Analysis of tax reforms: Implications from panel data techniques

Andreas Peichl
ZEW, U Mannheim

DG ECFIN Structural Reforms Workshop:
„Inequality and Structural Reforms: Methodological concerns",
Brussels, 2017-05-16

Measuring the effect of policy (reforms) on Inequality

- (Rising income) inequality big (policy) issue
- Role of (structural) policy reforms?

- Problem: measuring causal relationships empirically!
  - Different methods
  - Role of data

- This talk: focus on panel data methods

Evaluating Empirical Studies

- Consider ideal experimental design first; then formulate feasible design and analyze its flaws relative to ideal design
- Interested in causal relationships: "How does X affect Y?"
  - Identification = Concept of understanding causal relationship behind empirical results.
  - Important to distinguish between correlation and causation.
  - Counterfactual: What would have happened to Y in absence of X?
- Fundamental problem: we do not observe the counterfactual
  - What would have happened to the treated if they had not been treated?
  - What would have happened to the untreated if they had been treated?
  - → need to rely on 'identifying assumptions'

Identification

- A "research design" is a source of variation in outcomes that is credibly unrelated to selection into treatment

- Different methods with different 'identifying assumptions':
  - Experiments: generate random variation
  - Natural Experiments: know what generated the variation
  - Instrumental variables: have variable that provides quasi-randomized variation
  - Structural models: rely on (testable) econometric assumptions

  → trade-off btw. credibility & policy-relevance

Panel data techniques

- **Difference-in-Difference:**
  - Identifying assumption: "parallel trends":
    - Absent policy change, change btw t0&t1 would have been the same for T & C
    - Useful to plot long time series of outcomes for T and C
    - Can use placebo DiD to test parallel trends assumption
- **Fixed effects generalize DiD**
  - Usual setting: many countries, years and policy changes
  - Identification obtained from within-country variation over time
    - Advantage: more precise estimates by pooling several changes
    - Disadvantage: FE is black-box regression, more difficult to check trends non-parametrically as with a single change
- **Same parallel trends identification assumption as DiD**
  - Potential violation: policy reforms may respond to trends in outcomes

Example: My own work

- Dörrenberg / Peichl (2014): The impact of redistributive policies on inequality in OECD countries, Applied Economics, 2014, Vol. 46, No. 17, 2066–2086.

- Examining inequality trends and the role of policy changes in OECD countries between 1981 and 2005 using different country-level data sets

Overview

- Estimate the effect of 3 policy variables – government spending, social expenditure and progressive taxation – on income inequality.
- Exploit within-country variation in OECD countries over time (1981–2005) and apply different identification strategies (FE, IV)
- Results:
    - Government expenditures reduce inequality, despite potential offsetting second-round effects.
    - The effects of changes in tax progressivity are smaller and often insignificant implying that indirect behavioral effects play a (bigger) role with tax progressivity.
    - Discussion of endogeneity / identification
    - Our results also show that (different) inequality data sources should always be handled with caution.

Empirical approach I: FE

- Identification strategy I: Fixed effects panel estimations

$$Y_{i,t} = \beta_1 x_{i,t-1} + \beta_2 C_{i,t-1} + \gamma_t + \mu_i + \varepsilon_{i,t} \qquad (1)$$

where i denotes a country and t the point of time. $x_{i,t-1}$ is one of the lagged policy variables of interest and $C_{i,t-1}$ is a vector of several lagged control variables. Time and country fixed effects are captured by γ_t and μ_i, respectively. $Y_{i,t}$ indicates the level of inequality. $\varepsilon_{i,t}$ is a SE term and our coefficient of interest is β_1.

Dependent variable: measure of inequality

Panel A: LIS Gini					
Gov't spending	−0.380***	−0.356***			
(lagged)	(0.120)	(0.108)			
Social expenditure			−0.232**	−0.270***	
(lagged)			(0.092)	(0.091)	
Progressivity		0.007		−0.132**	−0.015
(lagged)		(0.108)		(0.061)	(0.097)
Observations	120	117	116	113	117
R2	0.546	0.545	0.564	0.588	0.515
Panel B: WIID Gini					
Gov't spending	−0.307**	−0.230**			
(lagged)	(0.113)	(0.104)			
Social expenditure			−0.051	−0.023	
(lagged)			(0.076)	(0.080)	
Progressivity		0.041		0.049	0.036
(lagged)		(0.046)		(0.045)	(0.050)
Observations	368	351	349	338	351
R2	0.314	0.299	0.328	0.339	0.280
Panel C: UTIP estimate					
Gov't spending	−0.103	−0.081			
(lagged)	(0.076)	(0.090)			
Social expenditure			0.001	−0.004	
(lagged)			(0.052)	(0.054)	
Progressivity		−0.025		−0.021	−0.031
(lagged)		(0.025)		(0.026)	(0.025)
Observations	437	389	378	353	389
R2	0.493	0.435	0.553	0.518	0.425

