UNIVERSITY OF MANNHEIM

GESS – Graduate School of Economic and Social Sciences

DSPI

UNIVERSITY OF OXFORD

Analysing Panel Data in Comparative Research

Kun Lee

DPhil Candidate in Social Policy

University of Oxford

Visiting PhD Student, CDSS, University of Mannheim

I will discuss…

- Concepts & use of panel/TSCS data
- Fundamental issues with pooled TSCS data
- Fixed/random effect models
- Hybrid models
- Use of dynamic specifications
- Panel GMM estimation
- Error Correction Models

- Focus on assumptions, interpretation, pros & cons

What is panel/TSCS data?

- I say panel, you say TSCS…
- “Repeated observation of same cross-section units” vs “pooled time-series of multiple units”
- Different terminology depending on long/wide structure & discipline
  - Large N, short T vs. small N, long T
  - Micro vs **macro** (country, state, region…) data
  - Microeconomics vs political science/CPE
- Different modelling strategies, c.f. Law of large numbers

Examples of macro-panel/TSCS analysis

- Democracy and economic growth (Helliwell, 1994; Acemoglu et al., 2008)
- Political/Institutional determinants of welfare state development (Huber & Stephens, 1993) & Wage inequality (Rueda & Pontusson, 2000; Iversen & Wren, 1998)
- Union power and Economic Performance (Hicks, 1994; Boreham & Compton, 1992)
- Trade with China and labour regulation (Adolph et al., 2017)
- Active labour market policy and (un)employment (Benda et al., 2020)

Reasons to use macro-panel/TSCS

- Interest on the role of macro-level institutions/structures
- Larger samples size N*T (=statistical power, more variables)
- Observing dynamics: effects change over time & space
- We know there are omitted variables almost always in our models (we cannot observe everything!)
- If omitted variables are significantly correlated with our main explanatory variable, our estimates will be biased
- Using panel data helps control some of the important omitted variables

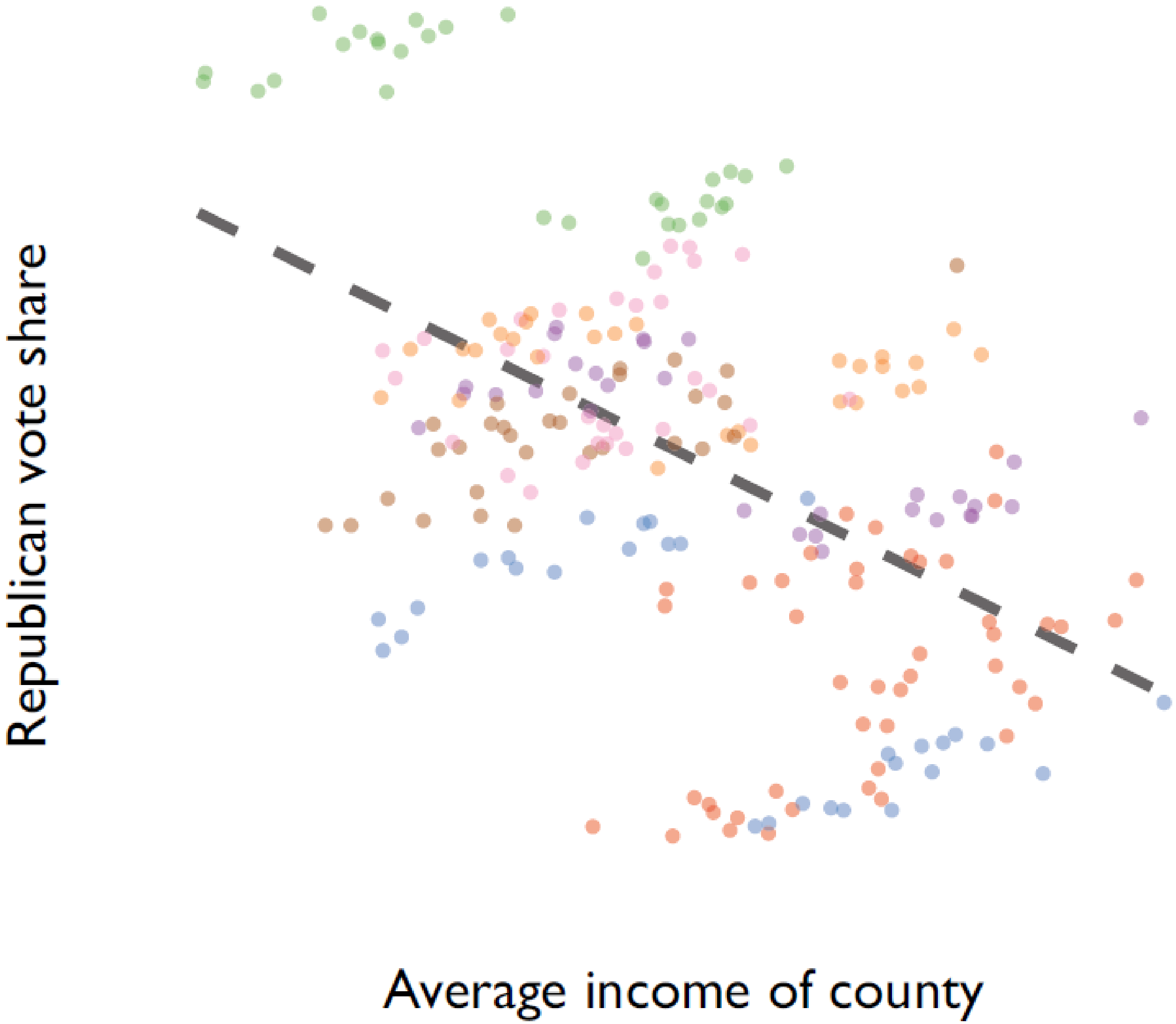
Fundamental Issues with TSCS data (1)

