21st Century Political Methodology: *Advances in All Modes of Empirical Analysis in Political Science*

Robert J. Franzese, Jr. (franzese@umich.edu)
Professor & Associate Chair, Department of Political Science,
Director, Program in International & Comparative Studies,
Research Professor, Center for Political Studies,
The University of Michigan, Ann Arbor

Fellow & 15th (former) President, *The Society for Political Methodology*

Prepared for Presentation to *The Asian Political Methodology Conference*, 5 January 2019

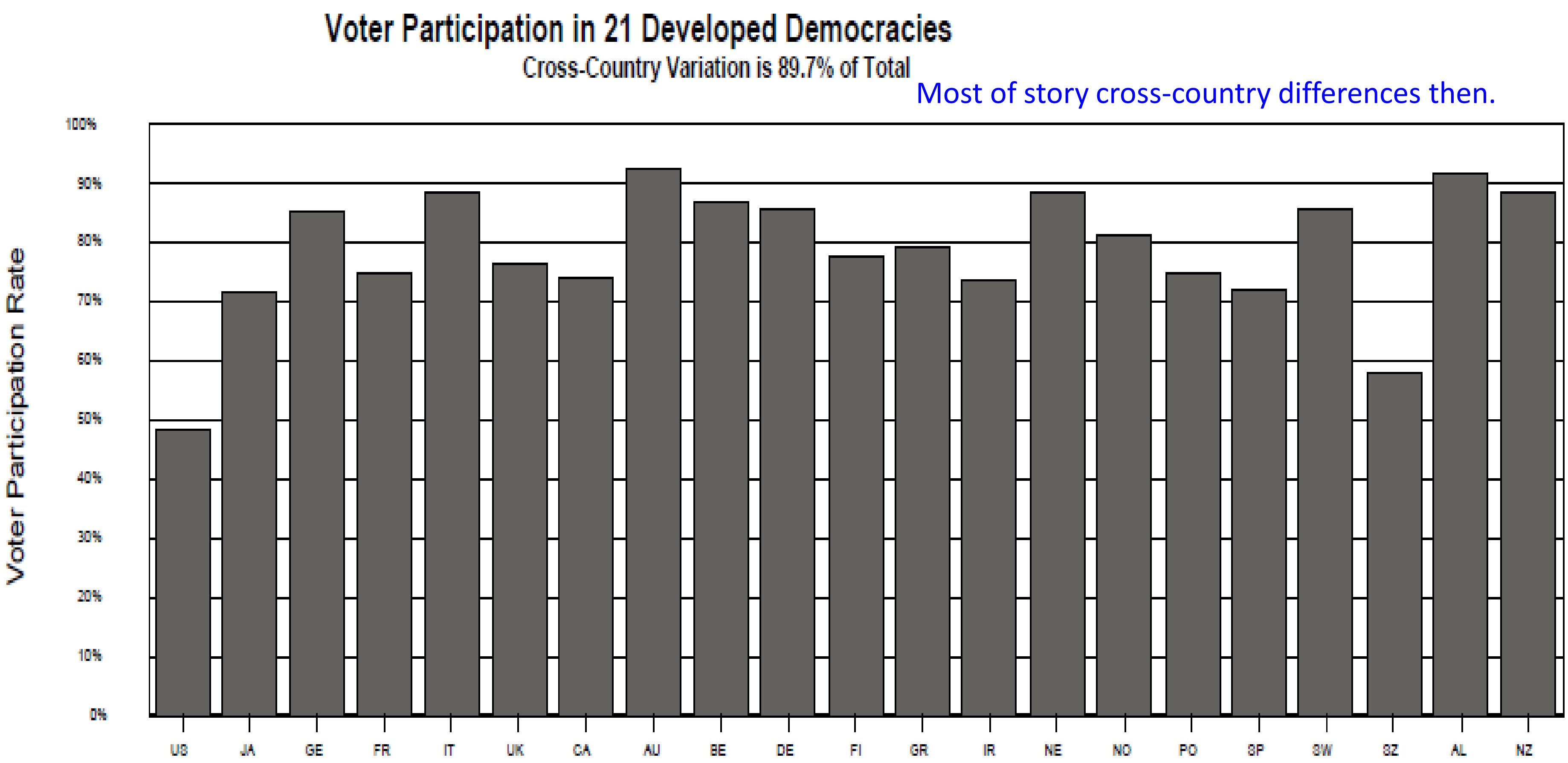
Some Comments on 21st Century Political Methodology

- An Exciting Time in Political Methodology:
  - Rapidly Advancing EMPIRICAL METHODOLOGIES for
  - Increasingly Sophisticated THEORIES, with
  - Seemingly Unbounded Potential DATA Richness.
- Modes of Empirical Analysis in Political Science:
  - Testing of Causal Theory
    - Ideal & Gold Standard=Experimental RCT
      Optimal to gauge evidence for ***existence*** of causal effect
  - Description & Measurement, Classification, & Forecast/Prediction
    - Ideal=Consistency & Accuracy, Performance relative to Expert;
      Gold Standard=**Out-of-Sample (Forecast) Error**
  - Empirical-Model & Causal-Response Estimation
    - Ideal=Empirical Model is Useful Empirical Simplification;
      Gold Standard=**Out-of-Sample (Causal-Response) Error**
- Kinds of Empirical Questions:
  - Factual: e.g., what % of population supports incumbent? (physical=statistical population)
  - Theoretical: e.g., what explains support incumbent? (stat.pop.=hypothetical, infinite)

Exciting Times: Increasingly Sophisticated Theories

- **Paradigm:** a shared set of assumptions & accepted theories in a scientific field.
  - Once a theory has become established as part of scientific knowledge in a field of study, researchers can build upon foundation that theory provides.
    - Scholars who study evolution of scientific fields of research lively & ongoing debate about where social sciences, political science, are in development.
    - The more-skeptical argue Political Science not sufficiently mature to have paradigm…
- A quick look at some of most developed & substantiated:
  - **Voter Participation:** know lot re: what sorts people vote & why voter-participation rates higher in some democracies & elections than in others
  - **Economic Voting:** know incumbents presiding over stronger economic times tend to do better in elections than incumbents presiding over weaker
  - **Electoral Cycles:** know incumbents ∴ incentives to try deliver voters stronger economic performance & other material benefits around election times and ∴ that policies & outcomes tend to exhibit electoral periodicity

- **Voter Participation:** know lot re: what sorts people vote & why voter-participation rates higher in some dem's & elect's than others

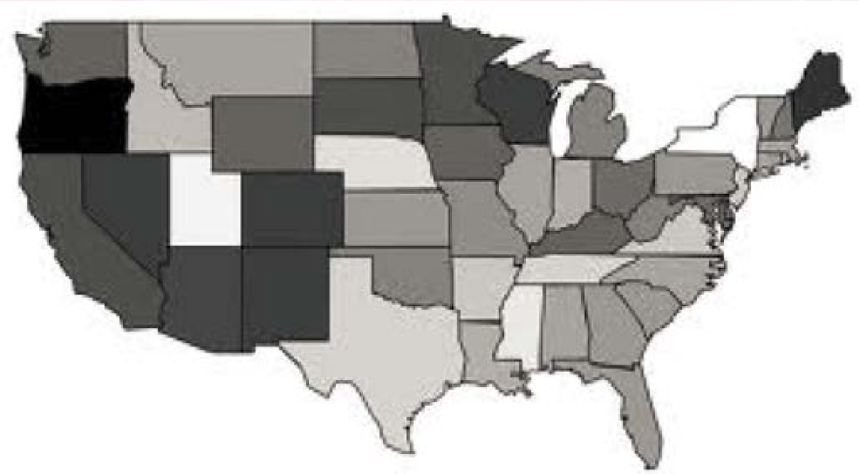


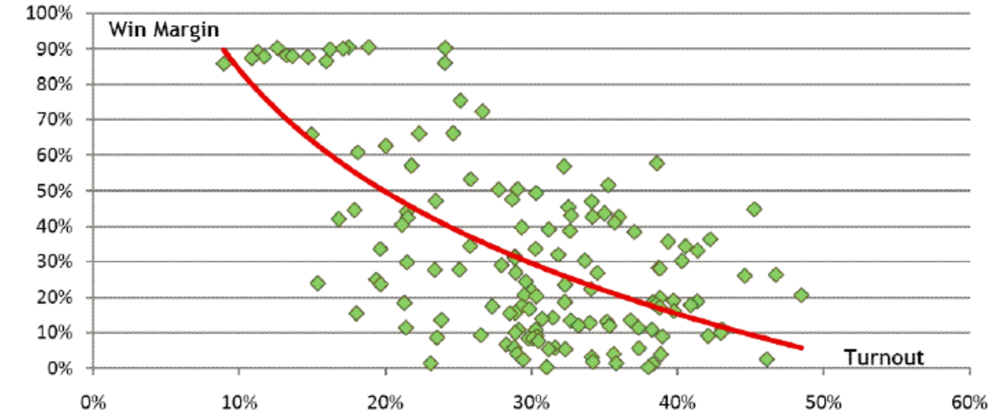
Most of story cross-country differences then.

$$\Pr(Vote) = f\left(pr\{pivotal\}\times\left[X_p - X_a\right] + B - C\right)$$

Voter Participation:
Who & How Many Vote?

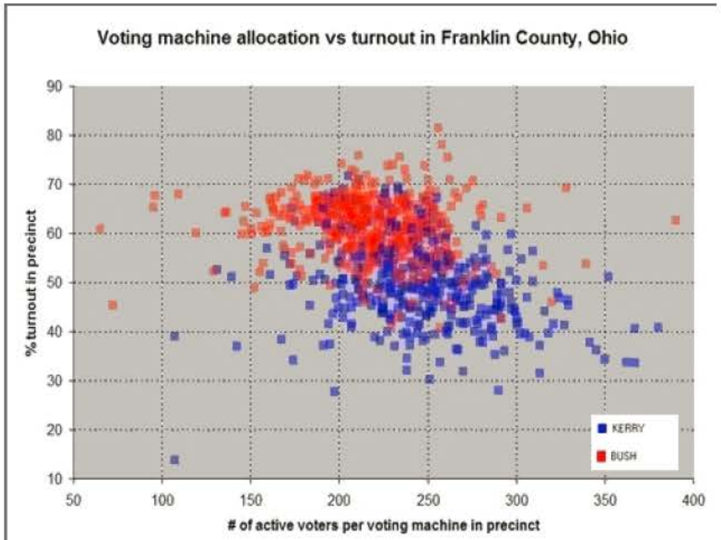
$$\Pr(Vote) = f\left(pr\{pivotal\}\times\left[X_p - X_a\right] + B - C\right)$$

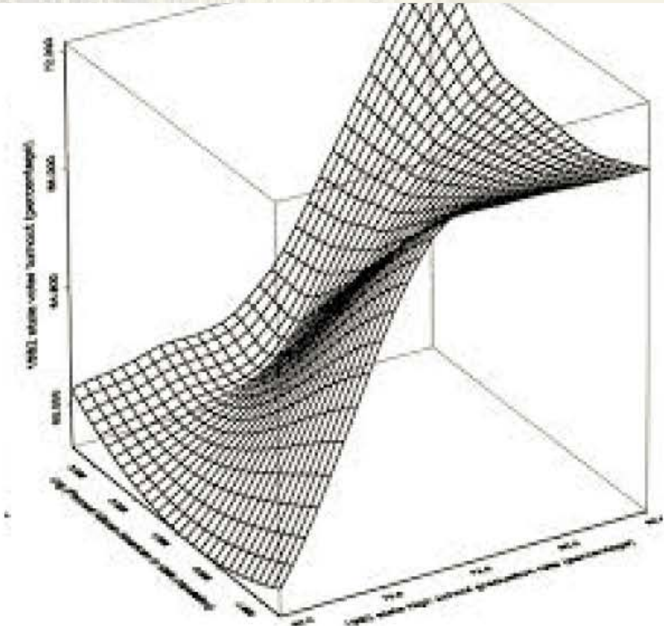




Sources: Virginia State Board of Elections. Results from 2003-2007. Iowa Secretary of State. Results from 2002-2006.

(Bottom-left: by voters/machine) (Bottom-right: by education (& by race))





Voter Participation:
Who & How Many Vote?

TABLE 7.5 *Two Models Explaining Turnout Variations in 31 Countries, 1945–1999*

Variable	Within country *b*	Within country SE	Panel corrected *b*	Panel corrected SE
Constant	25.06	(3.93)**	52.63	(2.14)**
Majority status (0–50%)	−0.13	(0.04)**	−0.16	(0.04)**
Margin of victory (0–70%)	−0.06	(0.04)*	−0.08	(0.03)**
Time since last election (0.6–5)	0.52	(0.18)**	0.37	(0.14)**
Disproportionality (1–20)	−0.01	(0.04)	−0.06	(0.04)
Compulsory voting (0,1)	5.99	(1.99)**	10.92	(0.76)**
Postal voting (0,1)	4.07	(1.96)**	6.79	(0.84)**
Weekend voting (0,1)	−1.57	(0.89)	−0.26	(0.54)
Size of electorate (million)	−0.01	(0.01)	−0.04	(0.01)**
Electoral salience (0,1)			25.46	(2.06)**
Turnout_{t-1}	0.66	(0.04)**		
Missing margin (0,1)	−5.59	(1.66)**	−5.89	(1.58)**
Adjusted R^2	0.506		0.709	
N	403		436	

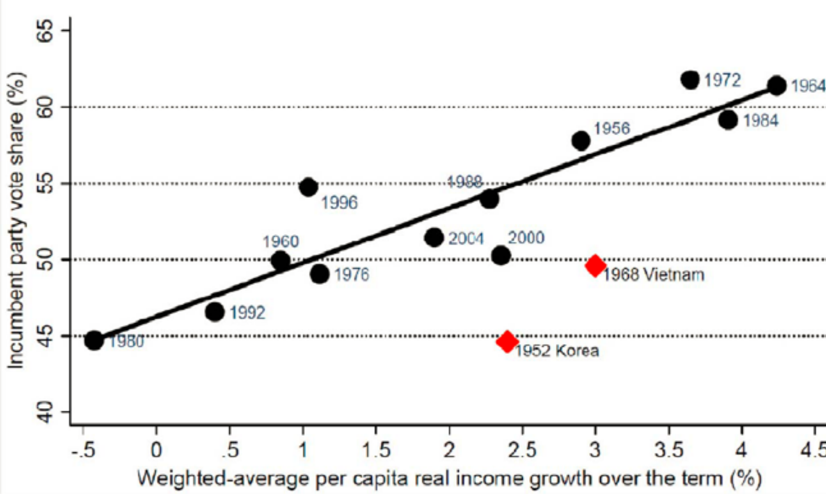
TABLE 8.3 Effects on Individual-Level Electoral Participation in 22 Countries

	Individual Level Only		*With National Effects Considered*		*With Missing Data Indicators*	
Variable	*b*	*SE*	*b*	*SE*	*b*	*SE*
Constant	.636	.017*	.065	.022	.069	.022
Age	.064	.002*	.063	.002*	.062	.002*
Strength of party identification	.010	.004	.040	.004*	.039	.004*
Political discussion	.097	.006*	.091	.006*	.093	.006*
Education	.005	.003	.025	.003*	.025	.003*
Religious participation	.008	.004	.024	.005*	.030	.004*
Union member	–.081	.006*	–.023	.006*	–.024	.006*
Income	.001	.001	.004	.009*	.004	.001*
Average country effect			.478	.017*	.489	.017*
Missing religious participation					–.041	.009*
Adjusted R^2	.055		.195		.195	
N	21,601		21,601		21,601	

Franklin, in *Comparing Democracies*

• **Economic Voting:** know incumbents presiding over stronger economic times tend to do better in elections than incumbents presiding over weaker

Figure 1. Bread and Peace Voting in US Presidential Elections

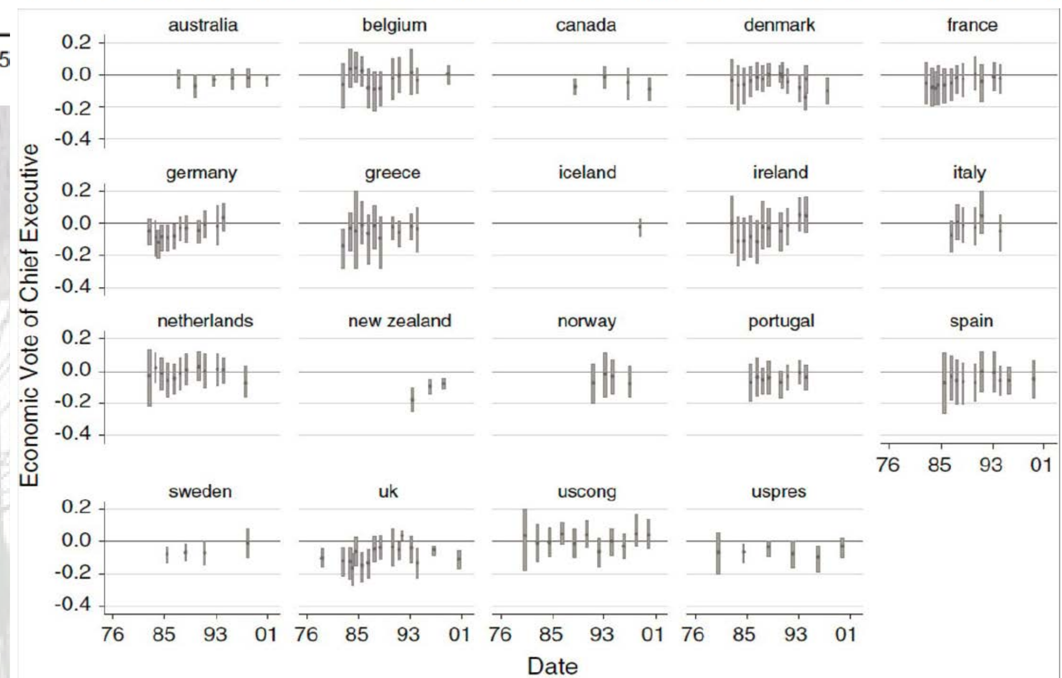


Hibbs' "Bread & Peace" model

Duch & Stevenson, *The Economic Vote*

$$logit(\pi_{ik}) = \beta_{0k} + \beta_{1k}X_{ik} + \sum_{j=1}^{J_k} \phi_{jk}Z_{jik}. \qquad (1)$$

In this notation, v_{ik} indicates a vote for the chief executive party by voter i in each of k ection surveys where $i = 1 \ldots n_k$. Likewise, X_{ik} are retrospective economic evaluations easured at the individual level and Z_{jik} are other characteristics of individuals that shape



g. 1 A map of economic voting for the party of the chief executive. The upper bound of on

Economic Voting: Increasingly Sophisticated Theories

Duch & Stevenson, *The Economic Vote*

Alternatively, same authors, different ways of presenting similar set of estimates:

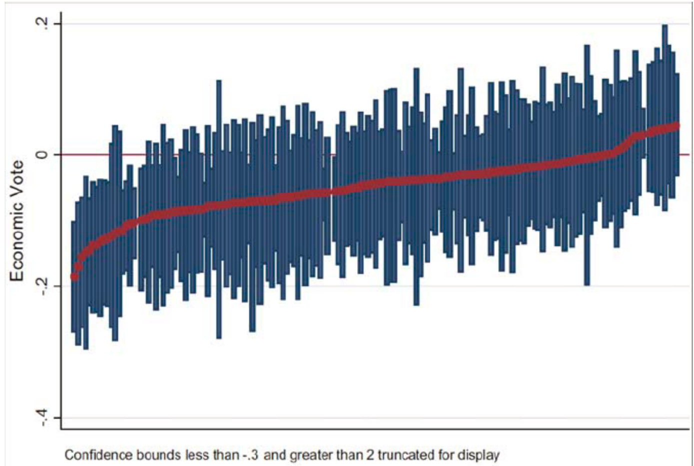
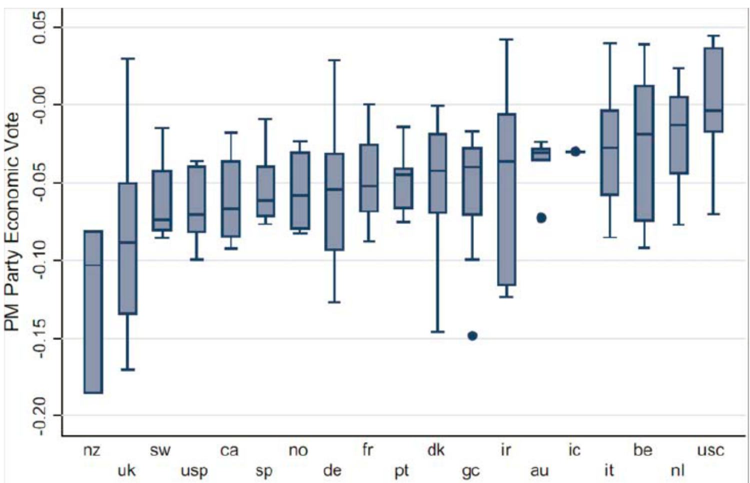


Fig. 1. Confidence bands for economic vote, PM parties.