Discussion of FE estimates

- FE estimates hint at that redistributive policies can reduce inequality
- "However, due to remaining endogeneity and imprecise coefficients, *we do not find clear-cut evidence that second-round effects do not offset redistributive policy measures* – especially for progressive taxation."
- Estimates (& SE) differ depending on data source. This raises concerns about the general data quality
- We must further be concerned that our identification strategy is not sufficient in order to obtain unbiased effects.
  - Problem of reverse causality not properly accounted for *if the effect of inequality levels on redistribution policies is not systematic within countries.*
  - It might, e.g., be the case that levels of inequality in year *t-1* affect policies in *t* in a *not systematic way*, yielding the necessity to control for lagged levels of inequality. Doing so in the above framework would however not be legitimate.

System GMM ?

- Common way of dealing with such dynamic problems: include lagged LHS variable on RHS using GMM (Arellano & Bover 1995; Blundell & Bond 1998)

- However, System GMM estimators not appropriate for models with small N and large T (Roine et al., 2009; Roodman, 2009).

- Another reason for GMM methods not to be applicable to our setting rests on our scepticism that the required assumption of 'weak exogeneity' is met.

  – This assumption allows the explanatory variables to be correlated with past and current levels of the error term, but not with its future realizations.

  – We have a reverse causality effect at hand, where redistribution does not only decrease inequality, but where inequality also positively affects levels of redistribution. Applying the assumption of weak exogeneity to our case implies that the system-GMM estimator, using internal instruments, would only be valid if future levels of inequality do not affect the current levels of measures of redistribution. However, governments, that expect increasing inequality in the future, could implement redistributive measures to offset the anticipated shock on income inequality. We are therefore skeptical that the assumption of weak exogeneity is likely to hold in our set-up.

Empirical approach II: IV

- Finding suitable instruments is generally a difficult task.
  In our set-up, we require IVs that are independent of ε in Equation 2, but correlated with one of our three respective explanatory variables of interest (government spending, total social expenditure and degree of progressivity).

- An ideal instrument would be randomly assigned to each observation and therefore generate a quasi-experimental way to estimate the causal effect.

- Obviously, no such instrument is available in our setting with country-level observations. Policies or other variables that may be correlated with measures of redistribution are usually not randomly assigned to a country and thus mostly endogenous.

- We are not aware of any study that uses a randomly assigned instrument to identify the effect of policy variables on inequality in a country-level setting.

IV approach II

- Our IVs take the initial value of the respective regressor in 1981 and then grow with the growth rate of GDP or the growth rate of the highest MTR.
    - Our instruments are exogenous in the sense that we do not use the actual observed annual levels but extrapolated figures that are based on the initial levels in 1981.
    - The instrument for one of the regressors may increase between 2 years within a country, whereas the actual level of the regressor remains stable.
    - Though, we observe a strong correlation between the extrapolated numbers and the actual ones for our measures of government spending and social spending, hence satisfying the IV relevance requirement.
- For the exclusion restriction to hold, we have to assume that, conditional on our control variables and fixed effects, the inequality trends are uncorrelated with the (average) growth rates used for extrapolation. This can be justified because GDP and the top tax rate are, among other variables, controlled for on both stages of our estimation.