- Constant-coefficient (Kittel, 1999): averaging effects across different countries & time points
  - Between-country effect: "Sweden's unemployment rate is lower than Germany because of higher ALMP spending"
  - Within-country effect: "Germany's unemployment rate decreased because of an increase in ALMP spending"
  - Two different substantive meanings are merged into one coefficient in the standard regression analysis
- Using constant coefficient implicitly assumes that the effects of X on Y is uniform across countries & over time
  - Between & within effects may change over time & space
  - You can model heterogeneous effects flexibly, but then parameters may be too many relative to sample size

Fundamental Issues with TSCS data (2)

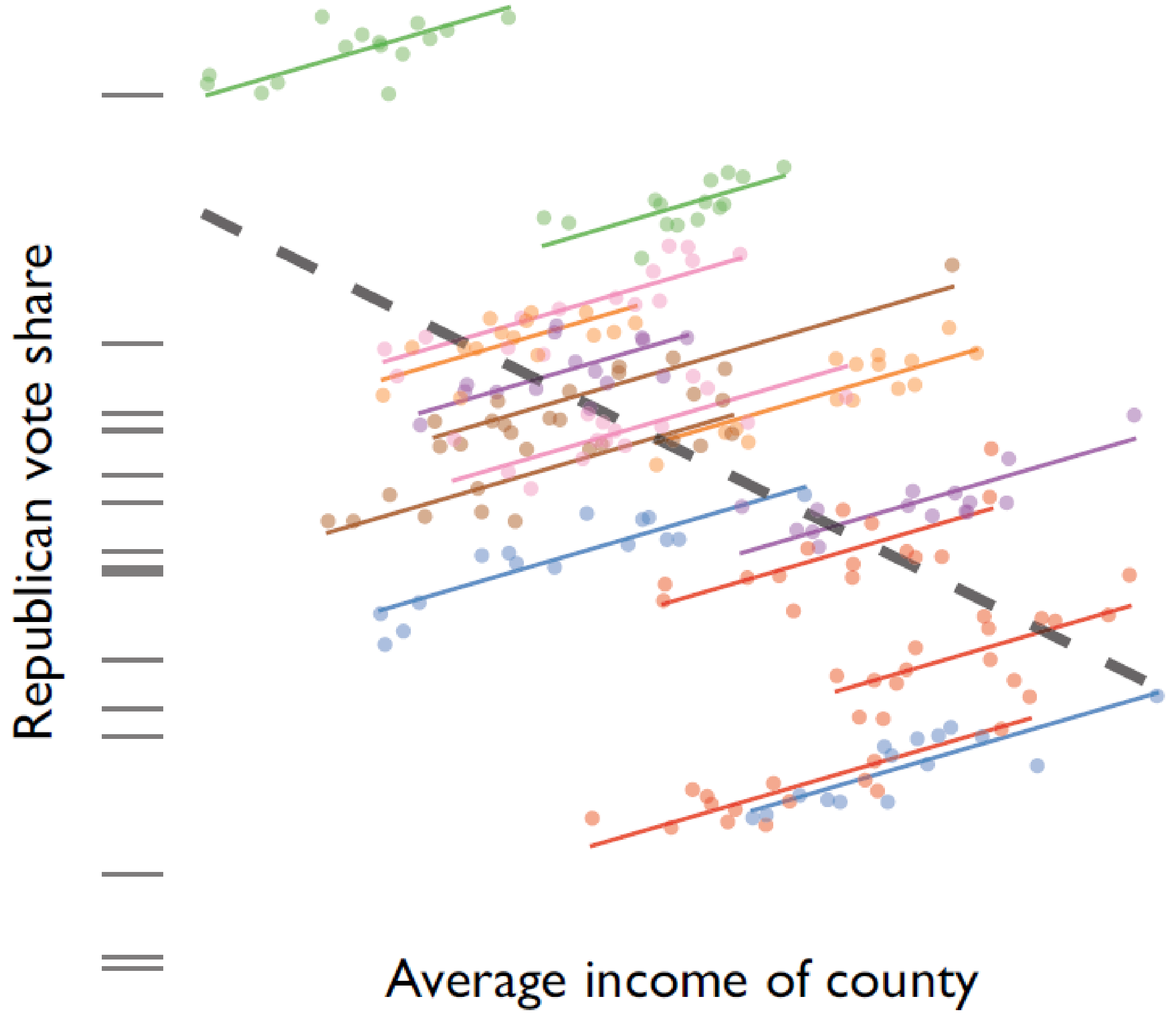
- TSCS data by nature accompanies selection problems
  - “We use data from 18 advanced welfare states…”
  - Some dishonest climate studies
- Classic assumptions of OLS (iid)
  - Heteroskedasticity: Unemployment rate fluctuations in Germany & Sweden
  - Contemporaneous correlation: Policy development in Germany & Austria; UK & US
  - Serial correlation: Unemployment rate in the UK in 2022 & 2021

Problems with Naïve OLS



Source: Adolph (2021)

Problems with Naïve OLS



Source: Adolph (2021)

Fixed Effects Models

- Workhorse model in panel data analysis across social sciences
- Least-square dummy variables: include N-1 country-specific dummy variables (or N dummies without intercept)
- Within-groups estimation: subtract group-specific means from all obs.
- Stata: xtreg Y X1 X2…Xk, fe
- R: plm(Y ~ X1+X2..+Xk, data=dataframe, model="within")
- Removes omitted variables bias from unobserved 'time-invariant' variables: geographical characteristics, entrenched culture
  - Reasonable in TSCS analysis, as many country-specific characteristics are time-constant
- Adding time dummies: "two-way" fixed effects
  - Controls unobserved effects that varies over time but uniform across countries
  - Economic crisis, pandemic effects (not observed as variables)

Fixed Effects Models: Issues (1)

- You are only using within-country variations ΔX → ΔY: averaged across countries
  - If you are interested in between-country institutional effects, FE is not suitable!!
- No substantive meaning in the dummy coefficient: "Germany is different because it's the Germany"
- Cannot include any substantive time-invariant variables
  - You are dumping all time-invariant characteristics into one dummy variable → controlling OVB has its cost
  - Ex) Welfare regime types, labour regulations, political institutions
  - R will return errors; Stata will automatically omit the variable
  - But they can still be included as an interaction term (not independently)

Fixed Effects Models: Issues (2)