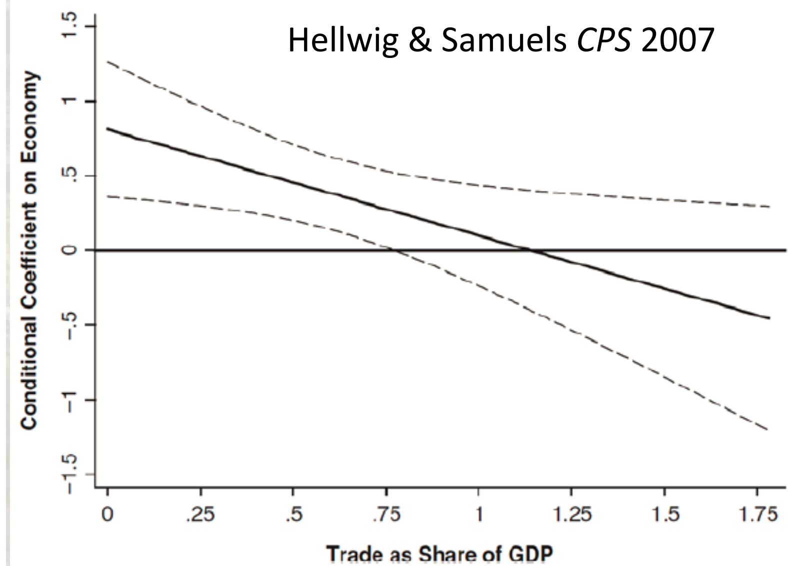
Notice that magnitude of economic vote generally greater in countries with typically single-party, majority governments (e.g., *inter alia*).



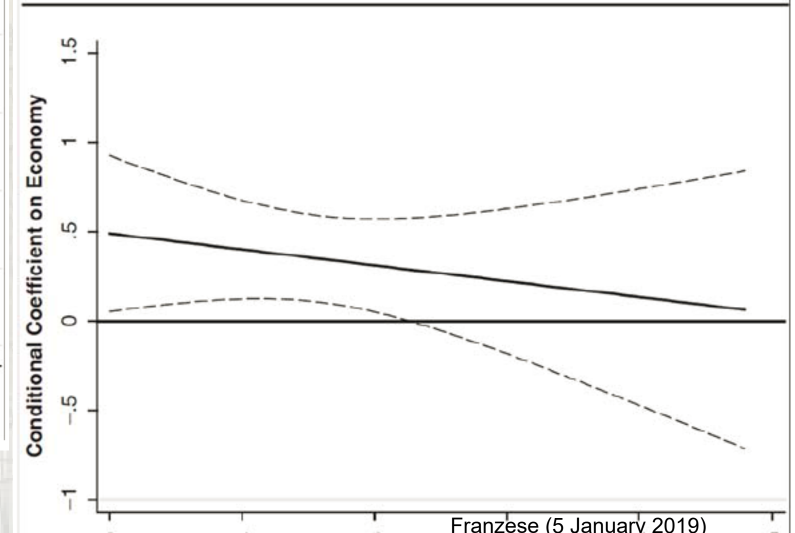
Note: PM Vote responding to economic perceptions that move one unit in worse direction

Effect of Economic Performance on Incumbent Vote Share Under Varying Levels of Trade Openness

Hellwig & Samuels *CPS* 2007



Effect of Economic Performance on Incumbent Vote Share Under Varying Levels of Capital Flows



- **Electoral Cycles**: know incumbents incentives try deliver voters ↑ economic performance & other benefits around election times & ∴ that policies & outcomes tend to exhibit electoral periodicity

Tufte, *Political Control of the Economy*

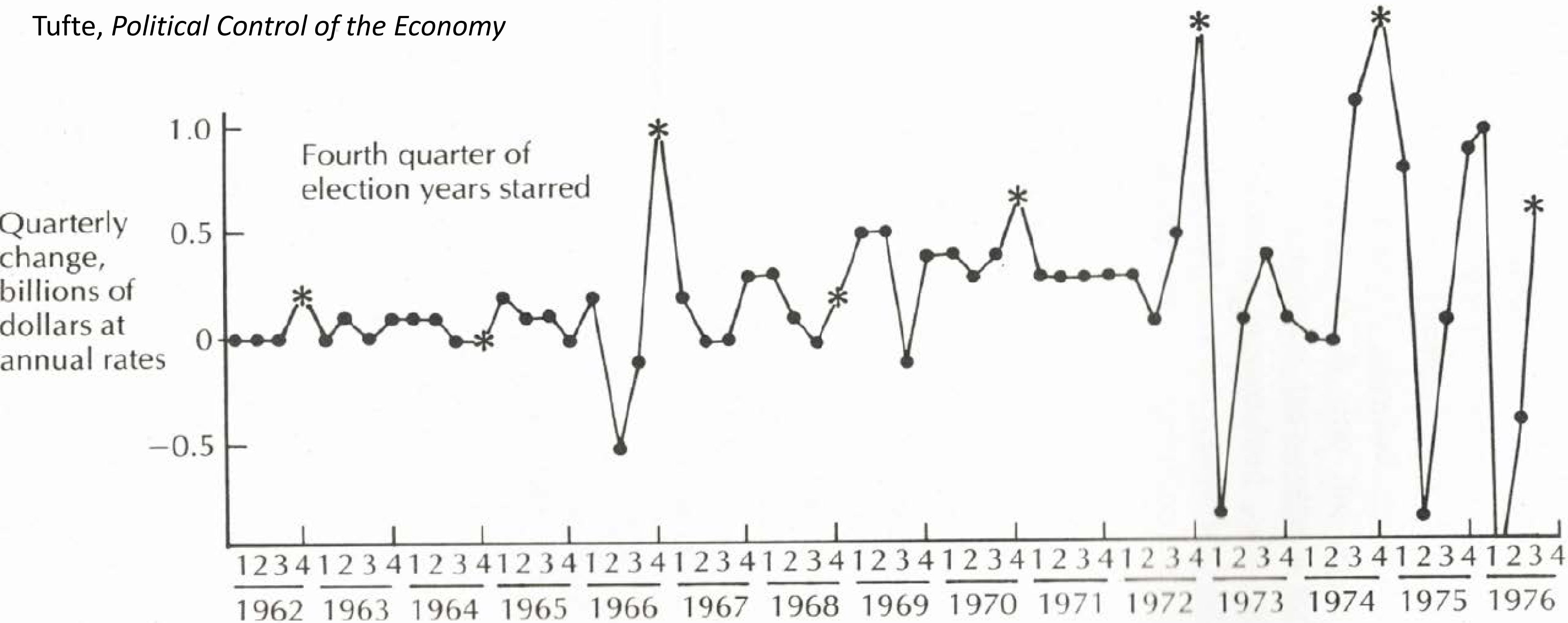


FIGURE 2-3
QUARTERLY CHANGES IN VETERANS BENEFITS

Electoral Cycles: Increasingly Sophisticated Theories

Expected Competitiveness of Election

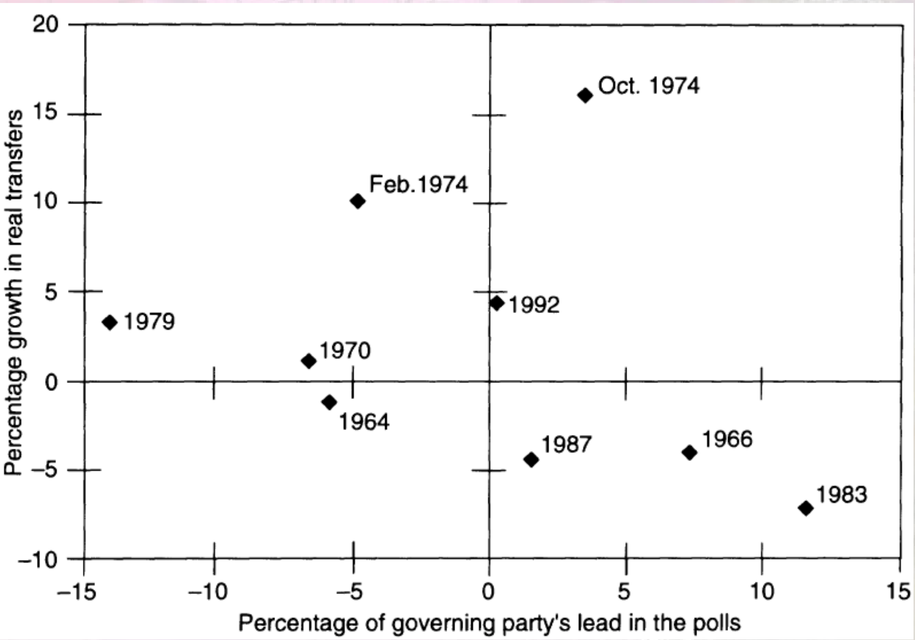


TABLE 2 *Conditional Coefficients on the Pre-Election Dummy*

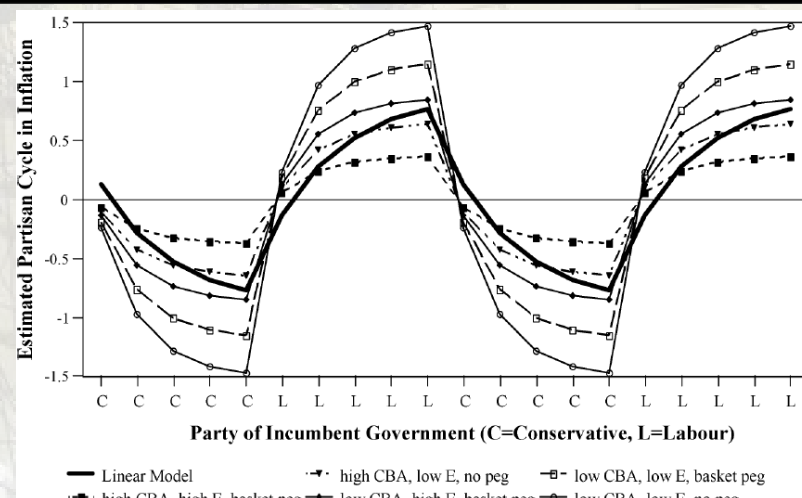
Lead	Conditional coefficient	Conditional standard error	t-statistic
12	− 506.76	206.80	− 2.45***
9	− 381.75	184.18	− 2.07**
6	− 256.74	164.75	− 1.56
3	− 131.73	149.79	− 0.88
0	− 6.72	140.71	− 0.05
− 3	118.29	138.68	0.85
− 6	243.30	144.00	1.69*
− 9	368.31	155.91	2.36***
− 12	493.32	173.07	2.85***

Openness, Exchange-rate Regime, & Crl. Bank Indep.

$$E(\pi) = B_0 + \beta_e E \beta_\pi \pi_a + (1-\beta_e E)\left\{\begin{bmatrix}(\beta_{gp}GP + \beta_{ey}EY + \beta_{up}UP + \beta_{be}BC + \beta_{aw}AW + \beta_{fs}FS \\ \quad + \beta_{te}TE + \beta_a \pi_a) \\ (1-\beta_{c1}C) + \beta_{c1}C\beta_{c2}\end{bmatrix} \\ (1-\beta_{sp}SP - \beta_{mp}MP) + \beta_{sp}SP\beta_{\pi*}\pi_{sp} + \beta_{mp}MP\beta_{\pi*}\pi_{mp}\right\}$$

$$\frac{\partial \pi}{\partial x} = (1-\beta_E E)\cdot\left\{(1-\beta_p P)\cdot\left[(1-\beta_c C)\cdot\beta_x\right]\right\}$$

	E=0.40			*E=0.65*			*E=0.90*		
	SP=MP=0	***MP=1***	***SP=1***	***SP=MP=0***	***MP=1***	***SP=1***	***SP=MP=0***	***MP=1***	***SP=1***
	Estimated Impact of 1-Unit Rightward Shift in Government Partisanship ($d\pi/dGP$)								
0.26	$-0.359^{.17}$	$-0.281^{.15}$	$-0.000^{.02}$	$-0.311^{.15}$	$-0.243^{.13}$	$-0.000^{.02}$	$-0.262^{.12}$	$-0.206^{.11}$	$-0.000^{.01}$
CBA = 0.46	$-0.257^{.12}$	$-0.202^{.10}$	$-0.000^{.01}$	$-0.223^{.10}$	$-0.174^{.09}$	$-0.000^{.01}$	$-0.188^{.09}$	$-0.147^{.08}$	$-0.000^{.01}$
0.66	$-0.156^{.07}$	$-0.122^{.06}$	$-0.000^{.01}$	$-0.135^{.06}$	$-0.106^{.05}$	$-0.000^{.01}$	$-0.114^{.05}$	$-0.089^{.05}$	$-0.000^{.01}$
	Estimated Impact of a Post-Election Year ($d\pi/dEY$)								
0.26	$+1.563^{.79}$	$+1.224^{.61}$	$+0.000^{.09}$	$+1.352^{.69}$	$+1.059^{.53}$	$+0.000^{.07}$	$+1.142^{.60}$	$+0.894^{.47}$	$+0.000^{.06}$
CBA= 0.46	$+1.120^{.57}$	$+0.877^{.44}$	$+0.000^{.06}$	$+0.970^{.50}$	$+0.759^{.39}$	$+0.000^{.05}$	$+0.819^{.44}$	$+0.641^{.34}$	$+0.000^{.05}$
0.66	$+0.678^{.37}$	$+0.531^{.29}$	$+0.000^{.04}$	$+0.587^{.32}$	$+0.459^{.25}$	$+0.000^{.03}$	$+0.495^{.28}$	$+0.388^{.22}$	$+0.000^{.03}$



Causal Inference for Theory Testing

- Yes, but are any of these relations b/w these characteristics of individuals & elections and participation, e.g., ***causal***?
- Neyman-Ruben Causal Model:

$$\text{Causal Effect} = Y_{it}(X=1) - Y_{it}(X=0)$$

  - **Fundamental Problem of Causal Inference…**
    - Compare *Treatment & Control Groups* such that identical in all ways except treatment status &, potentially, outcome.
    - Need rule out: (a) that $Y \Rightarrow X$ (endogeneity, reverse causality) and (b) that some $Z \Rightarrow Y$ *and* $Z \rightarrow X$ (spuriousness).
  - SUTVA: (conditions for ***internal validity*** of experimental causal-inference by difference means treatment & control group)
    - The probability one unit receiving/taking treatment, the (constant) magnitude of the treatment, & the effect of treatment independent of each other & of any other unit(s) receiving/taking treatment, sizes of treatments, or effects of treatments in those others.
    - "The 2 most common ways in which SUTVA can be violated [seems] when (a) there are versions of each treatment varying in effectiveness or (b) there exists interference between units" (Rubin 1990:282).

Strateg*ies* for (Distinctly) Identifying $X \Rightarrow Y$ from $Y \Rightarrow X$ and from $X \Rightarrow Y$ & $Z \rightarrow X$

- **Logical Impossibility:** Occasionally can rule out *a priori* (few *Y* could logically cause race or gender *X*, e.g.)
- **Temporal Precedence:** (*poor man's exogeneity*) If *X* before *Y*, then *Y* cannot $\Rightarrow X$. (potentially problematic in social-science contexts; highly susceptible to specification error)
- **System Specification:** if can specify how $\mathbf{X} \Leftrightarrow \mathbf{Y}$, can get both/all $\mathbf{X} \Rightarrow \mathbf{Y}$ & $\mathbf{Y} \Rightarrow \mathbf{X}$.
- **Instrumentation:** if can establish some $\mathbf{V} \rightarrow \mathbf{X}$ but not $\mathbf{V} \Rightarrow \mathbf{Y}$, except via $\mathbf{V} \rightarrow \mathbf{X}$ and $\mathbf{X} \Rightarrow \mathbf{Y}$, then can use $E(\mathbf{X}|\mathbf{V}) \Rightarrow \mathbf{Y}$.
  - By selves, above not nec'ly block spuriousness (left to statistical control by partialing).
- **Experimentation:** researcher *control* & *randomize* $X \Rightarrow$
  - Y cannot $\Rightarrow X$ (b/c controlled), & no $Z \leftrightarrow X$, even unknown *Zs* (b/c *X* randomized) $\Rightarrow$ not spurious
  - Create Pseudo-Experimental Conditions from Observational Data:
    - **Discontinuity Design:** idea = near cutoff value some indicator, above which X=1 & below X=0, random whether obs. above or below. [sorting; balance failure]
    - **Matching-Based Inference:** idea = if can measure all relevant *Z*, compare Y|X=1 & Y|X=0 for groups balanced (equal distributions) of all *Z*. [(=statistical control on steroids); fail if SUTVA violated (i.e., not clear if/how redress possible $Y \Rightarrow X$); not control *unobservables*]
    - **Difference-in-Difference:** idea = differencing (Y_{it+1}-Y_{it}) nets *all* constant obs-specific *Z*…

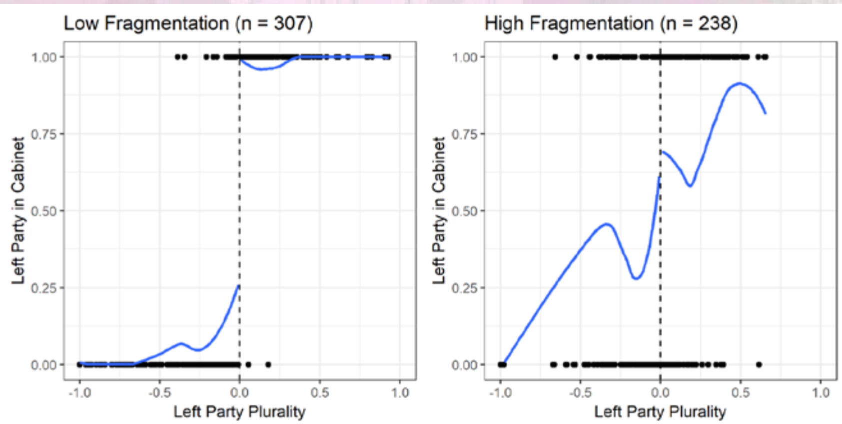
Experiments, the RCT

- **Experiments & Nonparametric Causal-Inference:**
  - Because treatment, X, (a) *randomized* & (b) *controlled*:
    - (a) will not correlate with any other Z (theoretically, in limit),
    - (b) cannot be caused by Y, because researcher controls (causes) it.
  - Also, insofar as $\text{Causal Effect} \equiv Y_{it}(X=1) - Y_{it}(X=0)$
    - Nonparametric, & so independent of functional form for X $\Rightarrow$ Y (and also of controls).
- Much advance in observational studies designed to yield pseudo-experimental conditions for this potential-outcomes framework causal 'effect', and yet, some **Limitations/Insufficiency of Nonparametric Causal-Inference**, to begin for example:
  - "Experiment will have nothing whatsoever to say about other causes. What it will do, and do well, is to determine ***whether*** […treatment…] had a positive or negative effect, or none at all…" (K&W; emph. added)
    - …***ideal to establish that causal effect exists***, not nec'ly great estimating that effect or gauging its substantive magnitude, especially relative to other causes.
      - …although some advances in this latter direction: ***conjoint analysis***.
  - ***Heterogeneous effects*** (e.g., nonlinearity, context conditionality) (***next***); ***External Validity… (later)***; Dynamics & Interdependence, etc.
- Will return to limitations & considerations other modes, but first an example