Dependent variable: measure of inequality

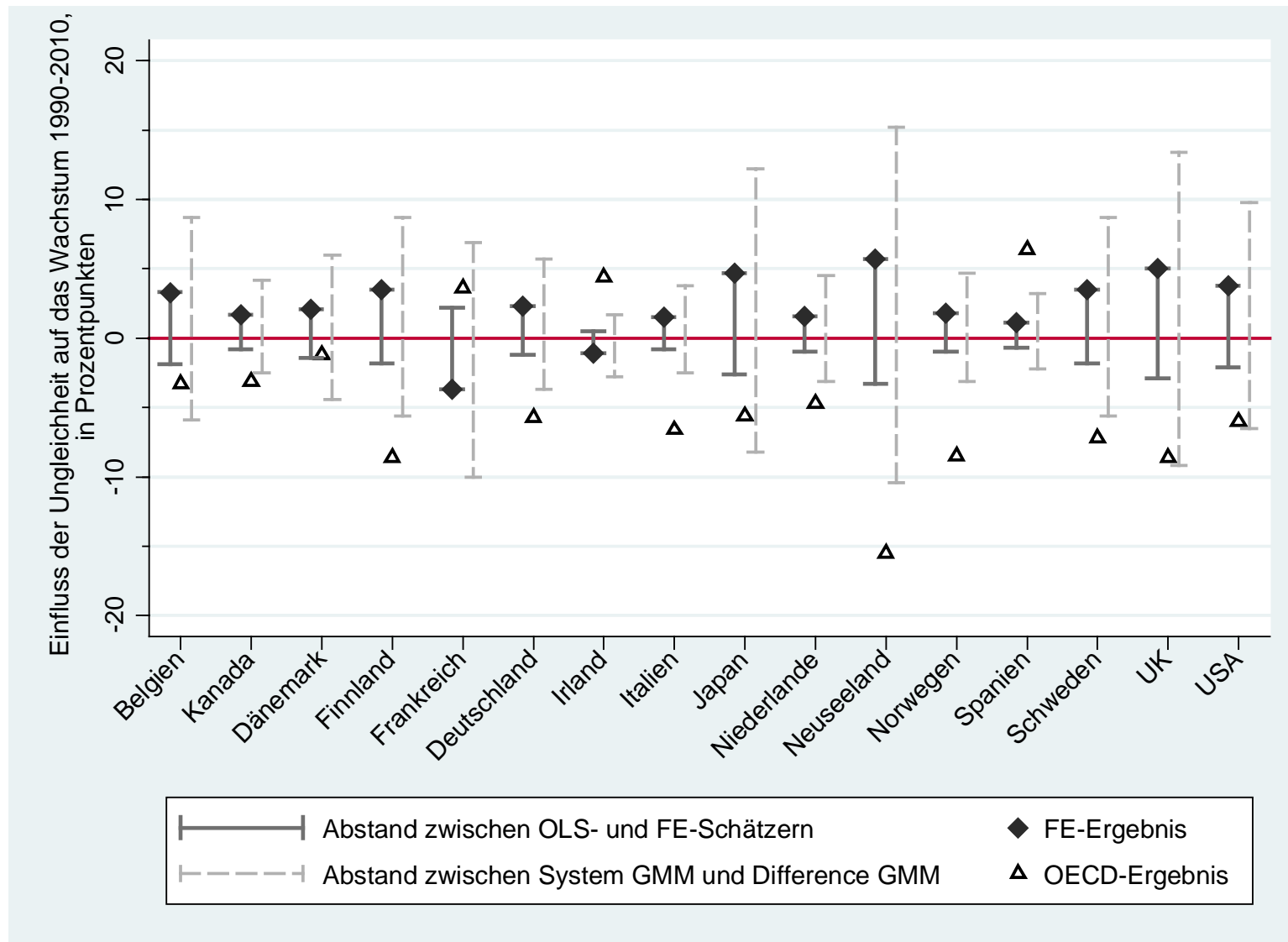|					
Panel A: LIS Gini					
Gov't spending	0.024	0.003			
	(0.218)	(0.226)			
Social expenditure			−0.216**	−0.295**	
			(0.106)	(0.116)	
Progressivity		0.035		−0.116*	1.252
		(0.070)		(0.069)	(0.789)
Observations	107	106	86	85	79
Panel B: WIID Gini					
Gov't spending	−0.152	−0.110			
	(0.153)	(0.162)			
Social expenditure			−0.107	−0.039	
			(0.138)	(0.139)	
Progressivity		0.098*		0.082	−0.223
		(0.053)		(0.053)	(0.854)
Observations	334	331	287	285	288
Panel C: UTIP estimate					
Gov't spending	−0.328***	−0.333***			
	(0.096)	(0.101)			
Social expenditure			−0.026	−0.016	
			(0.076)	(0.082)	
Progressivity		−0.017		−0.048**	0.167
		(0.022)		(0.024)	(0.414)
Observations	414	405	344	337	353