- Assuming that omitted variables (unobserved characteristics) are time-invariant → is that true?
  - Ex) Effects of family policy on female employment: are gender norms time-constant?
  - Reverse causality not allowed (simultaneity or feedback effect)
- Not so useful when your main explanatory variables have little variations over time (e.g. effects of democracy on growth)
- Huge loss of DoF if you have short T (e.g. N=30, T=4)
- You are not just loosing DoF: FE models discard the information of "levels", only use changes within countries (e.g. unemployment in Italy and the US)

Random Effect Models

- RE models use both within- and between-country effects
  - Does not discard the ‘level’ information
  - More efficient than FE models (= smaller variance, less uncertainty)
- Useful when FE is too costly (e.g. N=30, T=4)
- Including time-invariant explanatory variable (“level-2”) is also possible
- Treats unobserved country-specific characteristics (u_i) as random errors (not fixed) that has a distribution
- Estimated through generalised least square or maximum likelihood approach
- Stata: xtreg Y X1 X2…Xk, re
- R: plm(Y ~ X1+X2..+Xk, data=dataframe, model=“random”)
- Estimates are in between naive OLS and FE models

Random Effects Models: Issues

- Strong assumption: unobserved characteristics are uncorrelated with explanatory variables – $cov(X_{it}, U_i) = 0$
  - Almost always not true: estimates are mostly biased
  - Culture, longstanding institutions usually shape policy/politics
- The problem here is whether the bias is substantially large
- Hausmann Test: check whether FE and RE coefficients are similar
- Better in prediction than causal inference
  - When you can allow a bit of bias but want to include time-invariant variables (gender, race…), get better predictive power, more efficient estimates
  - Economics (FE) vs sociology, political science (RE/FE)

Still the constant-coefficient issue hasn't been solved…

Hybrid Model

- Combining within- & between-unit (country) effects within a random effects framework (Allison, 2009; Schunck, 2013, 2017)
- $y_{it} = \beta_0 + \beta_1(X_{it} - \bar{X}_i) + \beta_2\bar{X}_i + \beta_3 C_i + u_i + \varepsilon_{it}$
- β_1 (within-country effect) is identical to the counterpart in fixed-effect models (unbiased if no time-varying omitted variables)
- Country-specific means (levels) become one of level-2 (time-invariant) variables in multilevel modelling → between-country effect
- You can also include other time-invariant variables (C_i)
- If WE=BE, the model is identical to random-effects (intercept) model

Haapanala et al. (2022) “Decent Wage Floors in Europe: Does the Minimum Wage Directive Get it Right?”

Table 1. Results from random effects within-between (REWB) regression models.

	DV: share of workers on <60% median wage			DV: effective wage floor (P5 in PPS)		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Within: CBC	0.021	0.015	0.015	0.002	0.000	0.000
	(0.017)	(0.017)	(0.017)	(0.005)	(0.005)	(0.005)
Within: SMW dummy	3.250	3.009	3.010	-2.310***	-1.958**	-2.104**
	(3.040)	(3.109)	(3.109)	(0.887)	(0.916)	(0.915)
Within: SMW PPS	-0.462	-0.462	-0.462	0.318***	0.270***	0.285***
	(0.321)	(0.330)	(0.330)	(0.094)	(0.097)	(0.097)
Between: CBC	-0.123***	-0.128***	-0.129**	0.011	0.005	0.035***
	(0.030)	(0.029)	(0.063)	(0.009)	(0.007)	(0.013)
Between: SMW dummy	-5.599**	-11.027***	-11.031***	-5.302***	-1.767**	-1.525**
	(2.792)	(3.670)	(3.688)	(0.866)	(0.873)	(0.764)
Between: CBC*SMW dummy			0.001			-0.039***
			(0.069)			(0.014)
Between: SMW PPS	0.518	0.955*	0.953*	0.977***	0.637***	0.711***
	(0.385)	(0.515)	(0.532)	(0.118)	(0.122)	(0.111)
Constant	16.195***	20.086***	20.094***	8.970***	6.415***	5.990***
	(2.093)	(2.695)	(2.789)	(0.656)	(0.643)	(0.579)
Year FE	X	X	X	X	X	X
Controls		X	X		X	X