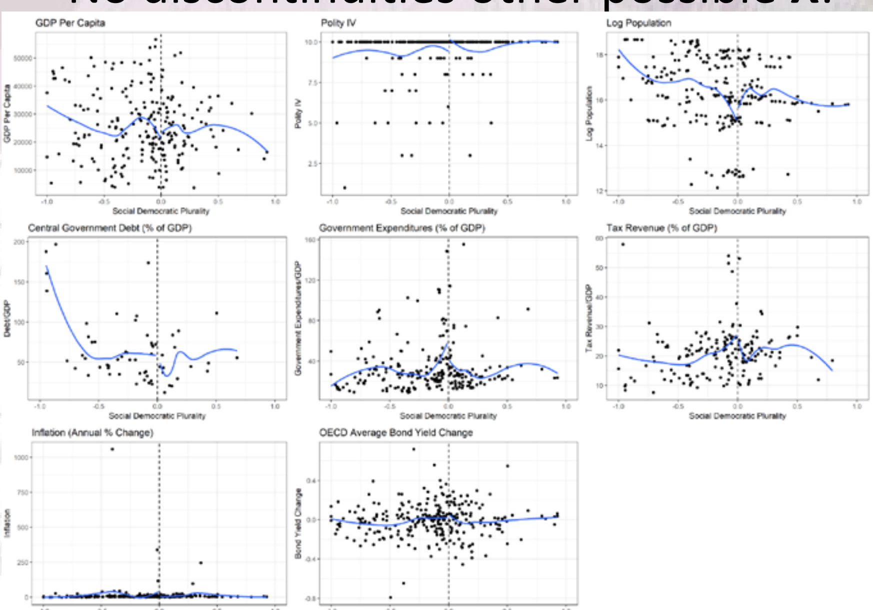
A Discontinuity-Design Test of Causal Effect of Left-Government on Govt-Bond Yields

Hays, Cook, & Franzese (2018)

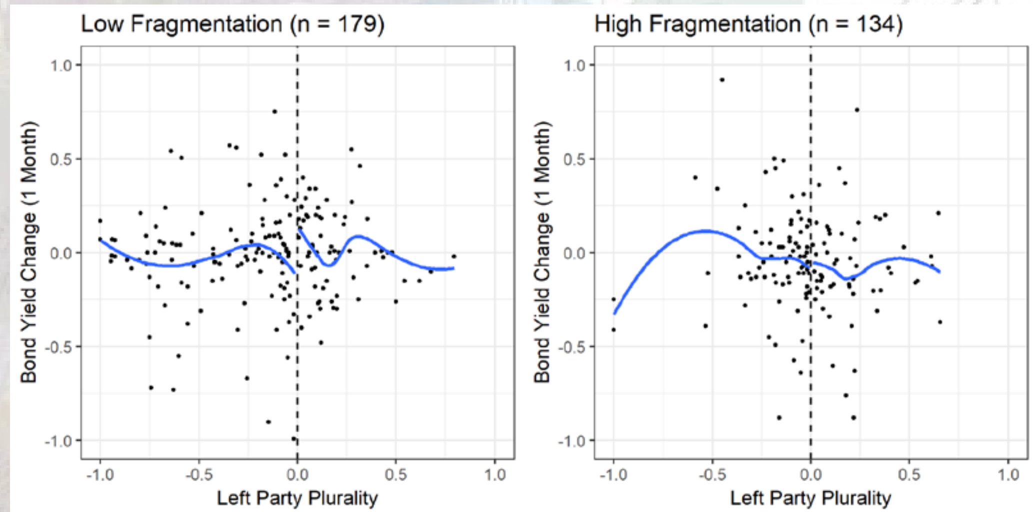
- A Discontinuity:



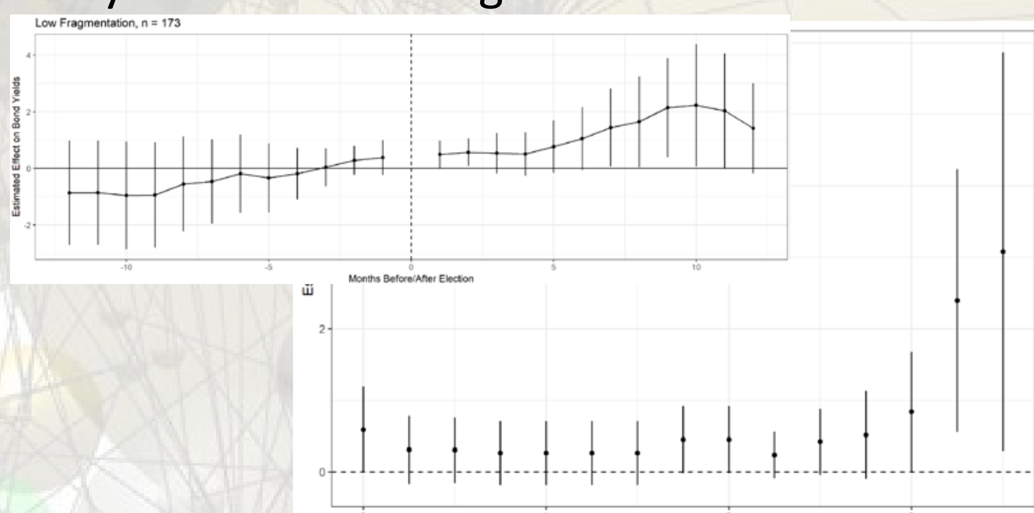
- No discontinuities other possible **X**:



Discontinuity-Design Test & 'Effect' Estimate:



- 'Dynamic' & Heterogenous Effect Estimates:

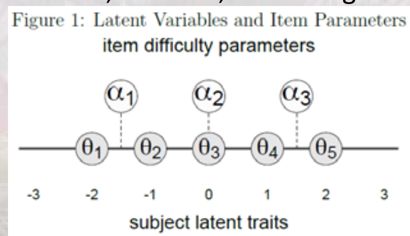


For Some Purposes, Causality is Irrelevant: Measurement, Description, Classification, Prediction

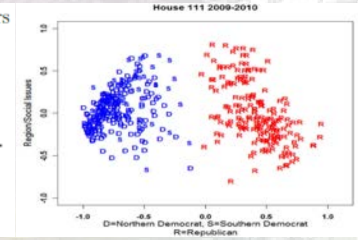
Particularly for Factual, as Opposed Theoretical, Questions…

- Data Resources Booming:
  - Event Data, CLEA, & web scraping, satellite imagery, social media, …
- Measurement Methodologies Advancing:
  - IRT & Bayesian Ideal-Point Est.
  - Network Measures
  - MR & MRS P
  - Scaling & Classifying Text, Sentiment Analysis
- Advances in Visualization
- Prediction:
  - Bayesian MLM & relatives (AMEN, e.g.)
  - Bayesian Model Averaging, Ensemble Methods
  - AI: Supervised, Unsup., & Deep Machine Learning, Natural Language Processing

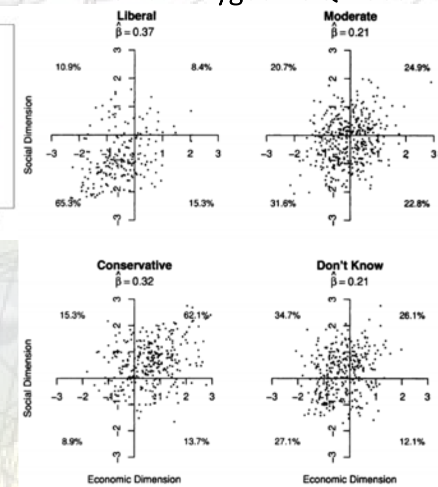
Fariss, Kenwick, & Reuning 2019



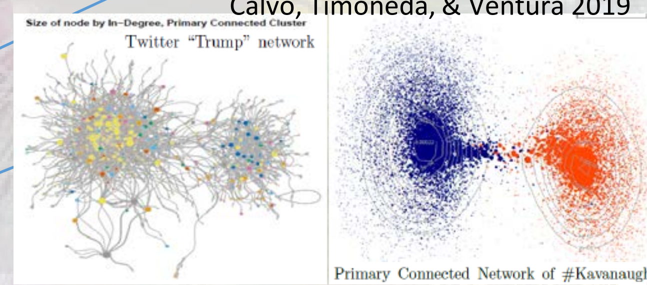
NOMINATE

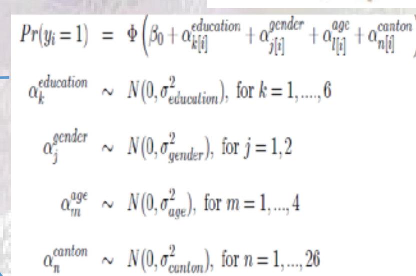


Treier & Hillygus *POQ* 2009

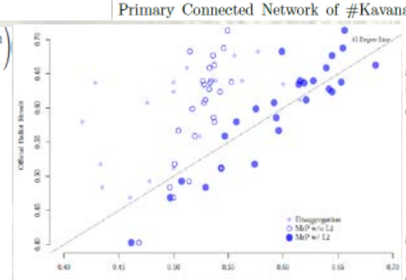


Calvo, Timoneda, & Ventura 2019



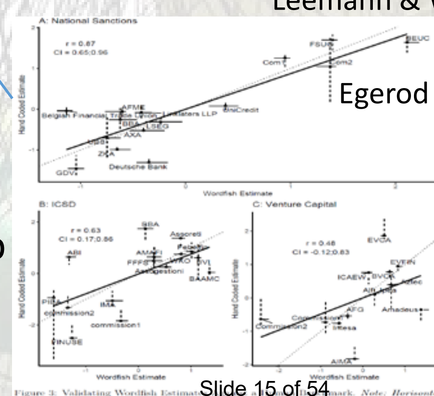


Leemann & Wasserfallen 2019

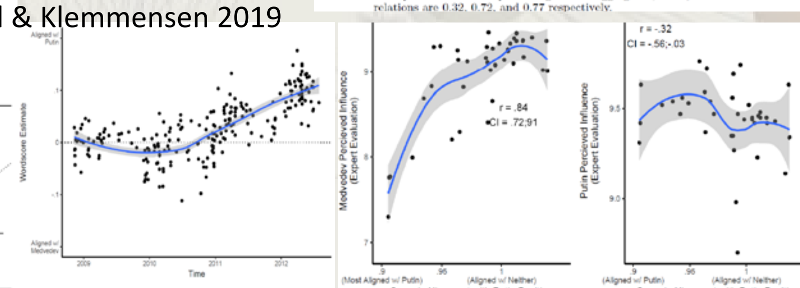


Ornstein 2018

Figure 3.4: Trump 2016 vote share plotted against disaggregated, MRP, and MLP estimates. relations are 0.32, 0.72, and 0.77 respectively.



Egerod & Klemmensen 2019

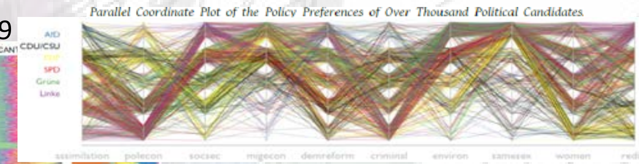


For Some Purposes, Causality may be Irrelevant: Measurement, Description, Classification, Prediction

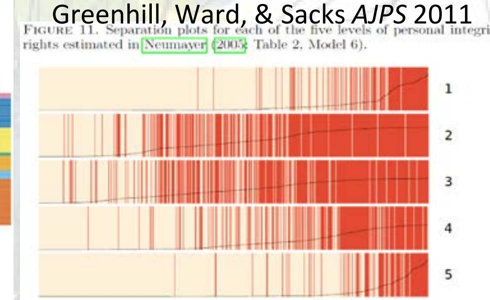
Particularly for Factual, as Opposed Theoretical, Questions…

- Data Resources Booming:
  - Event Data, CLEA, & web scraping, satellite imagery, social media, …
- Measurement Methodologies Advancing:
  - IRT & Bayesian Ideal-Pt. Est.
  - Network Measures
  - MR & MRS P
  - Scaling & Classifying Text, Sentiment Analysis
- Advances in Visualization
- Prediction:
  - Bayesian MLM & relatives (AMEN, e.g.)
  - Bayesian Model Averaging, Ensemble Methods
  - AI: Supervised, Unsup., & Deep Machine Learning, Natural Language Processing

Traunmüller 2019



Greenhill, Ward, & Sacks *AJPS* 2011



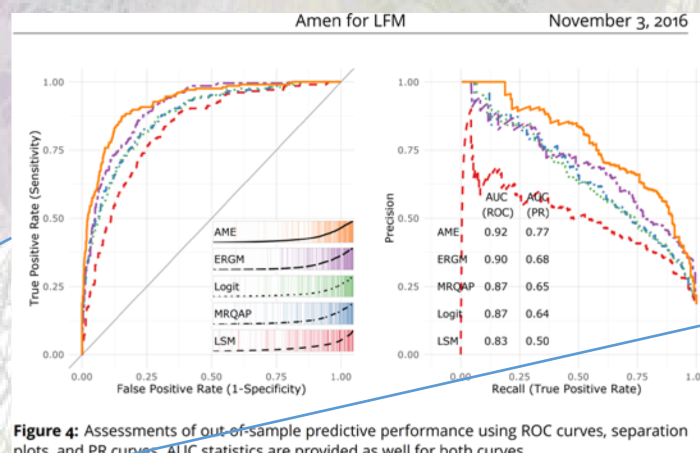
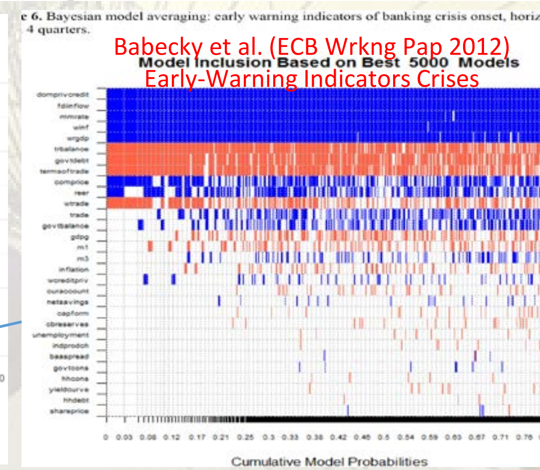
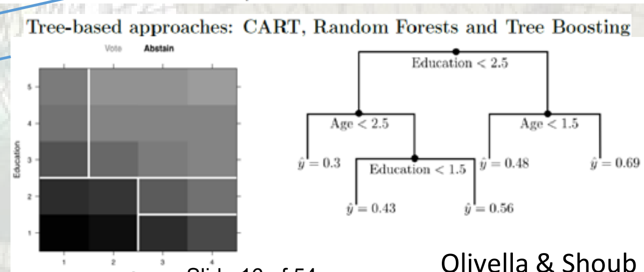
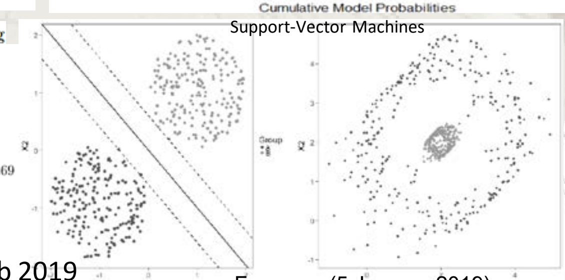


Figure 4: Assessments of out-of-sample predictive performance using ROC curves, separation plots, and PR curves. AUC statistics are provided as well for both curves.





Olivella & Shoub 2019



Franzese (5 January 2019)

…but in addition to *Causal Inference*, testing for existence of causal effects, & to Description/Prediction, another important aim/mode of empirical analysis: ***Empirical-Model & Causal-Effect Estimation***

The Fundamental Challenge*s* of Empirical Analysis

The Socio-Politico-Economic Reality we study is Characterized by:

- **Multicausality:** Just about everything matters…
- **(Heterogeneous Effects &) Context Conditionality:** how just about everything matters depends on just about everything else…
- **(Temporal, Spatial, & Spatiotemporal) Dynamics:** just about everything is moving, not static…
- **Endogeneity:** just about everything causes just about everything else.
  - (*Micronumerosity*: …& we usually have far too little empirical information to figure it all out; n.b., useful variation, not exactly number of observations)
  - (The target (truth, estimand) is moving, but that's just unobserved 2. again…)

A Collection of Concerns about Some Current Fashions in Social-Science Empirical-Research Methodology

- On limits experimentalism as standard for all empirical research in social science.
  - Or why observational research can be a *first* choice (not just when can't do experiment).
  - Out-of-Sample Error: an alternative (better?) gold standard.
- Beyond Causal *Inference* & Toward Causal *Estimation*
  - **Effect Heterogeneity** ⇒ fully non-parametric, model-free estimation not possible.
  - **Dynamics**: highlight difference b/w inferring the existence of a causal effect of treatment & estimating outcome response caused by shock. Cannot estimate latter w/o a (dynamic) model.
  - **Simultaneity**: when x ⇔ y, "nonparametric causal inference" paradoxically estimates causal parameters, and ***not*** causal responses. Cannot estimate latter w/o (system-of-eqtns) model.
- On Empirical Models & Why We Both Need & *Want* Them
  - Curse of dimensionality & logical impossibility fully model-free/nonparametric estimation.
- *To FE or Not To FE* (a usually not *Mostly Harmless* question)
  - "Fixed Effects" cost much more than "mere inefficiency".
  - The limitations of FE likely inherited by FE-like causal-inference strategies…
- In *Social* Phenomena, interdependence, interconnection endogeneity, and/or interdependence by endogenous interconnections (coevolution), imply not-SUTVA.
  - Even on own turf of identifying causal effects, let alone trying estimate causal responses, non-parametric causal-inference tends biased for *social* phenomena (by Rubin's own admission).

On the limits of experimentalism as the standard for ***all*** empirical research in social science.

- In the bible according to Freedman, Pisani, & Purves…
  - Chpt. 2 extols virtues of experimentation; which are two & great:
    - Rules out reverse causality, $Y \Rightarrow X$, because researcher controls X;
    - Rules out confounds, even unobserved ones (in large-samples), because randomized X.
      - [I suspect already here we can raise some doubts: when double-blind randomization is assumed vindicated b/c doctors who know health of patients & nature of their ills yielded better surgical results whereas blinded ones not significantly so…suggests effect heterogeneity that Doc's know & would also use in actual application.]
  - Ch. 3 warns dangers observational research, lacking those 2 great virtues
    - Interesting pattern develops however…each example observational-study conclusion is overturned later by…
    - …another observational study! [with *argument* that better designed]
      - The examples have also shifted from primarily medical in chapter 2 to primarily epidemiological in chapter 3, and epidemiology, like (macro)economics [& political-science!], "is not an experimental science" [Sims 2010].
    - …because ***causality is ultimately a theoretical, not an empirical, matter***

On the limits of experimentalism as the standard for *all* empirical research in social science.