Discussion of IV results

- The results rest on the assumption of instrument validity.
    - That is, conditional on all control variables as well as the country and year FE, the IVs must affect inequality only through the independent variables of interest.
    - This assumption is untestable and its validity needs to be approached intuitively.
    - Clearly, our instruments are not randomly assigned to each country-year observation. However, we are able to exploit some exogenous variation as the extrapolated values are not directly related to inequality. Of course, the growth rate of the GDP and marginal tax rate, which we use for extrapolation, have some impact on the income distribution, but we control for GDP and tax rate on both the first and second stage and hence condition on these variables. We only exploit within-country variation in our instruments and control for any effects that are specific to the included countries and systematic across time.
- Although we find negative coefficients in most specifications, the SEs in some cases are large and imprecise. In IV estimations, weak first-stage results can increase the SEs on the second stage, but this is mostly not the reason in our case; first-stage results are sufficiently strong at least for specifications which use government spending and social expenditure.

General discussion: empirical approaches

- The difficulties of identifying a clear and clean causal effect are omnipresent and inherent to this literature.

- The problem underlying (almost) all country-level studies is: it is impossible to exploit a (quasi-)experimental identification strategy, i.e., to randomly assign redistributive policies to a set of countries.

- The FE and IV methods used here can certainly mitigate problems of endogeneity, but there remains doubt if they are sufficient.

- Alternatives:
    - DiD analyses of single policy changes may help to identify a clear effect. However, the required 'common trend' assumption across both countries is often difficult to defend. → *Synthetic control methods?*
    - System-GMM methods may also be able to handle some of the inherent problems of endogeneity, but require very rich data, which are often not available (*NOT small N and large T like here!*), and are highly sensitive to specification and the choice of internal instruments.
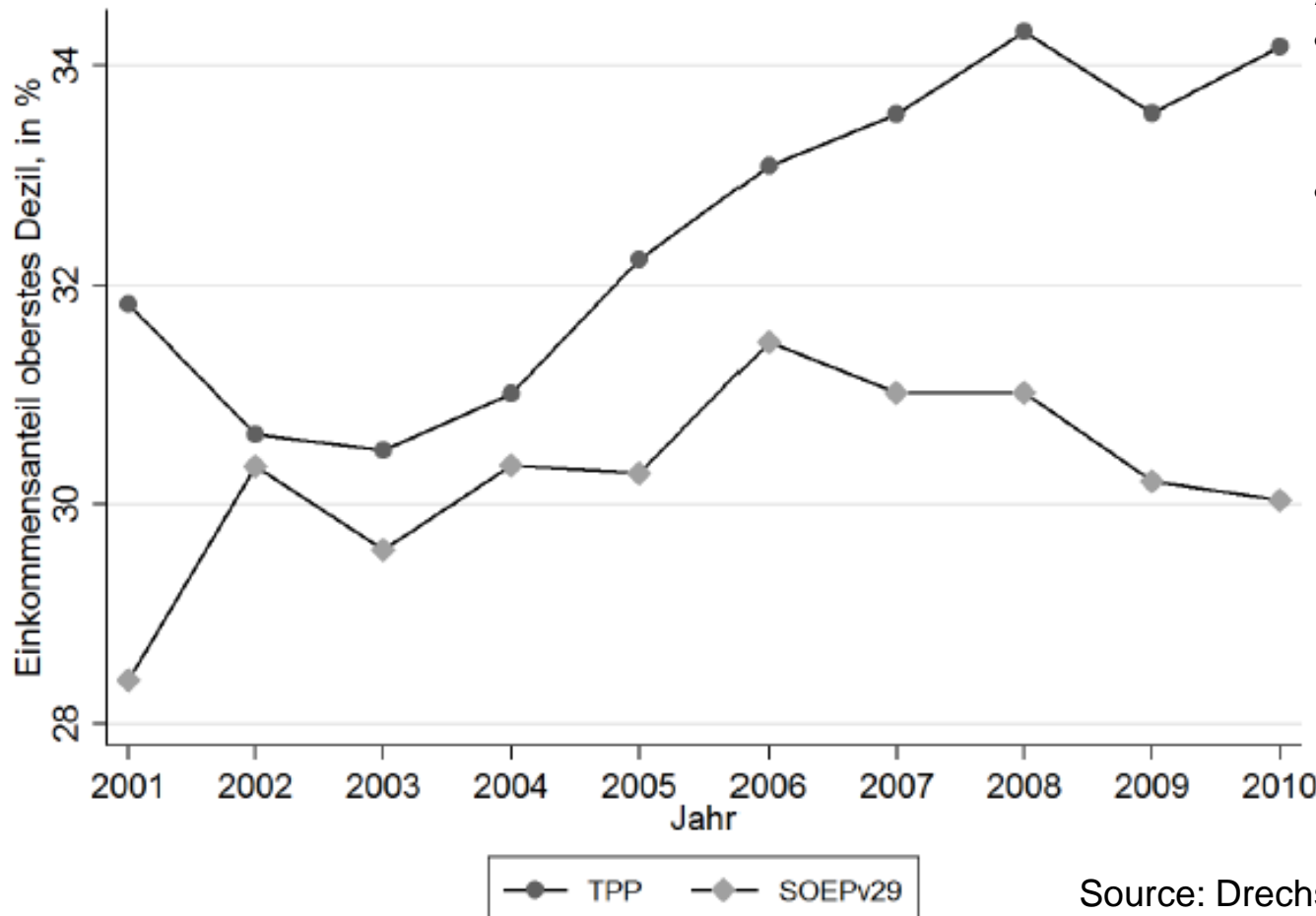
Impact of estimation methods (here Ineq & growth)