Hybrid Model: Issues

- But hybrid model is still bound to the RE assumption
- Unobserved country characteristics should not be correlated with $\bar{X}_i$ to get unbiased between-effect (or other time-invariant effect)

→ Otherwise between-effects would be biased

- Small-sample bias: another reason why the between-country effect is often unreliable (Bryan & Jenkins, 2016)
  - You need 25-30 countries for a simple model with single level-2 variable
- Using smaller N (<20) and for more complex modelling, Bayesian modelling performs better (Stegmueller, 2013; Elff et al. 2021)

But I still want to do some causal inference…

Use of Dynamic Specifications

- Use of lagged dependent variables (LDV) as an explanatory variable: serial correlation is included in the model and 'explains' the dependent variable ($Y_{it} = \alpha * Y_{i,t-1} + \beta * X_{it} + u_i + e_{it}$)
  - Ex) Current employment rate is explained by previous employment rate
- OLS with LDV: bias of β very small, as much of the endogeneity (time-varying & invariant characteristics) is correlated with LDV
  - α may be upward biased, standard errors are wrong & inefficient
  - Panel-corrected standard error (PCSE: Beck & Katz, 1995): allows contemporaneous correlation & heteroskedasticity across countries
  - Adding FE (country-specific dummies) may help reduce remaining bias
  - No need to have large N! (even unbiased with N = 1)

Dynamic Specifications: Issues

- But using LDV with FE introduces another type of bias (Nickell, 1981)
  - LDV is necessarily correlated with country-specific characteristics (u_i)
  - Creates downward bias on LDV coefficient (thus affecting X coefficient)
  - Bias wanes by the size of T → not so much concern when T is very long
- The substantive meaning of the coefficient changes: "Given the previous level of Y, one-unit change X increases/decreases Y by the size of β…"
  - Coefficient indicates a short-term, year-by-year response rather than a full effect

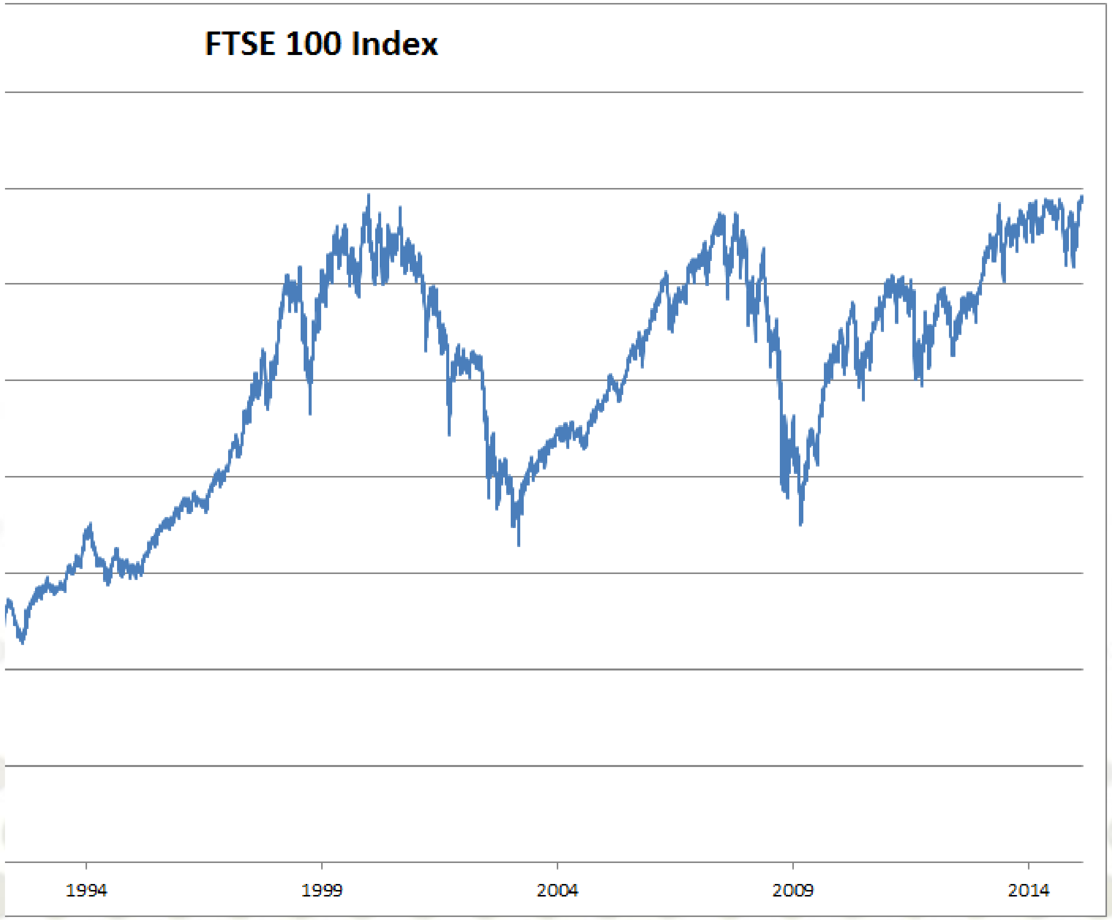
  →Short-term impulse of X on Y (there are delayed effects!)