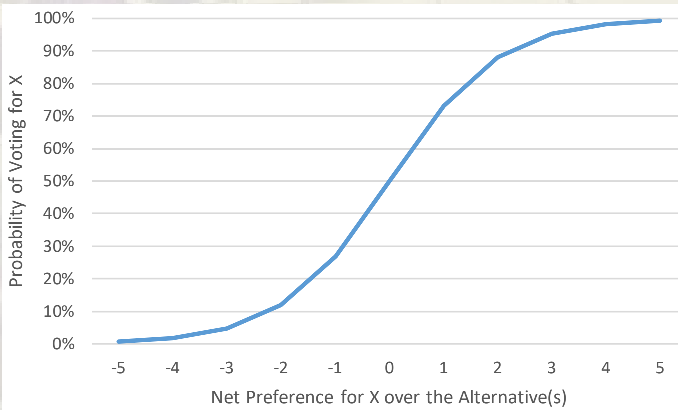
- More fundamentally, we know external validity is problematic
  - Standard Concerns:
    - External Validity of Samples: non-representative
    - External Validity of Treatment: one of the *Princess Bride* problems…
  - Plus, External Validity of Context:
    - Imbens (?2010? "Better LATE than nothing"): cannot imagine situation where could run experiment, and would prefer not to. **I can!**
      - E.g., Korea & Vietnam Wars era U.S. fighter-jet tests got kill ratio totally wrong.
    - [Silly argument about whether internal or external validity lexically primary: some claim that w/o internal validity don't care external; silly b/c want both of course, but if going to argue, obvious that only defensible position is opposite: external w/o internal still value in out-of-sample correlations; internal alone of only esthetic or historical interest, not theoretical scientific but factual descriptive]
  - Problem: by design, arising from their very causal-identification virtues, experiments [& related observational methods] tend to yield poor estimates of *effects*, understood as responses of *y* to exogenous movements *x*:
    - In a system with x ⇔ y, we know that *dx* ⇒ *dy* ⇒ *dx*…
    - The well-designed experiment, & methods designed to isolate the impact of *x* on *y*, like single-eqtn 2SLS or RDD, by design, get only that initial impulse to y…
    - …so, by design, they give lousy estimates of response of *y* to some exogenous impulse to *x*. [Some relevant math will be shown up-close later…]
      - [Not *design trumps control* but *general equilibrium trumps partial equilibrium*.]

On the Limitations/Insufficiency of the Nonparametric, Experimental, Potential-Outcomes-Framework, Causal-Inference Paradigm for Social Science

- Ideal for testing, for evaluating empirical evidence for ***whether causal-effect exists***
- ***Not necessarily for estimating causal effects, understood as dy/dx, how outcomes of interest respond to some cause(s)***
- **External Validity**: of sample…, of treatment…, & of context…
  - [In fact, strictly under paradigm, cannot infer away from support (even though that often the point!).
  - At worst: one obtains cleanly identified estimate of the causal effect of a treatment that could never be applied, in a context that could never obtain, about which we didn't care in the first place…]
- **Multicausality** ⇒ poor gauge effect size, especially relative to others: that's what multiple-regression control is about; conjoint experimentation offers some progress.

Effect Heterogeneity & Context Conditionality:

- *Neyman-Holland-Rubin causal model, is a model*: 'effects' as estimated = additive, constant, separable.
  - E.g., nonlinearity: e.g., substance dictates that for binary outcomes, probabilities, or proportions, Y is sigmoidal *f*(X):
  - A model of probabilities that doesn't respect these first principles (taper toward 0-1 bounds, steeper somehow between) not yield very good estimates for *external* inference (i.e., beyond estimation sample, and esp. not beyond support). (& std NHR⇒d*p*/d*x*=*c*)



ATE's ain't where it's at when world ain't straight.

- **Interactions**, the effect of X on Y depends on Z, and vice versa, similarly challenging for a non-parametric framework.
  - but see Imai et al., e.g., for progress on that front.

Some Fallacies in Our Understanding of the Nonparametric Causal-Inference

- **The Model of the Neyman-Holland-Rubin Causal Model:** Simple not nec'ly = weak, unrestrictive
  - Discrete, Additive, Separable (within & across obs.) Effects of Causes.
    - **Discrete:** to allow interval-valued treatments would be structural. I.e., as applied, inter alia, we are going to select group w/in which treatment homogenous, and simply difference means that v. other groups.
    - **Additive:** mean differencing tends to suffice for the intended purpose (essentially: control), only for linear, purely separable effects
    - **Separable:** So model is a flat line, unconnected to any other treatment's (i.e., treatment of different size, sort, or context) flat line.
  - That's surely a model, incredibly simplistic, yes, but in many ways an extremely strong one. &, as always, insofar as model misspecified, estimates will have poor properties
    - Keane (*JEconometrics* 2006): "criticism of structural econometric work is that it relies on 'too many' assumptions. In fact, I have often seen structural work dismissed out of hand for this reason. In contrast, many believe 'simple' empirical work is more 'convincing.' I readily concede that the typical structural estimation exercise relies on a long list of maintained *a priori* assumptions. But we are kidding ourselves if we think 'simple' estimators don't rely on just as many [or as-strong] assumptions."
    - I.e., the design (& what's done with its estimates) *are* the model. (You say *design*, I say *specification*: *Toe-May-Toe*, *To-Mah-Toe*.) Hard to see how this necessarily any less "model dependence" or any less risk of arbitrariness in this model rather than some other.
- **Matching as a Causal-Inference Strategy:**
  - Matching is just regression control on steroids: latter controls linear-additive-separable affects of $\mathbf{X}_c$, former controls any separable effects of $\mathbf{X}_c$. As such:
  - Matching *per se* is not a causal-identification strategy; to get causal-parameter estimates, must both observe $\mathbf{X}_c$ & assume them exogenous (pre-treatment), just like regression.
- **Given potential arbitrary effect-heterogeneity, fully nonparametric estimation impossible**

An Alternative Approach Suited to *Causal-Response Estimation*: *Theory/Substance-Based Empirical Modeling*

So what to do with Complex Context-Conditionality?

- **Empirical Models of Theoretical Intuitions (EMTI):**
  - Core Implication Theory: $\mathbf{y}=f\left(\mathbf{X},\mathbf{B},\boldsymbol{\varepsilon}\right)\underset{\text{if sep.}}{\Rightarrow}E\left(\mathbf{y}\right)=f\left(\mathbf{X},\mathbf{B}\right),\boldsymbol{\varepsilon}\sim g\left(\boldsymbol{\varepsilon}\right)$
    - EMTI emphasizes far too little typically drawn from theoretically implied *f*(·), *g*(·)
    - Theoretical model or intuitions and substance tend suggest more about some specific *f*(·) than, & not always or even often, that linear-additive.
      - Usually theory used just to suggest *x* as arg's, entered linear-additively by default, to regression/likelihood. (Or, worse, some *T* to isolate for causal-effect inference.) Hypotheses confined to first *partial* derivatives, not responses.
- EMTI ⇒ ***Model it!*** ™…& then, when modeling it:
  - ***Specification* is everything***.
    - * Note: specification (or design) includes measurement & identification strategy.
- Example: Two Hands on Wheel (shared policy-control)
  - ⇒ $y=\underbrace{c(p)\times f(\mathbf{x}_p)}_{\text{principle cntrl}\,\times\,\text{p action}}+\underbrace{[1-c(p)]\times g(\mathbf{x}_a)}_{\text{agent control}\,\times\,\text{agent action}}\Rightarrow$ many interesting things…
  - E.g., the effect on *y* of any $x\in(\mathbf{x}_p\cup\mathbf{x}_a)$ to which principle & agent would respond differently, depends on *c*(*p*)…

An EMTI Strategy for the Pervasive, & often Complex, Context-Conditionality of SocPolEco Reality

- **Empirical Modeling of Theoretical Intuitions:**
  - Theory & substance indicate what sort of random variable makes sense as type for outcome.
  - Random variables have distributions/densities; those have parameters that correspond to aspects of interest about that RV (outcome).
  - Substance suggests an appropriate form for such a parameter and theory suggests a model linking explanators (covariates) to those parameters by such a function.
  - If first & second moments additively separable, least squares is an available & effective estimation strategy. If not, maximum likelihood is available & effective, and almost as simple if observations conditionally (on model) independent.
- **Least-Squares Estimation:** $\underbrace{E(y) = f(\mathbf{x}, \boldsymbol{\beta})}_{\text{substance \& theory}} \Rightarrow \underset{\mathbf{b}}{Min}(\mathbf{y} - f(\mathbf{x}, \mathbf{b}))'(\mathbf{y} - f(\mathbf{x}, \mathbf{b}))$
- **Maximum-Likelihood Estimation:**

$$\underbrace{p(y_i \mid \boldsymbol{\theta}), \text{ cond'l indep } \Rightarrow p(\mathbf{y} \mid \boldsymbol{\theta}) = \prod_i p(y_i \mid \boldsymbol{\theta}), \ \boldsymbol{\theta} = f(\mathbf{x}, \mathbf{b})}_{\text{substance and theory}} \Rightarrow \underset{\mathbf{b}}{Max} \sum_i \ln p(y_i \mid f(\mathbf{x}, \mathbf{b}))$$

(Complex) Context-Conditionality: (Hallmark of Modern Soc-Sci Theory?)

- ***Complex Context-Conditionality***:
  - Effect of (almost) anything depends on (almost) everything else, often *complexly*
- **Principal-Agent (Shared-Control) Situations, for example:**
  - Equilibrium PA/Bargaining Models some convex combination actors' ideals.
  - If fully agent, $y_1=f(\mathbf{X})$; if fully principal, $y_2=g(\mathbf{Z})$; institutions: $0\leq h(\mathbf{I})\leq 1$ (eg, $h(\mathbf{I})$:monitor+enforce cost)
- **RESULT:**

$$y = h(\mathbf{I})\,f(\mathbf{X}) + \{1 - h(\mathbf{I})\}\,g(\mathbf{Z})$$

- **In words…**

$$\Rightarrow \frac{\partial y}{\partial x} = h(\mathbf{I})\frac{\partial f(\mathbf{X})}{\partial x} \;;$$

…

…

$$\frac{\partial y}{\partial z} = -h(\mathbf{I})\frac{\partial g(\mathbf{Z})}{\partial z} \;;$$

…i.e., effect of anything depends on everything else!

$$\frac{\partial y}{\partial i} = \frac{\partial h(\mathbf{I})}{\partial i}\left[\,f(\mathbf{X}) - g(\mathbf{Z})\right]$$

"Multiple Hands on the Wheel" Model (Franzese PA '03)

- Start with CapMobility × ERpeg × CBindep:

$$\pi = \begin{cases} P\cdot E\cdot C\cdot \pi_1(\boldsymbol{X}_1) + P\cdot E\cdot (1-C)\cdot \pi_2(\boldsymbol{X}_2) \\ +P\cdot (1-E)\cdot C\cdot \pi_3(\boldsymbol{X}_3) + P\cdot (1-E)\cdot (1-C)\cdot \pi_4(\boldsymbol{X}_4) \\ (1-P)\cdot E\cdot C\cdot \pi_5(\boldsymbol{X}_5) + (1-P)\cdot E\cdot (1-C)\cdot \pi_6(\boldsymbol{X}_6) \\ +(1-P)\cdot (1-E)\cdot C\cdot \pi_7(\boldsymbol{X}_7) + (1-P)\cdot (1-E)\cdot (1-C)\cdot \pi_8(\boldsymbol{X}_8) \end{cases}$$

- Central Bank & Government Interaction (Franzese *AJPS* '99):

$$E(\pi) = c\cdot \pi_c(\mathbf{x}_c) + (1-c)\cdot \pi_g(\mathbf{x}_g)$$

$$\pi_c = \overline{\pi}_c \qquad \pi_g(\mathbf{x}_g) = \pi_g(GP, UD, BC, TE, EY, FS, AW, \pi_a)$$

- Full Monetary <u>E</u>xposure & Atomistic ⇒ zero domestic autonomy ⇒

$$\widehat{\pi_1(\mathbf{x}_1)} = \widehat{\pi_2(\mathbf{x}_2)} = \widehat{\pi_5(\mathbf{x}_5)} = \widehat{\pi_6(\mathbf{x}_6)} = \pi_a$$

$$\Rightarrow E\cdot \pi_a + (1-E)\cdot \begin{Bmatrix} P\cdot C\cdot \pi_3(\mathbf{x}_3) + P\cdot (1-C)\cdot \pi_4(\mathbf{x}_4) \\ +(1-P)\cdot C\cdot \overline{\pi}_c + (1-P)\cdot (1-C)\cdot \pi_g(\mathbf{x}_8) \end{Bmatrix}$$

- s.t. that, full e.r. fix ⇒ CB&Gov match peg ⇒

$$\widehat{\pi_3(\mathbf{x}_3)} = \widehat{\pi_4(\mathbf{x}_4)} = \pi_p \Rightarrow E\cdot \pi_a + (1-E)\cdot \left\{P\cdot \pi_p + (1-P)\cdot \left[C\cdot \overline{\pi}_c + (1-C)\cdot \pi_g(\mathbf{x}_8)\right]\right\}$$

“Multiple Hands on the Wheel” Model (Franzese PA ‘03)

- Compact & intuitive, yet gives all theoretically expected interactions, in the form expected

$$\pi = E\cdot\pi_a + (1-E)\cdot\left\{P\cdot\pi_p + (1-P)\cdot\left[C\cdot\overline{\pi_c} + (1-C)\cdot\pi_g(\boldsymbol{X_g})\right]\right\}$$

$$\Rightarrow$$

$$\frac{\partial\pi}{\partial E} = \pi_a\left(P^*,E^*,C^*,\boldsymbol{X}^*,\pi_a^*\right) - \left\{P\cdot\pi_p\left(P^*,E^*,C^*,\boldsymbol{X}^*,\pi_p^*\right) + (1-P)\cdot\left[C\cdot\overline{\pi_c} + (1-C)\cdot\pi_g(\boldsymbol{X_g})\right]\right\}$$

$$\frac{\partial\pi}{\partial P} = (1-E)\cdot\left\{\pi_p\left(P^*,E^*,C^*,\boldsymbol{X}^*,\pi_p^*\right) - \left[C\cdot\overline{\pi_c} + (1-C)\cdot\pi_g(\boldsymbol{X_g})\right]\right\}$$

$$\frac{\partial\pi}{\partial C} = (1-E)\cdot\left\{(1-P)\cdot\left[\overline{\pi_c} - \pi_g(\boldsymbol{X_g})\right]\right\}$$

$$\frac{\partial\pi}{\partial x} = (1-E)\cdot\left\{(1-P)\cdot\left[(1-C)\cdot\frac{\partial\pi_g}{\partial x}\right]\right\}$$

$$\frac{\partial\pi}{\partial z^*} = E\cdot\frac{\partial\pi_a}{\partial z^*} + (1-E)\cdot\left\{P\cdot\frac{\partial\pi_p}{\partial z^*} + (1-P)\cdot\left[(1-C)\cdot\frac{\partial\pi_g}{\partial\pi_a}\cdot\frac{\partial\pi_a}{\partial z^*}\right]\right\}$$

“Multiple Hands on the Wheel” Model (Franzese PA ‘03)

- Effectively Estimable, yet gives all theoretically expected interactions, in the form expected

$$E(\pi) = \boldsymbol{B_0} + \beta_e E \cdot \beta_{\pi *}\pi_a + \left(1-\beta_e E\right) \cdot \left\{ \begin{array}{l} \left[\begin{array}{l} \left(\beta_{gp} GP + \beta_{ey} EY + \beta_{up} UP + \beta_{bc} BC + \beta_{aw} AW + \beta_{fs} FS + \beta_{te} TE + \beta_a \pi_a\right) \\ \cdot \left(1-\beta_{c1} C\right) + \beta_{c1} C \cdot \beta_{c2} \end{array} \right] \\ \cdot (1-\beta_{sp} SP - \beta_{mp} MP) + \beta_{sp} SP \cdot \beta_{\pi *}\pi_{sp} + \beta_{mp} MP \cdot \beta_{\pi *}\pi_{mp} \end{array} \right\}$$

- Just 14 parameters (plus intercepts & dynamics, assuming those constant), just 3 more than lin-add!
- Parameters substantive meaning, too:
  - Degree to which…constrains certain set of actors.
  - Yields est. of inflation-target hypothetical fully indep CB
    - $\Rightarrow$ general strategy for estimating/measuring unobservables
      - If know role factor will play & explanators of factor well enough, can estimate unobservables conditional on both those theories, if both powerful enough & enough empirical variation.

“Multiple Hands on the Wheel” Model (Franzese PA ‘03)

- Neat, but does it work? (**Easy!** stata: **nl**; R **nls** in **dynlm**. Estimated Equation, w/ Std. Errs.:

$$E(\pi)\approx\left\langle \begin{array}{l} .53^{.30}+.55^{.05}\pi_{t-1}-.12^{.04}\pi_{t-2}+.44^{.14}E\cdot\pi_a+ \\ (1-.44^{.14}E)\cdot\left\{\begin{array}{l} 1.0^{.05}SP\cdot .59^{.07}\pi_{sp}+.22^{.12}MP\cdot .59^{.07}\pi_{mp}+ \\ (1-1.0^{.05}SP-.22^{.12}MP)\cdot\left[\begin{array}{l} 1.0^{.11}C\cdot(-.59^{1.2})+ \\ (1-1.0^{.11}C)\cdot\left(\begin{array}{l} -.60^{.30}GP+2.6^{1.3}EY+16^{4.6}UP-11^{2.4}BC \\ +1.2^{.49}AW-1.1^{.30}FS-8.2^{4.9}TE+.64^{.24}\pi_a \end{array}\right) \end{array}\right] \end{array}\right\} \end{array}\right.$$

- Estimated Effects (highly context-conditional):

$$E\left(\frac{d\pi}{dC}\right)=(1-.44\cdot E)\cdot\left\{(1-b_pP)\cdot\left[(.6GP-2.6EY-16UP+11BC-1.2AW+1.1FS+8.2TE-.64\pi_a)-.59\right]\right\}$$

$$E\left(\frac{d\pi}{dx}\right)=(1-.44E)\cdot\left\{(1-SP-.22MP)\cdot\left[(1-C)\cdot b_x\right]\right\}$$

$$E\left(\frac{d\pi}{dP}\right)=(1-.44E)\cdot b_p\cdot\left\{.59\pi_p-\left[(1-C)\cdot(-.6GP+2.6EY+16UP-11BC+1.2AW-1.1FS-8.2TE+.64\pi_a)-.59C\right]\right\}$$

$$E\left(\frac{d\pi}{dE}\right)=.44\cdot\left(\pi_a-\left\{b_pP\cdot .59\pi_p+(1-b_pP)\cdot\left[(1-C)\cdot(-.6GP+2.6EY+16UP-11BC+1.2AW-1.1FS-8.2TE+.64\pi_a)-.59C\right]\right\}\right)$$