Einfluss der Ungleichheit auf das Wachstum 1990-2010, in Prozentpunkten

Abstand zwischen OLS- und FE-Schätzern	◆	FE-Ergebnis
Abstand zwischen System GMM und Difference GMM	△	OECD-Ergebnis

Belgien · Kanada · Dänemark · Finnland · Frankreich · Deutschland · Irland · Italien · Japan · Niederlande · Neuseeland · Norwegen · Spanien · Schweden · UK · USA

Inequality trends in DE: DATA matters!

Income share of top 10%

Admin (tax return) data:
- Better information on capital and business income
- Includes top tail

HH survey data
- GSOEP, also basis for LIS
- Better at bottom of distribution

Source: Drechsel-Grau / Peichl / Schmid (2015)

General discussion: policy lessons

- Results: government expenditures reduce inequality, while effects of tax progressivity are smaller (and often insignificant) implying that indirect behavioral effects play a (bigger) role with tax progressivity.

- Given the disincentive and distorting effects of progressive taxation, our results might imply that governments should combat inequality through progressive transfers rather than increasing the progressivity.

- Findings might help explaining differences in inequality btw EU and US:
  - While the US has a very progressive income tax schedule, very little redistribution occurs through social benefits.
  - In contrast, European welfare states rely (on average) much more on benefits and government expenditure to fight inequality (Fuest et al. 2010).
  - Introduction of the EITC in the US has effectively reduced inequality without large disincentive effects (Chetty et al., 2013).
  - Expanding the EITC and other benefits might be a fruitful way forward in order to combat rising inequality.
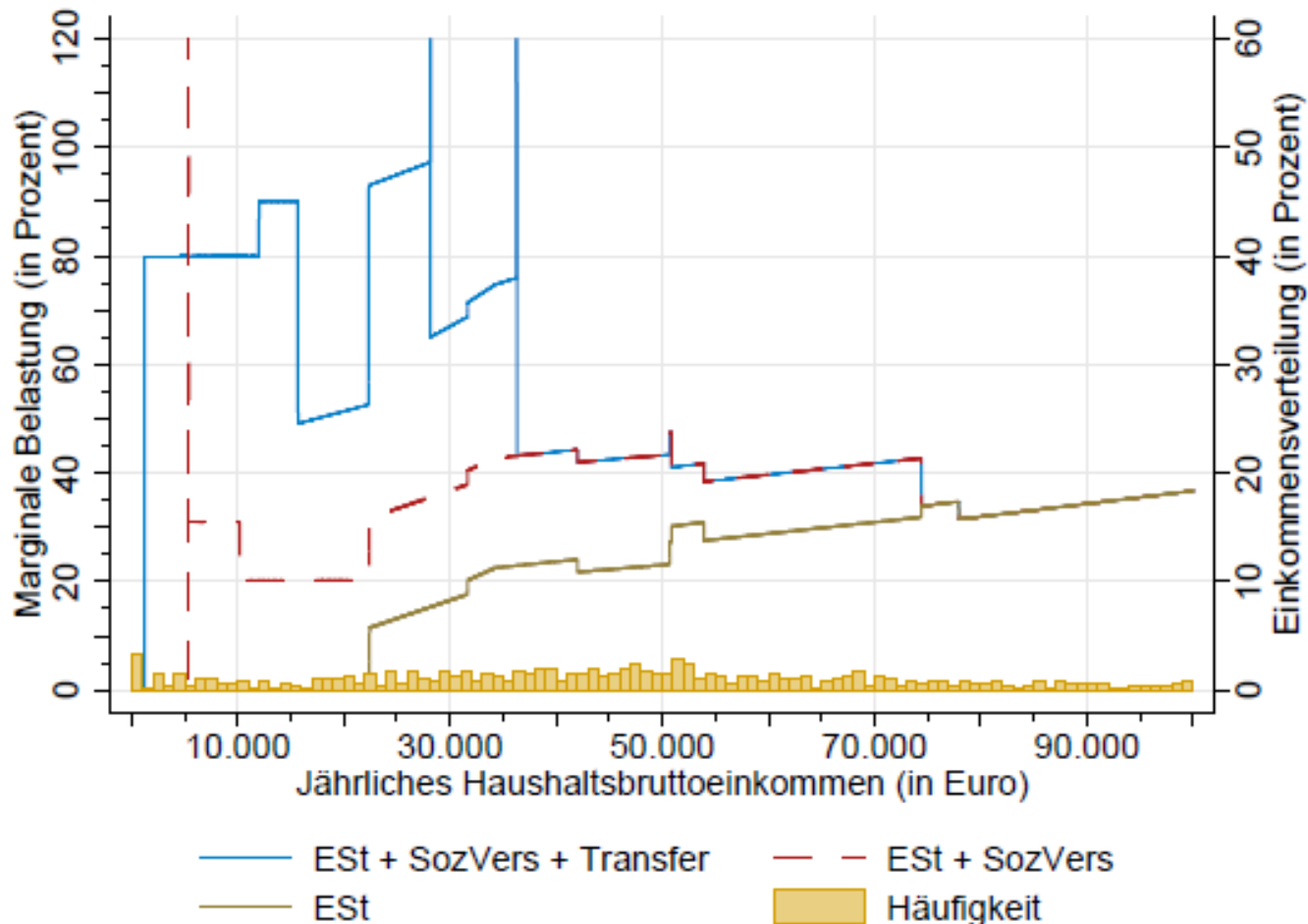
