dummies	Yes	Yes
Unemployment rate		
Proportion foreigners		
Constant	4.658*** (0.142)	
Log-likelihood	−37285.3	−15531.6
N observations	77,691	40,359
N subjects	15,694	6,181

RESULTS

- Most effects of class not statistically significant with FE
- Something time-constant seems to confound the effect of class on concerns
- Effect on unemployment stable

→ Unemployment matters for everyone

LINEAR PROBABILITY MODEL FIXED EFFECTS

LINEAR PROBABILITY MODEL

- Use linear FE for dummy outcome variable to get linear probability model (as in cross-sectional case)
- Circumvents many of the issues related to the use of non-linear models
  - Easy interpretation of effects
  - Meaningful (changes of) variance components
  - Comparability of coefficients between models
  - Fast estimation
- But also same problems as in linear case
  - Potentially predicting probabilities below 0 or above 1
  - Functional form misspecification (effects do not depend on the level of x)

EXAMPLE: CZYMARA, C. S., & DOCHOW, S. (2018). MASS MEDIA AND CONCERNS ABOUT IMMIGRATION IN GERMANY IN THE 21ST CENTURY: INDIVIDUAL-LEVEL EVIDENCE OVER 15 YEARS. EUROPEAN SOCIOLOGICAL REVIEW, 34(4), 381-401.

https://academic.oup.com/esr/article/34/4/381/5047112

All code available at: https://osf.io/y254k/

RESEARCH QUESTION

- Does media reporting on immigration in the 21 days before an interview increase one‘s concerns about immigration?
- Idea: Mass media as a source of threat perceptions

DATA

- German Socio-Economic Panel Survey(GSOEP, 1 2001 to 2015)
- 25,073 individuals and 190,049 person-years
- Outcome: Concerns about immigration (1: very; 0: somewhat or not concerned)
- Predictor: Media salience
- LPM FE model

RESULTS

	(1) Main model
Media salience, past 21 days	0.050****
	(0.001)
Party preference (ref.: no preference)	
CDU/CSU (Christian Democrats)	0.027****
	(0.004)
SPD (Social Democrats)	−0.007*
	(0.004)
Die Grünen (The Greens)	−0.012*
	(0.007)
Die Linke (The Left)	−0.005
	(0.008)
FDP (Free Democrats)	0.019**
	(0.009)
Others and mixed	0.015
	(0.010)
Radical right	0.144****
	(0.013)

- Interpretation: Like linear FE but change in probability to be concerned
- Effect of *Media salience* ($\beta = 0.05$)
  - Positive relationship: "*A one unit increase in media salience predicts an increase in the probability of being very concerned by five percentage points.*"

SUMMING UP

UNOBSERVED HETEROGENEITY

- FE: Fixed value
  - Linear model: Can be estimated (as dummies)
  - Logistic model: Cannot be estimated, but controlled
- RE: Random variable
  - Single value not interesting, but variation
  - ML (linear or logistic): Needs assumption about distribution of u_i
  - GLS (linear): Quasi-demeaning, no assumption about distribution of u_i neccessary
  - Assumption: u_i uncorrelated with x *and* z

		Idiosyncratic error: e_{it}	Unobserved heterogeneity: u_i	Extra assumptions
Linear models	Pooled	Assumed $\sim N(0, \sigma_u^2)$	-	• $Cov(u_i, z_i) = 0$ / $Cov(u_i, x_{it}) = 0$ • $Cov(e_{it}, e_{is}) = 0$
	RE	Variance estimated (Feasible Generalized Least Squares: FGLS)	Variance estimated (FGLS)	$Cov(u_i, z_i) = 0/$ $Cov(u_i, x_{it}) = 0$
	FE	Assumed $\sim N(0, \sigma_u^2)$	Estimated (LSDV) or controlled (FE)	-
Logistic models	Pooled	Assumed $\sim logistic(0, \frac{\pi^2}{3})$	-	• $Cov(u_i, z_i) = 0$ / $Cov(u_i, x_{it}) = 0$ • $Cov(e_{it}, e_{is}) = 0$
	RE		Assumed $\sim N(0, \sigma_u^2)$	• $Cov(u_i, z_i) = 0$ / $Cov(u_i, x_{it}) = 0$
	FE		Controlled	-

LINEAR AND LOGISTIC PANEL REGRESSION

	Linear	Logistic
Standard errors: pooled < RE < FE	✓	✓
Estimates: RE always between FE and BE	✓	✗
Hybrid models replicate FE and BE exactly	✓ (FE and BE uncorrelated by design)	✗ (only similar due to non-linearity)

QUESTIONS OR COMMENTS?

Thanks for your attention!

LITERATURE

- Andreß, Golsch & Schmidt (2014). Applied panel data analysis for economic and social surveys. Chapter 4.1 (119 - 174) and Chapter 5.1. Springer Science & Business Media.
- Bell, Fairbrother & Jones (2019). Fixed and random effects models: making an informed choice. Quality & Quantity 53 (2). 1051 - 1074.