Table 1 Alternative models of inflation in 21 OECD democracies, 1957–1990

Independent variable	*Linear-additive model (12)*	*Linear-interactive model (13)* $C=1, E=1, P=1$	$C=1, E=1, P=0$	$C=1, E=0, P=1$	$C=1, E=0, P=0$	$C=0, E=1, P=1$	$C=0, E=1, P=0$	$C=0, E=0, P=1$	$C=0, E=0, P=0$	*Theory-informed model (14)*
Intercept	+.80	+5.93								**+.53**
	(6.1)	(8.40)								**(.30)**
Lagged inflation	**+.65**	**+.51**								**+.55**
(π_{t-1})	**(.05)**	**(.06)**								**(.05)**
Twice-lagged	−.03	**−.10**								**−.12**
inflation (π_{t-2})	(.04)	**(.04)**								**(.04)**
Government partisanship	**−.14**	*+.39*	*−.09*	**−3.37**	−1.37	−.15	−.30	**+1.82**	−.39	**−.60**
(GP $\in \mathbf{X}_g$)	**(.08)**	*(.80)*	(1.29)	**(1.31)**	(8.16)	(.47)	(.97)	**(.74)**	(4.68)	**(.30)**
Postelection year	**+.59**	+.75	*−2.06*	+.50	*−.88*	*−2.31*	***+6.03***	+1.87	+3.81	**+2.60**
(EY $\in \mathbf{X}_g$)	**(.30)**	(.80)	*(2.31)*	(3.07)	*(14.67)*	*(1.56)*	***(3.46)***	(1.81)	(6.88)	**(1.32)**
Union power	**+2.19**	***−16.59***	+9.51	*−3.82*	*−2.46*	***+33.95***	+2.44	*−11.88*	*−3.32*	**+16.2**
(UP $\in \mathbf{X}_g$)	**(.74)**	***(6.43)***	(17.42)	*(13.91)*	*(59.24)*	***(7.64)***	(15.92)	*(13.56)*	*(37.49)*	**(4.61)**
Coordination of	**−1.36**	*+4.38*	***+11.27***	*+6.02*	*−39.11*	***−15.61***	−11.69	*+2.20*	*+9.27*	**−10.7**
bargaining (BC $\in \mathbf{X}_g$)	**(.41)**	*(3.50)*	**(5.33)**	*(4.91)*	*(30.32)*	***(3.97)***	(9.79)	*(3.86)*	*(23.64)*	**(2.35)**
Aggregate	+.13	−.76	−2.37	+1.94	***+13.70***	−.56	−.66	−2.24	−3.43	**+1.18**
wealth (AW $\in \mathbf{X}_g$)	(.71)	(1.15)	(1.51)	(1.43)	***(5.37)***	(1.10)	(1.38)	(1.91)	(2.35)	**(.49)**
Financial-sector size	**−.15**	**−.86**	***+2.00***	***+2.11***	***−11.13***	*+.55*	−1.64	−1.00	*+4.63*	**−1.09**
(FS $\in \mathbf{X}_g$)	**(.10)**	**(.36)**	***(.96)***	***(.79)***	***(4.61)***	*(.36)*	(1.26)	(.71)	*(3.90)*	**(.30)**
Trade exposure	−.04	***+31.74***	***−50.21***	*−54.49*	*+50.81*	***−37.33***	***+104.56***	*+48.70*	*−120.5*	**−8.23**
(TE$\in \mathbf{X}_g$)	(.99)	***(14.33)***	***(25.31)***	*(39.85)*	*(176.99)*	***(14.87)***	***(30.40)***	*(33.74)*	*(103.79)*	**(4.92)**
Inflation abroad	**+.39**	+.24	**+.89**	*−.07*	*−4.01*	**+.89**	+.18	**+.98**	*+2.65*	**+.64**
($\pi_a \in \mathbf{X}_g$)	**(.07)**	(.14)	**(.52)**	*(.59)*	*(3.94)*	**(.31)**	(.78)	**(.33)**	*(2.58)*	**(.24)**
Global-financial	+.29	—								**+.44**
exposure (E)	(.75)									**(.14)**
Single-currency	−.33	—								**+1.04**
(simple) peg (SP)	(.49)									**(.05)**
Multi-currency	−.37	—								**+.22**
(basket) peg (MP)	(.38)									**(.12)**
Peg or global inflation	—	—								**+.59**
($\pi_{sp}, \pi_{mp}, \pi_a$)										**(.07)**
Central bank	**−1.62**	—								**+1.03**
independence (C)	**(.68)**									**(.11)**
Central bank		—								**−.59**
target ($\bar{\pi}_c$)										**(1.18)**
Obs. (°Free)	660 (645)	660 (593)								660 (643)
R^2 (S.E.R.)	.72 (2.48)	.75 (2.31)								.76 (2.30)
D-W	1.91	2.03								1.96

Notes. Estimation by nonlinear least-squares, (14), or ordinary least-squares, (12) and (13), with Newey-West robust variance–covariance matrix. Standard errors in parentheses. Coefficients significant at $p = .10$ or better in bold; coefficients of implausible sign or magnitude in italic; and coefficients both significant and implausible in bold-italic. Independent variables labeled $x \in \mathbf{X}_g$ are the political–economic conditions modeled in (14) as those to which domestic governments respond, which response central bank independence, global-financial exposure, and exchange-rate pegs dampen.

- Notice the nonlinear model respecting the combinatorial form implied by substance & theory captures the complex context-conditionality with just 2 parameters more than the linear-additive model.
- Notice the crazy coefficient estimates in the multicolinear nightmare linear-inter. model
- Notice the nonlinear model obtains 5.5% improvement adjusted R^2 over linear & even a 1.33% gain over the 50-parameter larger linear-interaction model.

Context-Conditional Inflation Effects of Political-Economic Factors

Table 2: Estimated Effects of Domestic Political-Economic Conditions, $d\pi/x$, as Function of Central Bank Autonomy, *CBA*, International Monetary Exposure, *E*, and Exchange-Rate Regime, *P*

		Little Exposed (E=0.40)			*Moderately Exposed (E=0.65)*			*Highly Exposed (E=0.90)*		
		Float	*Basket Peg*	*Simple Peg*	*Float*	*Basket Peg*	*Simple Peg*	*Float*	*Basket Peg*	*Simple Peg*
		Estimated Impact of a Post-Election Year ($d\pi/dEY$)								
central	*0.26*	$+1.563^{.79}$	$+1.224^{.61}$	$+0.000^{.09}$	$+1.352^{.69}$	$+1.059^{.53}$	$+0.000^{.07}$	$+1.142^{.60}$	$+0.894^{.47}$	$+0.000^{.06}$
bank	*0.46*	$+1.120^{.57}$	$+0.877^{.44}$	$+0.000^{.06}$	$+0.970^{.50}$	$+0.759^{.39}$	$+0.000^{.05}$	$+0.819^{.44}$	$+0.641^{.34}$	$+0.000^{.05}$
auton.	*0.66*	$+0.678^{.37}$	$+0.531^{.29}$	$+0.000^{.04}$	$+0.587^{.32}$	$+0.459^{.25}$	$+0.000^{.03}$	$+0.495^{.28}$	$+0.388^{.22}$	$+0.000^{.03}$
		Estimated Impact of 10% Increase in Union Density ($0.1\cdot d\pi/dUP$)								
central	*0.26*	$+0.98^{.25}$	$+0.76^{.18}$	$+0.00^{.05}$	$+0.84^{.21}$	$+0.66^{.16}$	$+0.00^{.04}$	$+0.71^{.19}$	$+0.56^{.14}$	$+0.00^{.04}$
bank	*0.46*	$+0.70^{.18}$	$+0.55^{.13}$	$+0.00^{.04}$	$+0.61^{.15}$	$+0.47^{.11}$	$+0.00^{.03}$	$+0.51^{.14}$	$+0.40^{.10}$	$+0.00^{.03}$
auton.	*0.66*	$+0.42^{.13}$	$+0.33^{.10}$	$+0.00^{.02}$	$+0.37^{.11}$	$+0.29^{.08}$	$+0.00^{.02}$	$+0.31^{.10}$	$+0.24^{.08}$	$+0.00^{.02}$
		Estimated Impact of 1% Increase in Financial-Sector Employment-Share ($d\pi/dFS$)								
central	*0.26*	$-0.66^{.18}$	$-0.52^{.12}$	$-0.00^{.03}$	$-0.57^{.16}$	$-0.45^{.11}$	$-0.00^{.03}$	$-0.48^{.15}$	$-0.38^{.11}$	$-0.00^{.03}$
bank	*0.46*	$-0.47^{.13}$	$-0.37^{.09}$	$-0.00^{.02}$	$-0.41^{.12}$	$-0.32^{.08}$	$-0.00^{.02}$	$-0.35^{.11}$	$-0.27^{.08}$	$-0.00^{.02}$
auton.	*0.66*	$-0.29^{.10}$	$-0.22^{.07}$	$-0.00^{.01}$	$-0.25^{.09}$	$-0.19^{.06}$	$-0.00^{.01}$	$-0.21^{.08}$	$-0.16^{.06}$	$-0.00^{.01}$
		Estimated Impact of 1% Increase in Average Inflation Abroad ($d\pi/d\pi_{a}$)								
central	*0.26*	$+0.49^{.14}$	$+0.41^{.13}$	$+0.11^{.05}$	$+0.50^{.12}$	$+0.43^{.11}$	$+0.17^{.07}$	$+0.52^{.10}$	$+0.46^{.10}$	$+0.24^{.09}$
bank	*0.46*	$+0.38^{.10}$	$+0.32^{.09}$	$+0.11^{.04}$	$+0.41^{.08}$	$+0.36^{.08}$	$+0.17^{.06}$	$+0.44^{.08}$	$+0.39^{.08}$	$+0.24^{.08}$
auton.	*0.66*	$+0.27^{.06}$	$+0.24^{.06}$	$+0.11^{.04}$	$+0.32^{.06}$	$+0.28^{.06}$	$+0.17^{.06}$	$+0.36^{.06}$	$+0.33^{.06}$	$+0.24^{.08}$

NOTES: These are *first-year effects*, meaning before the estimated dynamics unfold. Standard errors noted in superscripts.

Context-Conditional Partisan Inflation-Cycles

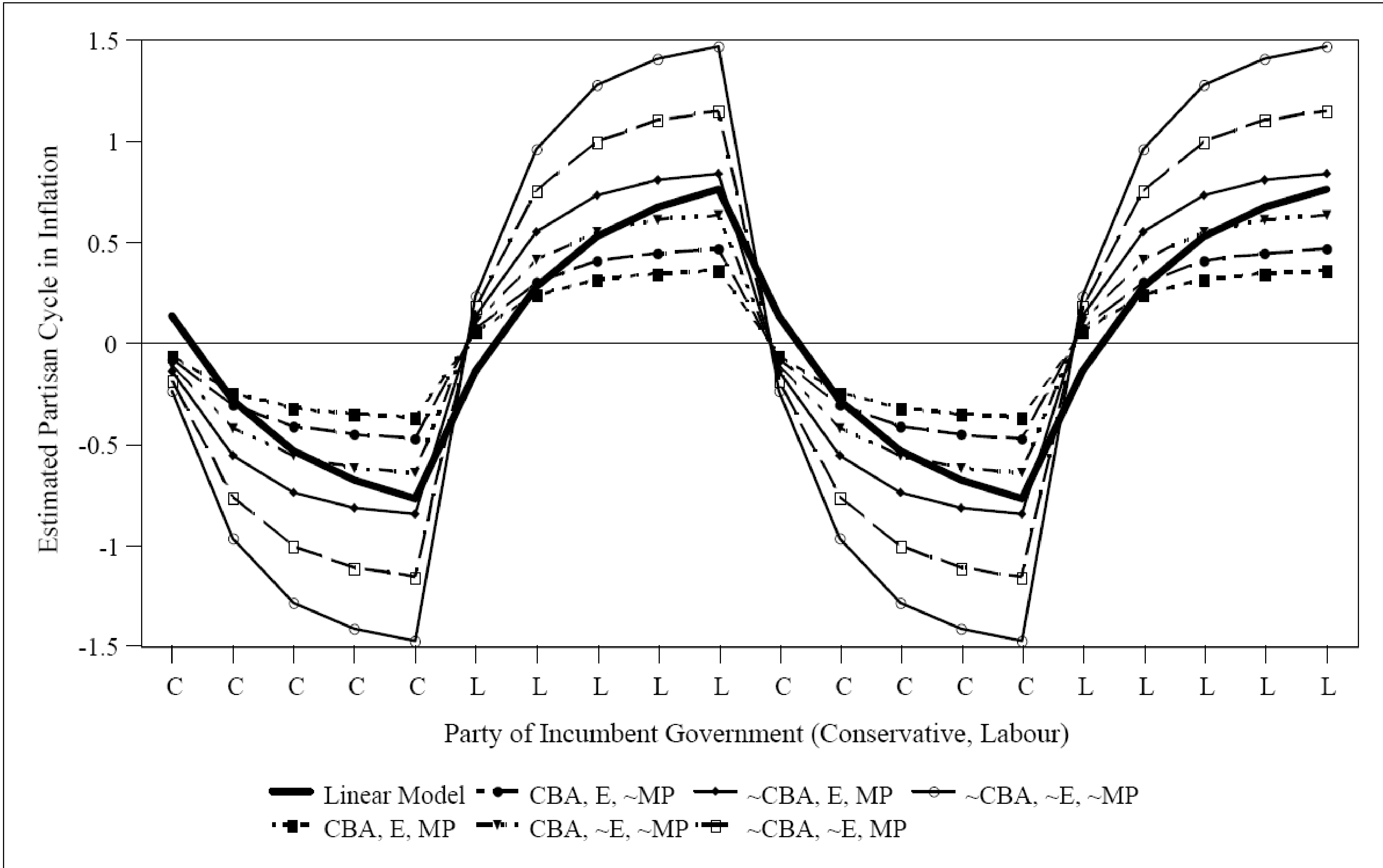


Figure 1: Estimated Partisan Cycles in the Linear & Theoretically Informed Models at High & Low *CBA*, *E*, & *MP*

Context-Conditional Inflation Effects of a Single-Currency Exchange-Rate Peg (to average currency)

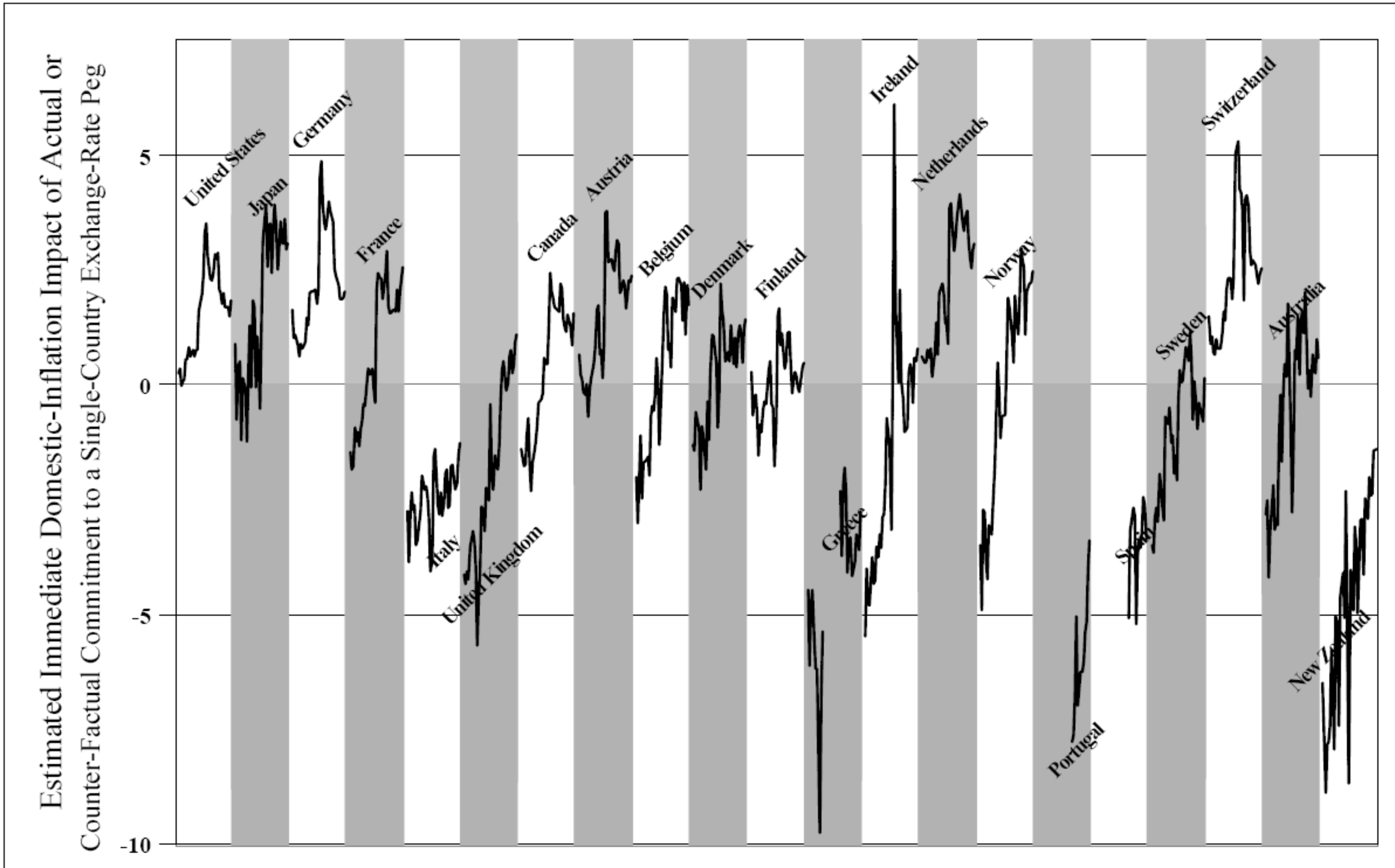


Figure 2: Estimated Domestic-Inflation Effect of Actual or Counter-Factual *SP* in 21 Countries, 1957-90. Estimates plotted for dINF/dSP at the values of all other variables in the equation actually occurring in that country-year. For counter-factual pegs, peg country assumed to have OECD-average inflation that year. Shading separates countries and extends from 1955 to 1990 in each country, left to right.

Context-Conditional Anti-Inflation Effects of Central Bank Independence

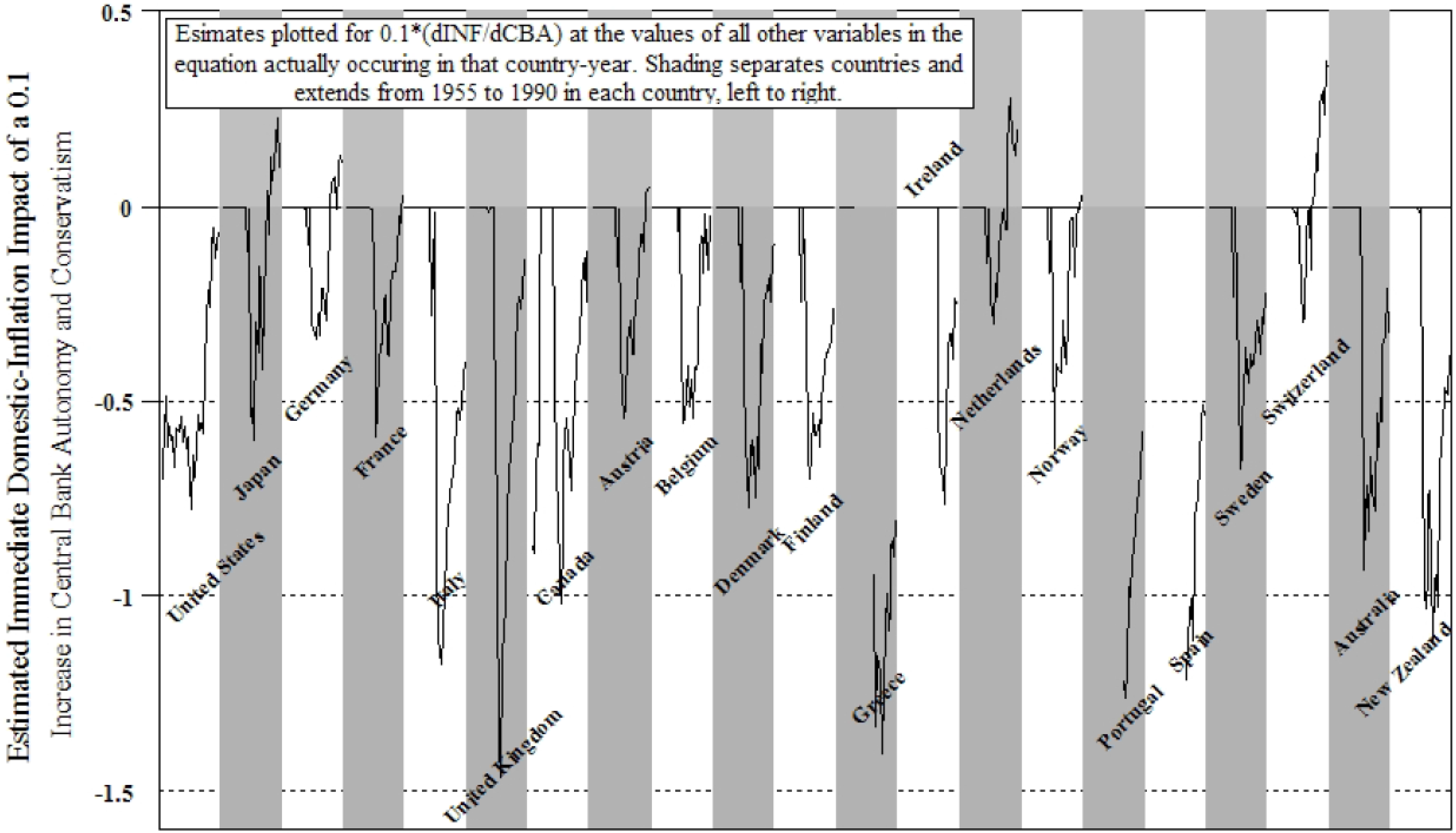


Figure 9: Estimated Immediate Domestic-Inflation Impact of 0.1 Increase in CBA in 21 Countries, 1957-90

(Temporal) Dynamics also suffice to make causal-effect inference insufficient for causal-response estimation

- **Another distinction worth elaborating:**
  - **Identifying that a causal effect exists (causal inference)**

    vs.

    estimating a causal response (causal estimation).
    - **Experiments tend be ideal for the former; Not necessarily so great at the latter.**
- The socio-politico-economic reality that we study is dynamic & interdependent. Approaches to empirical analysis that emphasize nonparametric causal-inference are static & insulated.