IPSP Chapter 3 on Inequality

Key policy message(s): „it depends":

- Policies can affect inequality pre-market (e.g. education, land), in markets (e.g. anti-trust laws), and post-market (e.g. redistribution)
- Common objectives, but priorities have to be country-specific
  - New model of the welfare state necessary?
- Political economy matters (e.g. size and strength of middle class, democracy, political coalitions, collective actions)
- Limited scope for international policies (e.g. tax avoidance, evasion; financial sector regulation; carbon pricing)

www.ipsp.org

Which policies do we consider?

Abbildung 8: Effektive Grenzbelastung Alleinverdiener-Paar, zwei Kinder

Legend:
- ——— ESt + SozVers + Transfer
- – – ESt + SozVers
- ——— ESt
- ▨ Häufigkeit

X-axis: Jährliches Haushaltsbruttoeinkommen (in Euro)
Left Y-axis: Marginale Belastung (in Prozent)
Right Y-axis: Einkommensverteilung (in Prozent)

Scope for reform?
Simulating (rev.-neutral) reduction in MTR (claw-back rate) at bottom:
- Increase in LS
- Decrease in Ineq (Buhlmann et al. 2017)

Conclusions

- Data matters → need for better (more admin) data
- Identification crucial – but problematic in policy relevant (macro) context (see also debate on inequality and growth)
- Do NOT rely on cross-country data estimates (for policy advice)!

Alternatives:

- Using simulation models (MSM-DSGE): generate exogenous variation by construction
  - Linking EUROMOD & QUEST: Barrios, Salvador, Mathias Dolls, Anamaria Maftei, Andreas Peichl, Sara Riscado, Janos Varga, Christian Wittneben (2017), Dynamic Scoring of Tax Reforms in the European Union, ZEW DP 17-017.
- Synthetic control methods
  - see Daniel's talk

Thank you!

peichl@zew.de