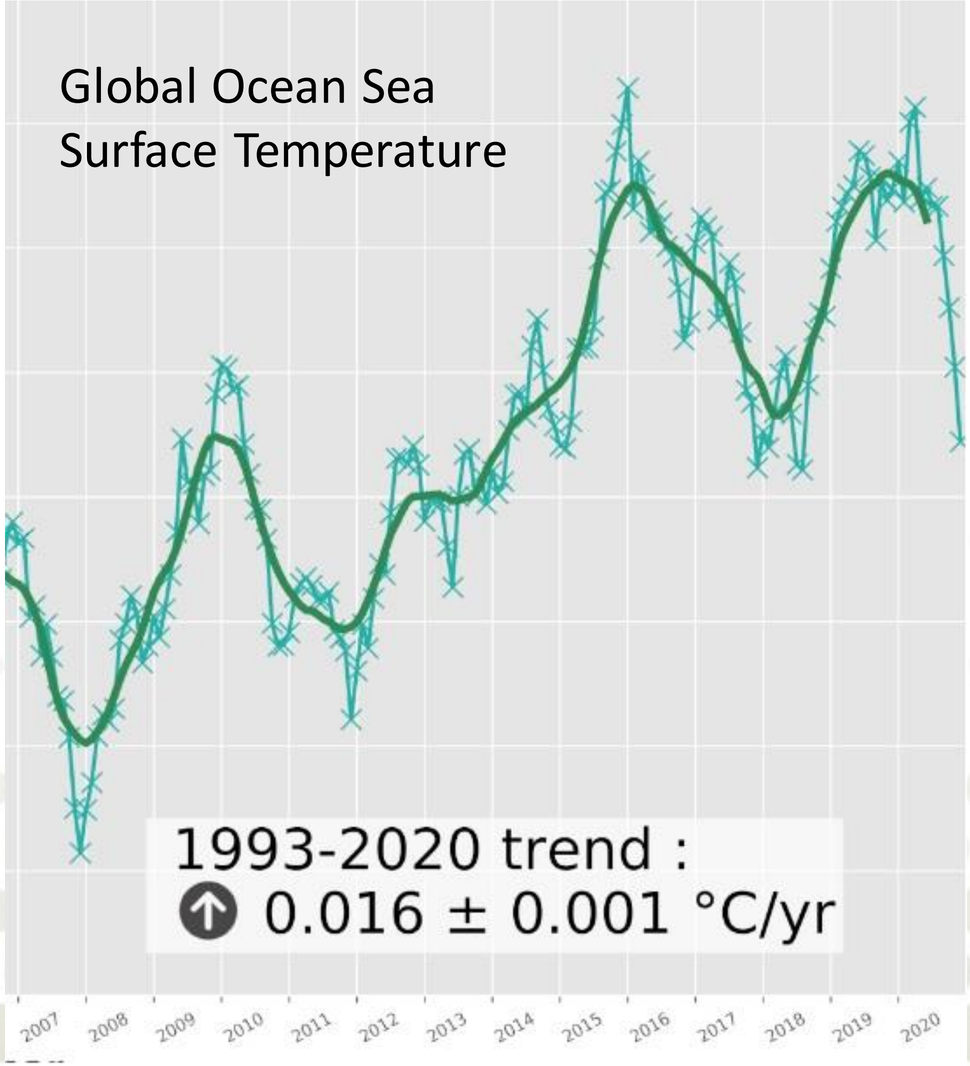
  * Interpretation of long-run, full effect: $\beta/(1-\alpha)$ (* α: Coefficient of LDV)
- Most of the DV's variation may be captured by LDV
  - Now, effects of X can be way underestimated, even appear non-significant