Temporal dynamics, for instance, mean a world like this…

B. Interpreting Dynamic Specifications:

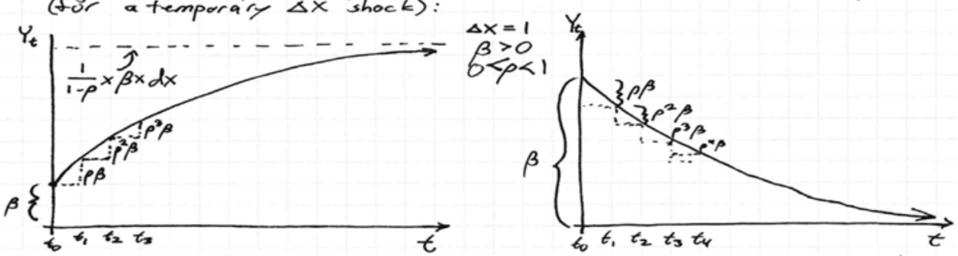
$$Y_t = \omega Y_{t-1} + X_t\beta + \varepsilon_t$$ with ε_t iid or autocorrelated doesn't matter here.

$$\frac{\partial Y_t}{\partial X_t} = \beta$$ same as always

a) This is *literally* response of y in period t to a unit increase in x in period t. The model is dynamic, so there is more to the response than just this one-period effect.

b) Next period, t+1, y_t is larger (smaller) by $\beta\times\Delta x_t$ units, which means y_{t+1} will be larger (smaller) by $\rho\times\Delta y_t = \rho\times\beta\times\Delta x_t$, in addition to the $\beta\times\Delta x_t$ from this period, and…

- After some impetus, ΔX, the effect persists into the future to a degree/extent ρ (for permanent ΔX shock) or fades at rate $(1-\rho)$ (for a temporary ΔX shock):



…but *causal inference* gives a dY/dX=a (single, scalar) number.

- Consider the many well-designed causal-inference studies of turnout effects of motor-voter laws & the like, for example. Typically small-to-modest effects found.
- Consider also the evidence that voting is a long-term acquired habit, the aggregate implication of which is that voter participation evolves dynamically like this.
- Response of voter participation isn't one snapshot-in-time scalar, it's a vector over time.

Context-Conditional Temporal Dynamics
The Multiple Effects of Multiple Policymakers

- Theory:
  - The ***multiple effects operate through different aspects of policymaker fragmentation, polarization, & partisanship***:
    - **Veto-Actor Effects**: raw number of parties (fragmentation) & ideological ranges (polarization)
    - **Common-Pool Effects**: effective numbers (fragmentation)
    - **Delegation-Bargaining Effects**: power-wtd mean ideologies (partisanship)
  - ***Different ways these distinct effects manifest in policy***:
    - **V-A** (primarily) work to slow policy-adjustment (*delay stabilization*);
    - **C-P** induces over-draw from common resources (incl. from future as in debt); under-invest in common properties (*incumbents less electioneering*), log-proportionately
    - **D-B** induces convex-combinatorial (compromise) policies, incl. *greater left-activist/right-conservative Keynesian-countercyclical/conservative pro-cyclical*, in proportion to degree left/right controls pol.
- **Empirical Model of this Theoretical Synthesis:**
  - *Absolute number* (frag.) VAs & their *ideological range* (polar.) modify policy-adjust rates
  - *(log) Effective number* pol-mkrs & *s.d. of their ideology* (wtd measures) gauge extent of C-P problem in *electioneering* (+debt-lvl effect?)
  - Some barg. process among partisan pol-mkrs (e.g., Nash ⇒ wtd-influence) determines combo reflected in net policy responsiveness to macro (º K-activism)
  - ⇒

$$\begin{aligned} D_{it} = \alpha_i &+ \left(1+\rho_n NoP_{it} + \rho_{ar} ARwiG_{it}\right) \times \left(\rho_1 D_{i,t-1} + \rho_2 D_{i,t-2} + \rho_3 D_{i,t-3}\right) \\ &+ \left(\beta_{\Delta Y} \Delta Y_{i,t} + \beta_{\Delta U} \Delta U_{i,t} + \beta_{\Delta P} \Delta P_{i,t}\right) \times \left(1+\beta_{cg} CoG_{it}\right) \\ &+ \left(\gamma_{e1} E_{it} + \gamma_{e2} E_{i,t-1}\right) \times \left(1+\gamma_{en} ENoP_{it} + \gamma_{sd} SDwiG_{it}\right) + \mathbf{x}'_{it}\boldsymbol{\eta} + \mathbf{z}'_{it}\boldsymbol{\omega} + \varepsilon_{it} \end{aligned}$$

Empirical Model Specification & Data

$$D_{it} = \alpha_i + \left(1+\rho_n NoP_{it} + \rho_{ar} ARwiG_{it}\right)\times\left(\rho_1 D_{i,t-1} + \rho_2 D_{i,t-2} + \rho_3 D_{i,t-3}\right) + \mathbf{x}'_{it}\boldsymbol{\eta} + \mathbf{z}'_{it}\boldsymbol{\omega} + \varepsilon_{it}$$
$$+\left(\beta_{\Delta Y}\Delta Y_{i,t} + \beta_{\Delta U}\Delta U_{i,t} + \beta_{\Delta P}\Delta P_{i,t}\right)\times\left(1+\beta_{cg} CoG_{it}\right) + \left(\gamma_{e1}E_{it} + \gamma_{e2}E_{i,t-1}\right)\times\left(1+\gamma_{en} ENoP_{it} + \gamma_{sd} SDwiG_{it}\right)$$

- D_{it} = Debt (%GDP);
- *NoP* & *ARwiG* = raw *N*um *o*f *P*rtys in *G*ovt & *A*bs *R*ange *w/i G*ovt:
  - VA conception, so modify dynamics. Expect ρ_n & ρ_{ar} >0.
  - By thry & for efficiency: modify all lag dynamics same.
- *CoG* (govt center, left to right, 0-10):
  - Modifies response to macroecon (equally, by thry & for eff'cy) : β_{cg}<0.
  - Macroec: ΔY = real GDP growth; ΔU = Δ unemp rate; ΔP = infl rate.
- **x'η** = controls: set pol-econ cond's response to which not partisan-differentiated or gov comm-pool: (e.g., E(real-int)-E(real-grow), *ToT*)
- *ENoP* & *SDwiG* = *E*ffective *N*um *o*f *P*rtys in govt & *S*td *D*ev *w/i G*ovt:
  - Frag & Polar by *wtd-influence* concept. CP lvl-effects modify (at same rate) electioneering, E_t, pre-elect-yr, & E_{t-1}, post-elect-yr.: γ_{en} & γ_{sd}<0.
- **z'ω** = set of constituent terms in the interactions:
  - *ENoP*, *SDwiG* <u>*may*</u> have positive coeff's by CP-effect on lvl debt, but issue is *temporal fract* > curr. govt *fract*. Thry o/w says omit.

- Pace Brambor et al. ('06), but joint-significance of multiple-policymaker conditioning effects (γ_{en}, γ_{sd}, ρ_n, ρ_{ar}, β_{cg}) overwhelmingly rejects excluding (p≈.001), whereas joint-sig coeff's on constit. terms, **z**, clearly fails reject (p≈.602) exclusion. (Almost) All theory says should be zero, so…

		Coeff.	**Std. Err.**	***t*-Stat.**	**Pr(*T*>|*t*|)**
Lagged Dependent Variables	D_{t-1}	1.207	0.060	20.290	0.000
	D_{t-2}	-0.158	0.085	-1.851	0.065
	D_{t-3}	-0.117	0.045	-2.577	0.010
$\boldsymbol{\rho_n}$ (*veto-actor effect: fractionalization*)		**0.011**	**0.005**	**2.369**	**0.018**
$\boldsymbol{\rho_{ar}}$ (*veto-actor effect: polarization*)		**-0.002**	**0.004**	**-0.437**	**0.662**
Macroeconomic Conditions	ΔY	-0.375	0.087	-4.332	0.000
	ΔU	1.095	0.286	3.829	0.000
	ΔP	-0.207	0.053	-3.889	0.000
$\boldsymbol{\beta_{cg}}$ (*partisan-compromise bargaining*)		**-0.051**	**0.020**	**-2.484**	**0.013**
Controls	x_1 (*open*)	16.128	5.314	3.035	0.002
	x_2 (*ToT*)	0.414	1.728	0.239	0.811
	x_3 (*open·ToT*)	-10.780	5.194	-2.076	0.038
	x_4 (*dxrig*)	-0.038	0.066	-0.578	0.563
	x_5 (*oy*)	1.898	1.100	1.724	0.085
Pre- and Post-Electoral Indicators	E_t	0.475	0.420	1.133	0.258
	E_{t-1}	1.146	0.562	2.040	0.042
$\boldsymbol{\gamma_{en}}$ (*common-pool effect: fractionalization*)		**-0.570**	**0.209**	**-2.727**	**0.007**
$\boldsymbol{\gamma_{sd}}$ (*common-pool effect: polarization*)		**0.881**	**0.586**	**1.503**	**0.133**
		Summary Statistics			
N (Deg. Free)		735 (696)		s_e^2	2.522
$\mathbf{R^2}$ ($\overline{\mathbf{R}}^2$)		0.991 (0.990)		**DW-Stat.**	2.099

Veto-Actor Effects: Estimates of Policy-Adjustment Rate						
Adjustment Rates	***NoP*=1**	***NoP*=2**	***NoP*=3**	***NoP*=4**	***NoP*=5**	***NoP*=6**
Lag Coefficient[a]	0.943	0.952	0.960	0.969	0.978	0.986
Policy-Adjust/Yr[b]	0.057	0.048	0.040	0.031	0.022	0.014
Long-Run Mult.[c]	17.498	20.639	25.154	32.200	44.727	73.208
½-Life[d]	11.778	13.956	17.087	21.971	30.654	50.397
90%-Life[e]	39.127	46.362	56.761	72.985	101.832	167.415

Bargaining Effects: Estimates of Keynesian Fiscal Responsiveness						
	Mean Econ. Performance -2 std. dev.	*Mean Econ. Performance -1 std. dev.*	*Mean Economic Performance*	*Mean Econ. Performance +1 std. dev.*	*Mean Econ. Performance +2 std. dev.*	
Growth	-2.354	0.454	3.261	6.069	8.877	
d(UE)	1.915	1.034	0.153	-0.728	-1.608	
Infl	-3.593	1.230	6.054	10.877	15.701	
CoG	*E(D|Econ)*[f]	*E(D|Econ)*	*E(D|Econ)*	*E(D|Econ)*	*E(D|Econ)*	*Fiscal-Cycle Magnitude*[g]
3.0	3.157	0.599	-1.959	-4.516	-7.074	10.231
4.2	2.930	0.556	-1.818	-4.192	-6.566	9.496
5.4	2.703	0.513	-1.677	-3.867	-6.058	8.761
6.6	2.476	0.470	-1.536	-3.543	-5.549	8.026
7.8	2.250	0.427	-1.396	-3.218	-5.041	7.291
9.0	2.023	0.384	-1.255	-2.894	-4.533	6.555

Collective-Action/Common-Pool Effects: Estimates of Electoral Debt-Cycle Magnitude					
	***ENoP*=1**	***ENoP*=2**	***ENoP*=3**	***ENoP*=4**	***ENoP*=5**
Electoral-Cycle Magnitude[h]	1.07410	0.86454	0.65497	0.44541	0.23585

Some Dynamic Effect Estimates

(From a Different, but Similar Political Economy of Public Debt Project)

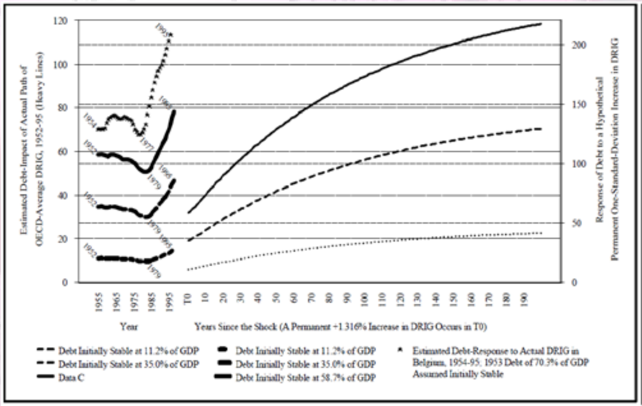


Figure 8: Estimated debt-responses to hypothetical, permanent, 1-standard-deviation adverse DRIG shocks, to the actual OECD-average DRIG path, and to the actual DRIG path in Belgium.

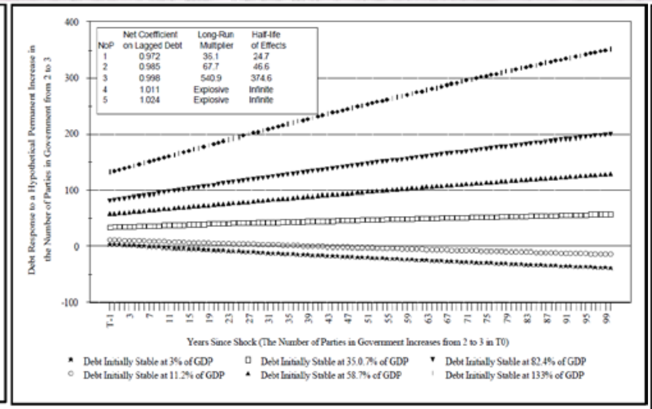


Figure 10: Estimated Debt-Response to a Hypothetical Permanent Increase in NoP from 2 to 3 at Various Initially Stable Outstanding-Debt Levels

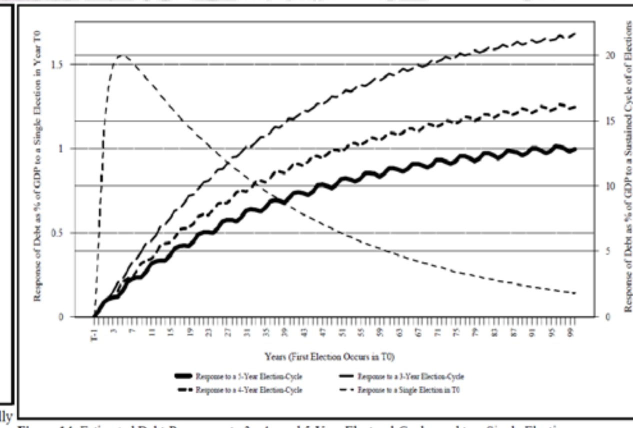


Figure 14: Estimated Debt-Responses to 3-, 4-, and 5-Year Electoral-Cycles and to a Single Election

(From a Different, but Similar Political Economy of Public Transfers Project)

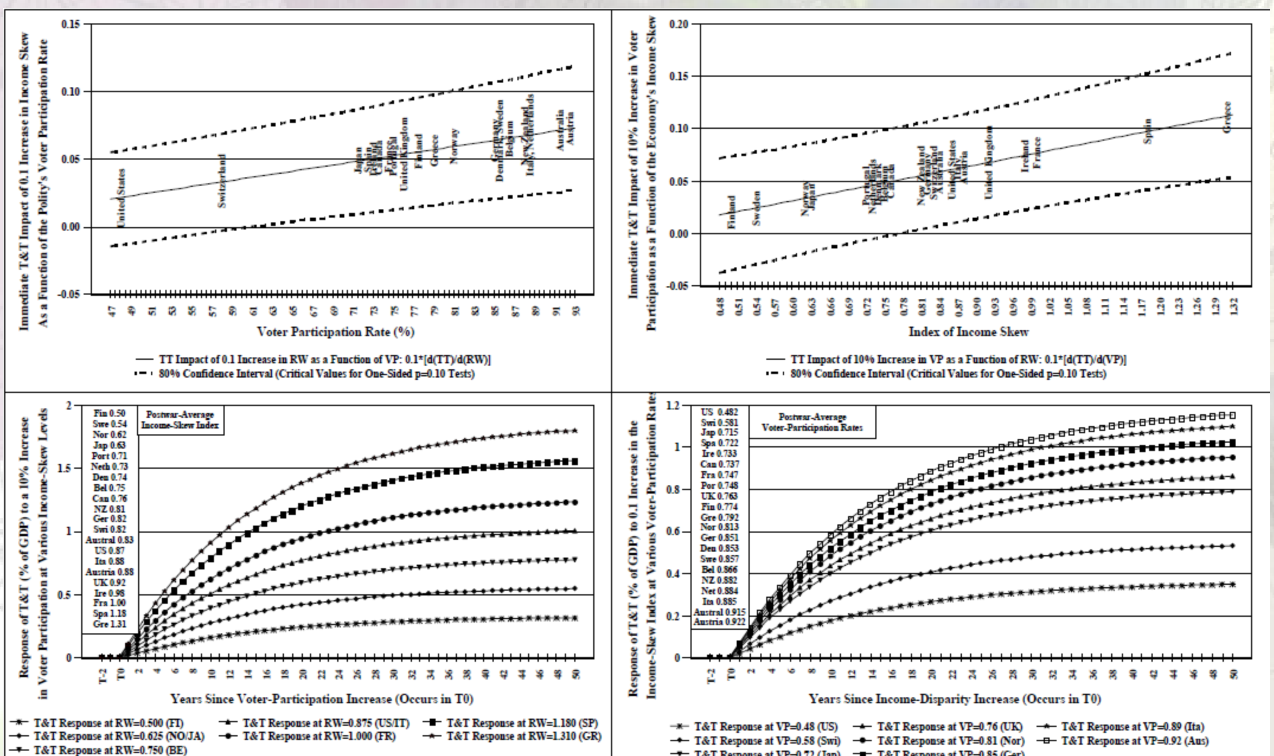


Figure 4: Estimated Immediate and Longer-term T&T Response to Increases in Income Skew as a Function of Voter Participation and to Increases in Voter Participation as a Function of Income Skew

Return to those 'dynamic' estimates of left-govt interest-costs: those were just static snapshots; being 'non-parametric', offer no clue to t+13 etc, & also susceptible to small-sample peculiarities in monthly events.

Spatial (Cross-Unit) Interdependence imply Spatial Dynamics, and are a form of Simultaneity ($y_1 \Leftrightarrow y_2$), & so also suffice to make causal-effect inference insufficient for causal-response estimation

Systems with cross-unit interdependence (contagion), or with simultaneous causality, y⇔x, like more or less all of social science, mean a world like this:

$$y = \alpha_0 + \alpha_1 x + \alpha_2 z_y + \varepsilon_y$$
$$x = \beta_0 + \beta_1 y + \beta_2 z_x + \varepsilon_x$$

, which imply:

$$y = \alpha_0 + \alpha_1 \underbrace{\left(\beta_0 + \beta_1 y + \beta_2 z_x + \varepsilon_x\right)}_{x} + \alpha_2 z_y + \varepsilon_y$$

$$y - \alpha_1 \beta_1 y = \alpha_0 + \alpha_1 \underbrace{\left(\beta_0 + \beta_2 z_x + \varepsilon_x\right)}_{\text{exogenous part of } x} + \alpha_2 z_y + \varepsilon_y$$

$$y(1 - \alpha_1 \beta_1) = \alpha_0 + \alpha_1 \underbrace{\left(\beta_0 + \beta_2 z_x + \varepsilon_x\right)}_{\text{exogenous part of } x} + \alpha_2 z_y + \varepsilon_y$$

$$y = (1 - \alpha_1 \beta_1)^{-1} \left[\alpha_0 + \alpha_1 \left(\beta_0 + \beta_2 z_x + \varepsilon_x\right) + \alpha_2 z_y + \varepsilon_y\right]$$

, meaning: $\frac{dy}{dx} = \frac{\alpha_1}{1 - \alpha_1 \beta_1}$ **and not** $\frac{dy}{dx} = \alpha_1$.

Experiments work to identify *existence* of causal effects by preventing *estimation* of *responses* in the actual simultaneous system of interest. They estimate causal parameters, not causal effects.

- The experimentally or quasi-experimentally derived estimates of causal ‘effects’ of *X* in cases where *X*⇔*Y* in the context we care about (i.e., not in the lab) will be of the impulses, i.e. of the parameters, β, and not of the response, the *effect, d*Y/*d*X.
- In quasi-experimental contexts, may very well be biased estimates of β as well, simultaneity, including spatial-simultaneity, being sources of “interference” so Control likely contaminated.

Given Ubiquitous Endogeneity of Social Phenomena, Must Estimate Systems Models

- This discussion regards causal-parameter estimation (which is what exper. or well-designed non-parametric causal-inference strategy will uncover also, $\frac{\partial y}{\partial x}$ and not $\frac{dy}{dx}$).
  - Notice, btw, that can say quite a bit about the simultaneity bias in this case. Simply not true that it's a unique advantage of design-based strategies that can bound these sorts of biases (or ones from other *confounds*)

c) *Mutual Causality*, $y \Leftrightarrow x$, & so single-equation model is incomplete (violating Assumpt 1), implying Covariance Regressor w/ Residual (violating Assumpt 4):

$$\left.\begin{array}{l} y=\beta x+\gamma z+\varepsilon_y \\ x=\theta y+\lambda w+\varepsilon_x \end{array}\right\} \Rightarrow \left\{\begin{array}{l} Cov(x,\varepsilon_y)=Cov(\varepsilon_y,\theta y+\lambda w+\varepsilon_x)=Cov(\varepsilon_y,\theta y) \\ \qquad =Cov(\varepsilon_y,\theta(\beta x+\gamma z+\varepsilon_y))=Cov(\varepsilon_y,\theta\varepsilon_y)=\theta Var(\varepsilon_y) \\ Cov(y,\varepsilon_x)=Cov(\varepsilon_x,\beta x+\gamma z+\varepsilon_y)=Cov(\varepsilon_x,\beta x) \\ \qquad =Cov(\varepsilon_x,\beta\varepsilon_x)=\beta Var(\varepsilon_y) \end{array}\right.$$

$$\begin{array}{l} \mathbf{y}=\beta\mathbf{x}+\gamma\mathbf{z}+\boldsymbol{\varepsilon}_y \\ \mathbf{x}=\theta\mathbf{y}+\lambda\mathbf{w}+\boldsymbol{\varepsilon}_x \end{array}\text{, but we estimate instead just } \mathbf{y}=b\mathbf{x}+g\mathbf{z}+\mathbf{e}_y :$$

$$\Rightarrow \begin{bmatrix} b \\ g \end{bmatrix}=\left\{[\mathbf{x}\ \ \mathbf{z}]'[\mathbf{x}\ \ \mathbf{z}]\right\}^{-1}[\mathbf{x}\ \ \mathbf{z}]'\left[\beta\mathbf{x}+\gamma\mathbf{z}+\boldsymbol{\varepsilon}_y\right]$$

$$=\left\{[\mathbf{x}\ \ \mathbf{z}]'[\mathbf{x}\ \ \mathbf{z}]\right\}^{-1}[\mathbf{x}\ \ \mathbf{z}]'[\beta\mathbf{x}+\gamma\mathbf{z}]+\left\{[\mathbf{x}\ \ \mathbf{z}]'[\mathbf{x}\ \ \mathbf{z}]\right\}^{-1}[\mathbf{x}\ \ \mathbf{z}]'\boldsymbol{\varepsilon}_y$$

$$=\begin{bmatrix} \beta \\ \gamma \end{bmatrix}+\left\{[\mathbf{x}\ \ \mathbf{z}]'[\mathbf{x}\ \ \mathbf{z}]\right\}^{-1}\begin{bmatrix} \mathbf{x}'\boldsymbol{\varepsilon}_y \\ \mathbf{z}'\boldsymbol{\varepsilon}_y \end{bmatrix}=\begin{bmatrix} \beta \\ \gamma \end{bmatrix}+\frac{1}{\left|[\mathbf{x}\ \ \mathbf{z}]'[\mathbf{x}\ \ \mathbf{z}]\right|}\begin{bmatrix} \mathbf{z}'\mathbf{z} & -\mathbf{x}'\mathbf{z} \\ -\mathbf{z}'\mathbf{x} & \mathbf{x}'\mathbf{x} \end{bmatrix}\begin{bmatrix} \mathbf{x}'\boldsymbol{\varepsilon}_y \\ \underbrace{\mathbf{z}'\boldsymbol{\varepsilon}_y}_{=0} \end{bmatrix}$$

$$=\frac{1}{\left|[\mathbf{x}\ \ \mathbf{z}]'[\mathbf{x}\ \ \mathbf{z}]\right|}\begin{bmatrix} C(\mathbf{x},\boldsymbol{\varepsilon}_y)\times V(\mathbf{z}) \\ -C(\mathbf{x},\boldsymbol{\varepsilon}_y)\times C(\mathbf{z},\mathbf{x}) \end{bmatrix}=\begin{bmatrix} \beta \\ \gamma \end{bmatrix}+\frac{1}{|\underbrace{\mathbf{R}'\mathbf{R}}_{\substack{\text{Regresor} \\ \text{V-Cov}}}|}\begin{bmatrix} \theta\sigma^2_{\varepsilon_y}\times V(\mathbf{z}) \\ -\theta\sigma^2_{\varepsilon_y}\times C(\mathbf{z},\mathbf{x}) \end{bmatrix}$$

Simultaneity bias generally has sign of & is proportionate in magnitude to omitted causal arrow, &, as usual in multiple regression, it induces biases in other regressors, generally of smaller magnitude (b/c Var gen'ly > |Cov|), in opposite direction (same direction if Cov<0), and magnitudes of induced biases distributed across regressors in proportion to their correlation w/ endogenous regressor (OVB intuition).

Given Ubiquitous Endogeneity of Social Phenomena, Must Estimate Systems Models

E. Generalize for (linear) system of M equations with K regressors:

1. Recall that for one observation on this system, we could write:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_m \end{bmatrix}_i' \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mm} \end{bmatrix} + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix}_i' \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{k1} & \beta_{k2} & \cdots & \beta_{km} \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \vdots \\ \varepsilon_m \end{bmatrix}_i$$

$$\underbrace{\mathbf{y}_i'}_{1\times M} \underbrace{\mathbf{\Gamma}}_{M\times M} + \underbrace{\mathbf{x}_i'}_{1\times K} \underbrace{\mathbf{B}}_{K\times M} = \underbrace{\boldsymbol{\varepsilon}_i}_{1\times M}$$

2. Normalizing γ_{mm} coeff's on y_m (diagonals of Γ) to 1 in above (so explain $1\times y$ rather some other $1\times y$) makes these diagonals γ_{mm} of Γ below $= 0$:

$$\underbrace{\mathbf{Y}}_{N\times M} = \underbrace{\mathbf{Y}}_{N\times M} \underbrace{\mathbf{\Gamma}}_{M\times M} + \underbrace{\mathbf{X}}_{N\times K} \underbrace{\mathbf{B}}_{K\times M} + \underbrace{\mathbf{E}}_{N\times M}$$

$$\Rightarrow \mathbf{Y} - \mathbf{Y\Gamma} = \mathbf{Y}(\mathbf{I} - \mathbf{\Gamma}) = \mathbf{XB} + \mathbf{E}$$

$$\Rightarrow \mathbf{Y} = (\mathbf{XB} + \mathbf{E})(\mathbf{I} - \mathbf{\Gamma})^{-1}$$

Can also generalize the simultaneity-bias formula from previous slide thusly:

3. Can also use this to generalize formula simultaneity bias to sys M eqtns:

$$\underbrace{\mathbf{Y}}_{N\times M} = \underbrace{\mathbf{Y}}_{N\times M} \underbrace{\mathbf{\Gamma}}_{M\times M} + \underbrace{\mathbf{X}}_{N\times K} \underbrace{\mathbf{B}}_{K\times M} + \underbrace{\mathbf{E}}_{N\times M}$$

$\Rightarrow$ Reg $\mathbf{Y}$ on $\tilde{\mathbf{Y}}$ and $\mathbf{X}$, where $\tilde{\mathbf{y}}_m$ normalized to $\mathbf{0}$ in m^{th} equation

$$\Rightarrow [\hat{\mathbf{\Gamma}} \quad \hat{\mathbf{B}}] = \{[\tilde{\mathbf{Y}} \quad \mathbf{X}]'[\tilde{\mathbf{Y}} \quad \mathbf{X}]\}^{-1}[\tilde{\mathbf{Y}} \quad \mathbf{X}]'[\mathbf{Y\Gamma} + \mathbf{XB} + \mathbf{E}]$$

$$\Rightarrow [\hat{\mathbf{\Gamma}} \quad \hat{\mathbf{B}}] = [\mathbf{\Gamma} \quad \mathbf{B}] + \{[\tilde{\mathbf{Y}} \quad \mathbf{X}]'[\tilde{\mathbf{Y}} \quad \mathbf{X}]\}^{-1}[\tilde{\mathbf{Y}} \quad \mathbf{X}]'\mathbf{E}$$

BIAS $= \mathbf{A'E}$, which analogously to multivariate measurement-error case, means bias concentrates in "most-endogenous" & induces biases in OVB fashion.

- **Y** here is matrix of endogenous variables data, which were *y* & *x* in previous slide; **X** here is another set of exogenous variables **Z**, *z* & *w* in prev. (sorry).
- An exogenous shock to X from before can only be expressed in ε_x, but once it is, we see its effect, i.e. the full causal response, is given by $(\mathbf{I}-\mathbf{\Gamma})^{-1}\times d\boldsymbol{\varepsilon}\times\mathbf{B}$.

Simulation Demonstration of Inadequacy of Causal Inference to Causal Estimation

SIMULTANEITY BIAS, 2x2 case:

$$\text{TRUTH: } \begin{array}{l} y = .5x + z + \varepsilon_y \\ x = .5y + w + \varepsilon_x \end{array}$$

```
drop x y z
gen err_y=rnormal()
gen err_x=rnormal()
gen z=rnormal()
gen w=rnormal()
```

$$\text{TRUTH: } \begin{array}{l} y = .5x + z + \varepsilon_y \\ x = .5y + w + \varepsilon_x \end{array} \Rightarrow \begin{array}{l} y = .5(.5y + w + \varepsilon_x) + z + \varepsilon_y = (1-.25)^{-1}\left[.5(w+\varepsilon_x) + z + \varepsilon_y\right] \\ x = .5(.5x + z + \varepsilon_y) + w + \varepsilon_x = (1-.25)^{-1}\left[.5(z+\varepsilon_y) + w + \varepsilon_x\right] \end{array}$$

```
gen y=(1/(1-.25))*(.5*w+.5*err_x+z+err_y)
gen x=(1/(1-.25))*(.5*z+.5*err_y+w+err_x)
reg y x z
reg x y w
```

- In this case, for example, $\frac{\partial y}{\partial x} = .5$, but $\frac{dy}{dx} = .67$ (i.e., causal-parameter estimation fails to give the causal effect, understood causal response of **Y**, d**Y**, to d**X**).

Of course, social phenomena are dynamic systems of endogenous equations, so…

- Vector Autoregression should get attention as potentially "useful empirical simplification" also:

- In the simple two-variable case, the *structural version* of the *first-order* VAR model is

$$y_t = b_{10} - b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_{yt}$$
$$z_t = b_{20} - b_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_{zt}$$

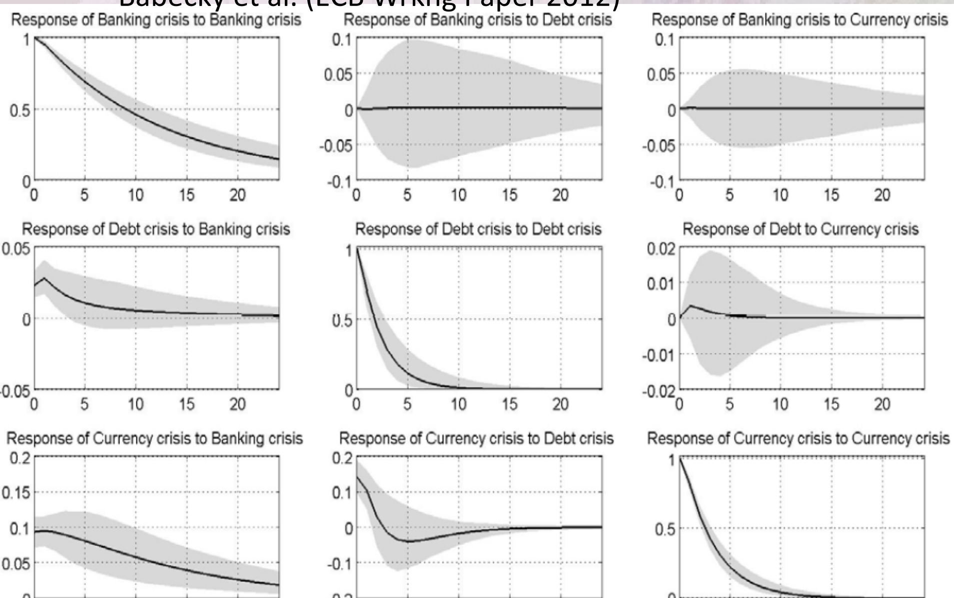
where y_t and z_t are assumed stationary and ε_{yt} and ε_{zt}, the *structural disturbances*, are uncorrelated white-noise disturbances with standard deviations σ_y and σ_z respectively.

- Note that we can rewrite this system as

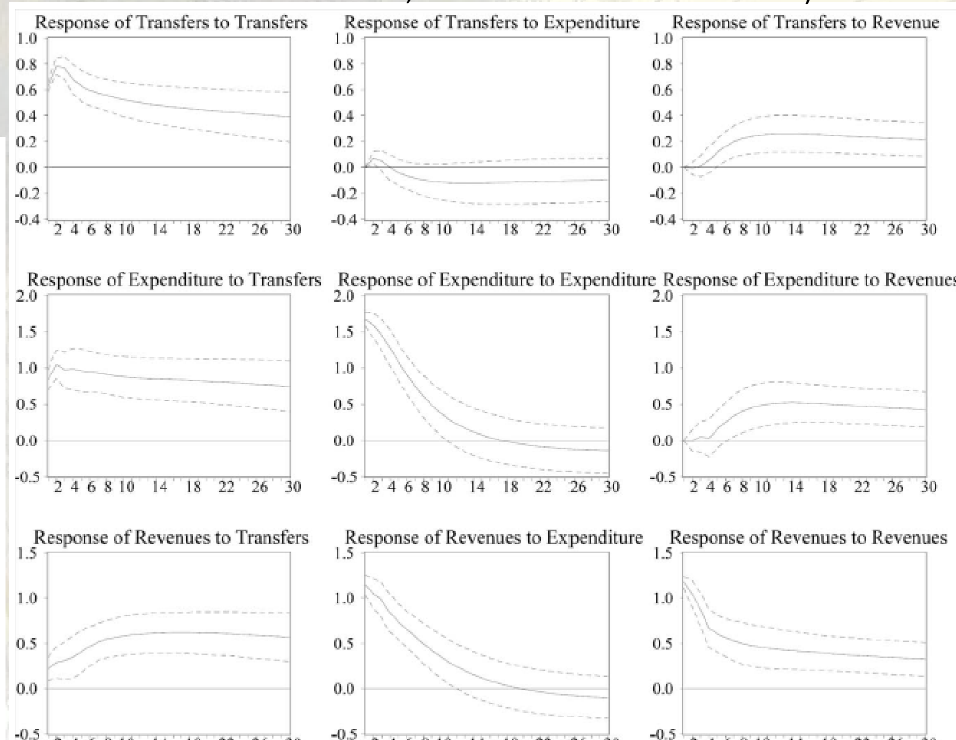
$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}\begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

And these examples of (1) Banking, Debt, & Currency Crises, and (2) Public Transfers, Total Expenditures, & Revenues may also illustrate how plausible external validity but questionable internal validity still interesting & useful btw…

Babecky et al. (ECB Wrkng Paper 2012)



Franzese, *Macroeconomic Policies* 2002)



Interpreting Spatiotemporal (=Dynamic Interdependent) Effects

- The Model: $\mathbf{y}_t = \rho \mathbf{W}_n \mathbf{y}_t + \phi \mathbf{I}_n \mathbf{y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t$
  - Convenient, for interpretation, to write model this way too:

$$\mathbf{y}_t = \rho \mathbf{W}_n \mathbf{y}_t + \phi \mathbf{y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t$$

  - Coefficients, $\boldsymbol{\beta}_x$ are the pre-spatial, pre-temporal—and wholly unobservable!—impulse from some *x* to y.
- Spatiotemporal <u>*Effects*</u>:
  - Post-spatial, pre-temporal "instantaneous effect" of *d*x:

$$d\left\{\left[\mathbf{I}_N - \rho \mathbf{W}_N\right]^{-1}\left(\mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right)\right\} \Big/ dx_i \text{ for some (set of) } i; \text{ i.e., } \left[\mathbf{I}_N - \rho \mathbf{W}_N\right]^{-1} d\mathbf{x}_k^i \boldsymbol{\beta}$$

    - Spatiotemp Response Paths, use this:

$$\mathbf{y}_t = \left[\mathbf{I}_N - \rho \mathbf{W}_N\right]^{-1}\left\{\phi \mathbf{y}_{t-1} + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right\}$$

    - LR Multiplier & LR-SS, use this:

$$\mathbf{y}_t = \rho \mathbf{W}_N \mathbf{y}_t + \phi \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t = \left(\rho \mathbf{W}_N + \phi \mathbf{I}_N\right) \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t$$
$$= \left[\mathbf{I}_N - \rho \mathbf{W}_N - \phi \mathbf{I}_N\right]^{-1}\left(\mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t\right)$$

Maps of Response-Estimates (F&H *EUP*)

Figure 1. Short-run Spatial Effects of a Positive One-unit Shock to German LMT Expenditures