Panel GMM Estimation

- But if T is not too long (=<25): Nickell bias may be a concern
  - Bias is equal to 20% when T=30 (Judson & Owen 1999; Roodman 2009)
- Way to combine fixed effects with LDV
- $\Delta Y_{it} = Y_{it}$ - $Y_{i,t-1}$ is not correlated with fixed effects
- Difference GMM (Arellano & Bond 1991)
  - Use $Y_{i,t-2}$ as an instrument for $\Delta Y_{i,t-1}$ in first-differenced model
- System GMM (Arellano & Bover 1995; Blundell & Bond 1998)
  - Use $\Delta Y_{i,t-2}$ as an instrumental variable for $Y_{i,t-1}$
- But this method is designed for large N, small T setting → very sensitive to small-sample bias (Roodman, 2009)

Spurious Correlations in Time-Series Data





Error Correction Models

- Most commonly used is Engle & Granger (1987) two-step method
- $\Delta y_{it} = \beta_1 \Delta X_{it} + \beta_2 \left(y_{i,t-1} - \beta_3 X_{i,t-1} \right) + \varepsilon_{it}$

  where $\Delta X_{it} = X_{it} - X_{i,t-1}$
- This can be rewritten as,
- $y_{it} = \alpha\, y_{i,t-1} + \beta_1 \Delta X_{it} + \beta_2 X_{i,t-1} + \varepsilon_{it}$
- Include levels & differences in a single dynamic model
  - Avoids non-stationarity & estimates short/long-term effects
- Core assumption: X and Y has a long-term equilibrium relationship (→ residuals from a level model follow a stationary process)
- Similar issues as in “OLS with LDV” models

Remaining issues

- Cross-sectional dependence: “Are policy changes in Germany and France independent?”
  - Use of time dummies, panel-corrected standard error (Beck & Katz 1995)
  - Use of spatial models: modelling diffusion process
- Heterogeneous effects across time & space
  - Hypothetically, effect of X in the first half of the period is -0.5, and the second half is +0.5 → zero pooled effect
  - Same in cross-country aggregation: ex) Effects of ALMP may differ in Sweden and the UK
- Use of interaction effects: but you cannot include interactions with all countries or all time points! (use theory to group countries & times)
- Matching methods for causal inference in time-series cross-section analysis (Imai et al., 2021)

Summary & Conclusion

- Fixed-effects models give you within-country effects (often unbiased) but at (sometimes large) costs
- Random-effects models give you efficient but often biased estimates
- Hybrid models can separately estimate within- and between-country effects but this is still a random-effects model
- Using LDVs may provide some causal evidence but the interpretation becomes different
- Panel GMM allows combining LDV with FE but is sensitive to biases in small N, long T setting
- Spurious correlations should also be considered (feat. error correction model)

So many different approaches, but which model should I choose?

"All models are wrong, but some are useful“

“Since all models are wrong the scientist must be alert to what is importantly wrong”

George Box (1976), "Science and statistics", *Journal of the American Statistical Association*, 71 (356): 791–799.

Every method has its assumptions, pros & cons.

No one-size-fits-all solution

Need to choose carefully which one would be the “least worrying” model, given your data structure, assumptions of the models & research questions.

Thank you!
Q & A