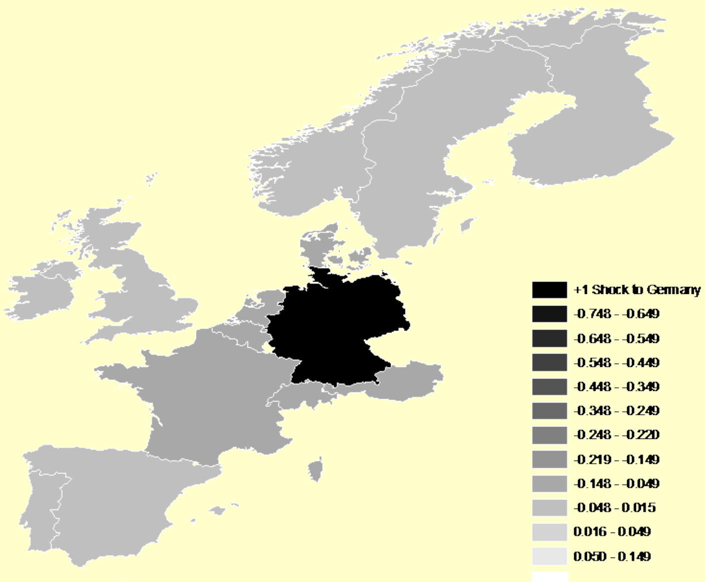
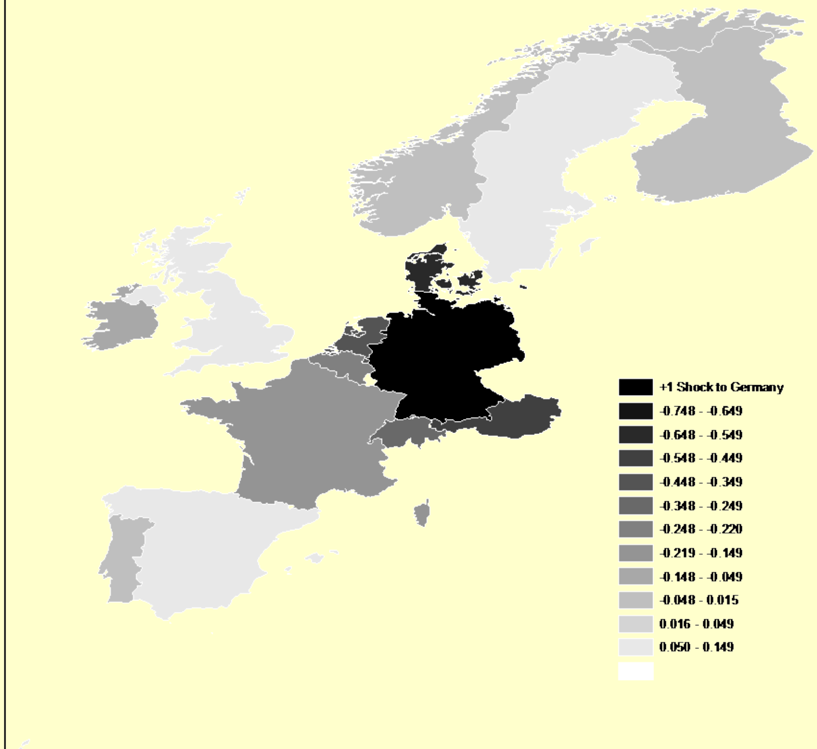


Figure 2. Steady-state Spatial Effects of a Positive One-unit Shock to German LMT Expenditures



Actually, can demonstrate that some manifestations of Spatiotemporal Interdependence make even NHR-Based Causal-Inference (well, specifically: Matching) Problematic

$$\text{DGP: } \mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{x} + \mathbf{T} + \boldsymbol{\varepsilon} \text{ where } w_{ij}^{*} = -3 - a_1\left|x_i - x_j\right| + v_{ij},$$

$$w_{ij} = 1,0 \text{ if } w_{it}^{*} > 0, \leq 0, \text{ with } x \sim N(0,1),\ v \sim Logistic\ (0,1),$$

$$T_i = a_2 x_i + u_{ij}, \text{ with } u \sim Logistic\ (0,1), \text{ and } \varepsilon \sim N(0,1).$$

- **Experimental Cases:**
  - 1 Exogenous Network (a_1=0), Orthogonal Treatment (a_2=0), No Spillover (ρ=0).
  - 2 Exogenous Network (a_1=0), Orthogonal Treatment (a_2=0), Spillovers (ρ=.5).
  - 3 Endogenous Network (a_1=1), Orthogonal Treatment (a_2=0), No Spillover (ρ=0).
  - 4 Endogenous Network (a_1=1), Orthogonal Treatment (a_2=0), Spillovers (ρ=.5).
  - 5 Exogenous Network (a_1=0), Treatment Not Orthogonal (a_2=1), No Spillover (ρ=0).
  - 6 Exogenous Network (a_1=0), Treatment Not Orthogonal (a_2=1), Spillovers (ρ=.5).
  - 7 Endogenous Network (a_1=1), Treatment Not Orthogonal (a_2=1), No Spillover (ρ=0).
  - 8 Endogenous Network (a_1=1), Treatment Not Orthogonal (a_2=1), Spillovers (ρ=.5).
- **Estimators:**
  - Naïve Regression: Y on X and T, OLS.
  - Matching: Nearest Neighbor using propensity scores by logit: $T_i = \alpha + a_2 x_i + u_{ij}$
  - Spatial Autoregression: Y on X, WY, and T, by spatial-ML.
  - Spatially Lagged Treatment: Y on X, T, and WT, OLS.

Some Quick MC's to Illustrate Some Challenges and Estimation-Strategy Effectiveness

Case 1: Exogenous Network, Orthogonal Treatment, No Outcome Contagion				
	Naïve Regression	Matching	Outcome Contagion	Treatment Diffusion
Coeff ($\beta = 1$)	1.003	1.033	0.999	1.003
Std	0.201	0.272	0.201	0.201
RMSE	0.201	0.274	0.201	0.201
Coeff ($\rho = 0$)			-0.035	0.017
Std			0.189	0.375
RMSE			0.192	0.376

Case 2: Exogenous Network, Orthogonal Treatment, with Outcome Contagion				
	Naïve Regression	Matching	Outcome Contagion	Treatment Diffusion
Coeff ($\beta = 1$)	1.05	1.019	0.999	1.014
Std	0.213	0.279	0.204	0.208
RMSE	0.219	0.279	0.204	0.209
Coeff($\rho = .5$)			0.428	0.411
Std			0.171	0.478
RMSE			0.186	0.486

Case 3: Endogenous Network, Orthogonal Treatment, No Outcome Contagion				
	Naïve Regression	Matching	Outcome Contagion	Treatment Diffusion
Coeff ($\beta = 1$)	1.015	1.031	1.015	1.013
Std	0.218	0.293	0.219	0.219
RMSE	0.219	0.295	0.219	0.22
Coeff($\rho = 0$)			-0.038	-0.029
Std			0.18	0.432
RMSE			0.184	0.433

Case 4: Endogenous Network, Orthogonal Treatment, with Outcome Contagion				
	Naïve Regression	Matching	Outcome Contagion	Treatment Diffusion
Coeff ($\beta = 1$)	0.943	0.944	0.994	1.04
Std	0.203	0.252	0.21	0.251
RMSE	0.21	0.258	0.21	0.254
Coeff($\rho = .5$)			0.429	0.145
Std			0.182	0.582
RMSE			0.195	0.682

Case 5: Exogenous Network, Treatment Non-Orthogonal, No Outcome Contagion				
	Naïve Regression	Matching	Outcome Contagion	Treatment Diffusion
Coeff ($\beta = 1$)	0.994	0.964	0.99	0.994
Std	0.2	0.258	0.2	0.2
RMSE	0.2	0.26	0.2	0.2
Coeff($\rho = 0$)			-0.024	0.007
Std			0.138	0.303
RMSE			0.14	0.303

Case 6: Exogenous Network, Treatment Non-Orthogonal, with Outcome Contagion				
	Naïve Regression	Matching	Outcome Contagion	Treatment Diffusion
Coeff ($\beta = 1$)	1.026	1.055	1.007	1.018
Std	0.22	0.273	0.205	0.217
RMSE	0.221	0.279	0.205	0.217
Coeff($\rho = .5$)			0.453	0.359
Std			0.125	0.348
RMSE			0.134	0.376

Case 7: Endogenous Network, Treatment Non-Orthogonal, No Outcome Contagion				
	Naïve Regression	Matching	Outcome Contagion	Treatment Diffusion
Coeff ($\beta = 1$)	1.005	1.031	1.003	1.004
Std	0.214	0.296	0.213	0.214
RMSE	0.214	0.297	0.213	0.214
Coeff($\rho = 0$)			-0.025	-0.005
Std			0.135	0.281
RMSE			0.137	0.281

Case 8: Endogenous Network, Treatment Non-Orthogonal, with Outcome Contagion				
	Naïve Regression	Matching	Outcome Contagion	Treatment Diffusion
Coeff ($\beta = 1$)	1.25	1.542	1.021	1.197
Std	0.244	0.405	0.229	0.234
RMSE	0.349	0.677	0.23	0.306
Coeff($\rho = .5$)			0.465	-0.03
Std			0.105	0.362
RMSE			0.11	0.641

Contagion, Network Selection, & Especially Coevolution Pose Large Problems for Nonparametric Causal-Inference

- **Some highlights of results**
  - Combination of network selection & network contagion by far the most problematic for all the incorrect estimation strategies.
    - Worst of all if furthermore treatments non-orthogonal (i.e., not perfectly experimental), but even if random-control assigned, "indirect effects" esp. poorly estimated.
  - Propensity-score matching (perhaps surprisingly) dominated by simple regression; quite appreciably so in worst cases (selection & contagion).
  - In these relatively clean conditions, the problems for matching or treatment-spillover models show mainly as inefficiency (as expected), and much worse for the "indirect" than the "direct" effects.
    - In worst case, *treatment-effect* estimate bias is +20% & *indirect effects* horribly estimated.
  - Correctly specified estimation model with appropriate estimator dominates, of course, dramatically so when selection & contagion, & even more dramatically when treatment non-orthogonal (i.e., outside experimental contexts)

- And/but this is all taking the ATE/Causal-Parameter (not d**y**/d**x**) as estimand:
  - If instead causal-*response* is estimand, then even when less-structural estimation strategy gets the parameter right, it's horribly mistaken about response (because no feedback & can't be). In fact, estimate not even in the right dimensionality!

The Curse of Dimensionality & the Logical Impossibility of Truly *Nonparametric* or *Model-Free* Inference

C. Consider, e.g., a system of M endogenous equations like this:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ y_m \end{bmatrix}_i' \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1m} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{m1} & \gamma_{m2} & \cdots & \gamma_{mm} \end{bmatrix} + \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix}_i' \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{k1} & \beta_{k2} & \cdots & \beta_{km} \end{bmatrix} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \vdots \\ \varepsilon_m \end{bmatrix}_i$$

1. In matrix notation, system written compactly as:

$$\underbrace{\mathbf{y}_i'}_{1\times M} \underbrace{\mathbf{\Gamma}}_{M\times M} + \underbrace{\mathbf{x}_i'}_{1\times K} \underbrace{\mathbf{B}}_{K\times M} = \underbrace{\boldsymbol{\varepsilon}_i}_{1\times M} .$$

2. Even just $V(\varepsilon) \equiv \mathbf{\Sigma}$ has ½M^2+½M > M things to learn, in general, from each M things observed in each context i… (& assuming that VCov fixed over i).

3. Causal estimation & inference from any sort of data, observational or experimental, requires that this number of parameters (things to learn) per observation be reduced to less than 1 (i.e., parameters/observations<1).

- Point simply that, being fully non-parametric, the number things to estimate grows at least exponentially in the number of observations: generally impossible w/o model to reduce parameterization.
- So, ***models***: I want them; in fact, point of exercise is to estimate them: **Useful Empirical Simplifications**. To infer out of sample (& often beyond support as well); simply cannot without model. But, even if you don't like models, you cannot infer much (anything?) w/o one. & not so sure simpler model necessarily implies less-restrictive model… I'd rather try theory & substance first & appeal to simplicity second ⇒ ***EMTI***

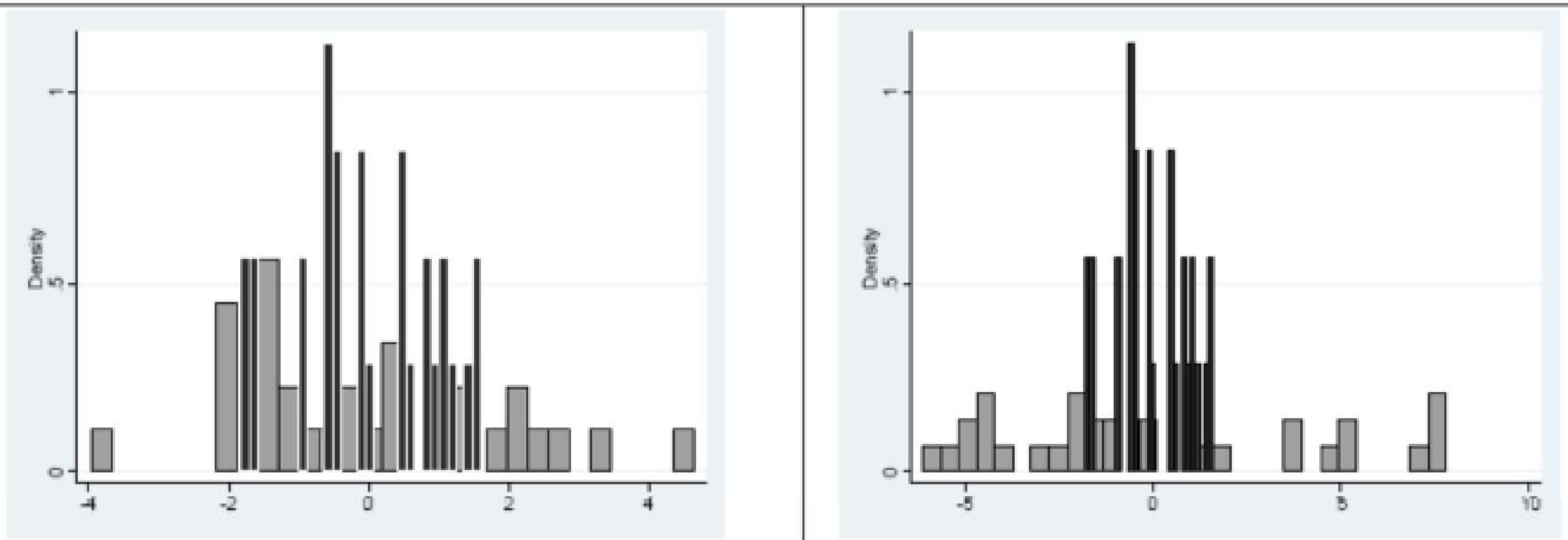
To FE or Not To FE: The Not-So-Harmless DLMFE estimator / estimation strategy

4. Part of how FE manifests is tendency to *pick up* too much heterogeneity and call it unit-fixed &, in LSDV case, part systematic.

   a) I.e., the *sweep* sweeps both fixed & stochastic unit-specific effects.

   b) I.e., classic *overfitting* = another way see incidental-param problem

   c) Troeger's MC's illustrate problem: note severe overdispersion of estimated relative to actual unit-specific effects:

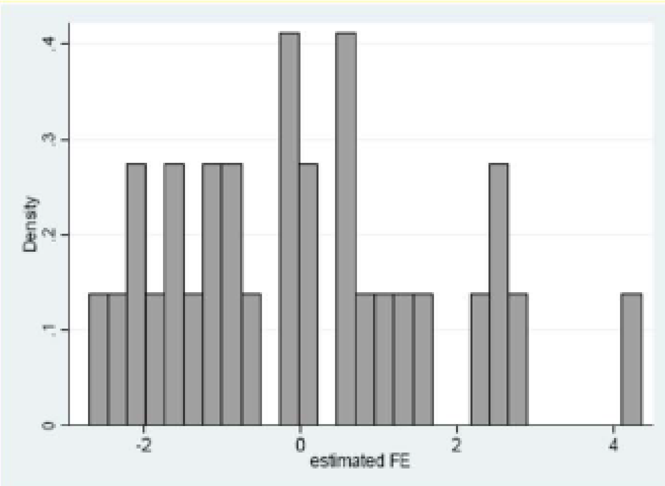


black: true FE, grey: estimated FE
Settings: FE in DGP ~ N(0,1), 1 RHS variable, SD(within)=SD(between)=1

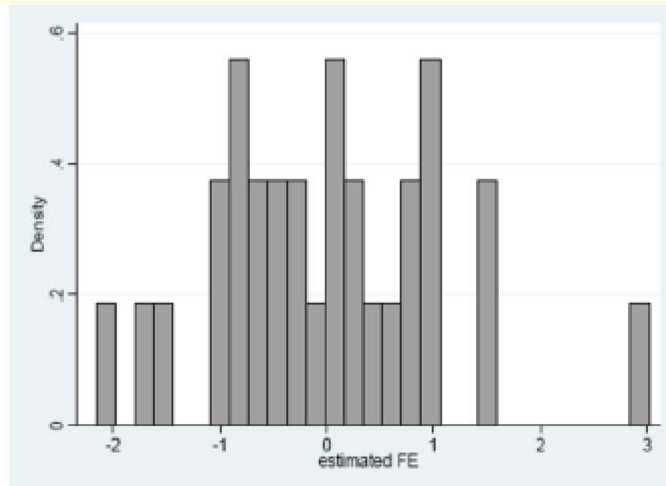
black: true FE, grey: estimated FE
Settings: FE in DGP ~N(0,1), 1 RHS variable SD(within)=1 SD(between)=3

To FE or Not To FE: The Not-So-Harmless DLMFE estimator / estimation strategy

d) Can be even worse. Will even find fixed-effects where they ain't:



Settings: no FE in DGP, 1 RHS variable, SD(within)=SD(between)=1



Settings: no FE in DGP, 3 RHS variables, SD(within)=SD(between)=1

<u>Notice</u>: in both this & previous case, the unit-effects not obviously biased (I think may/should be a small-sample inflation bias in the FE's and a corresponding small-sample attenuation bias in the **b**'s), but at least highly inefficient. Even if this latter is the case, in limited (in *T*) samples, these "*mere* inefficiency" issues can be *severe*.

Then, if other problems w/ model &/or estimation strategy also, e.g. dynamic misspecification, this overfitting will furthermore induce biases in other parameter estimates that can easily make FE the worst option of the family of panel/TSCS models. Worse even than just the rampant Type II error that often attendant DLMFE

The view that FE at worst merely inefficient & RE biased "insufficiently nuanced".

Seems to me FE-based strategies like D-in-D should inherit these shortcomings.

An Unfortunate Syllogism for the Current Orthodoxy as Applied to Social Science…

- The Four (no Five, no Six) Fundamental Problems of Empirical Analysis in Social Science: [An empirical comparativist's manifesto: *Context Matters* –]
  - **<u>1. Multicausality</u>**: just about everything matters.
  - **<u>2. Heterogeneity & Context-Conditionality</u>**: the way just about everything matters depends on just about everything else.
  - **<u>3. Temporal, Spatial, Spatiotemporal Dynamics</u>**: just about everything is dynamic, not static.
  - **<u>4. Ubiquitous Endogeneity</u>**: just about everything causes j.a. everything else.
    - [0. **<u>Micronumerosity</u>** (Goldberger): We have precious little data/useful variation with which to sort it all out.]
    - [-1. & the truth is probably moving on us (but that's just unobserved #2 again)]
  - [I.e., a conjecture: if Social, Political, &/or Economic, then not SUTVA.]
- **No (limited, very limited) way forward to *<u>CAUSAL ESTIMATION</u>* without imposing structure, i.e. models, ideally as theoretical & substantively motivated/specified as possible, & estimate as close as possible in actual contexts to which wish to infer…but that's fine with me. I like models. I think they're very much of the point of the empirical exercise…to obtain *<u>USEFUL EMPIRICAL SIMPLIFICATIONS</u>*.**