Marginal Tax Rates and Income: New Time Series Evidence

by Karel Mertens

Lorenz Kueng

Northwestern University's Kellogg School of Management
and NBER

In a nutshell:

- this is a **great paper!!** (really)
- lots of substance in it
- very carefully done

I will use my discussion time to

- give you my (limited) understanding of it,
- provide some suggestion how you might improve it further,
- ask many questions (mostly reflecting my ignorance).

Setting

$$\Delta ln(e_t) = \eta \Delta ln(1 - \tau_t) + v_t$$

- e is taxable income or earnings
- $(1 - \tau_t)$ is net-of-tax rate, where τ is the AMTR
- Define β^{OLS} as linear projection of e_t on $(1 - \tau_t)$

2 Problems:

1. η **might not reflect the micro elasticity** ϵ of taxable income to MTR changes,
   - which since Feldstein (1995) is the main focus of a large micro literature in PF
   - and an important parameter for optimal taxation

   However, η might still be what macroeconomists care about: total causal effect of AMTR change on aggregate income, including any GE effects.

2. η **might not** correspond to any **causal parameter**

Example of 1st Problem

Classic wage endogeneity

- HH supplies hours $h_s = [(1-\tau)w]^{\epsilon}$
- Firms demand hours h_d such that $w = f'(h) = \alpha h_d^{\alpha-1}$
- Suppose taxes exogenous
- $$\ln(e) = \underbrace{\alpha \frac{\overbrace{\epsilon}^{\frac{d\ln(h)}{d(1-\tau)}}}{1+(1-\alpha)\epsilon}}_{\eta=\beta^{OLS}\neq\epsilon} \ln(1-\tau) + ...$$

Suggestion

- paper doesn't really try to recover micro elasticity ϵ
- however you devote quite some space discussing this literature
- I suggest you move this to the appendix (at least I got confused at first...)

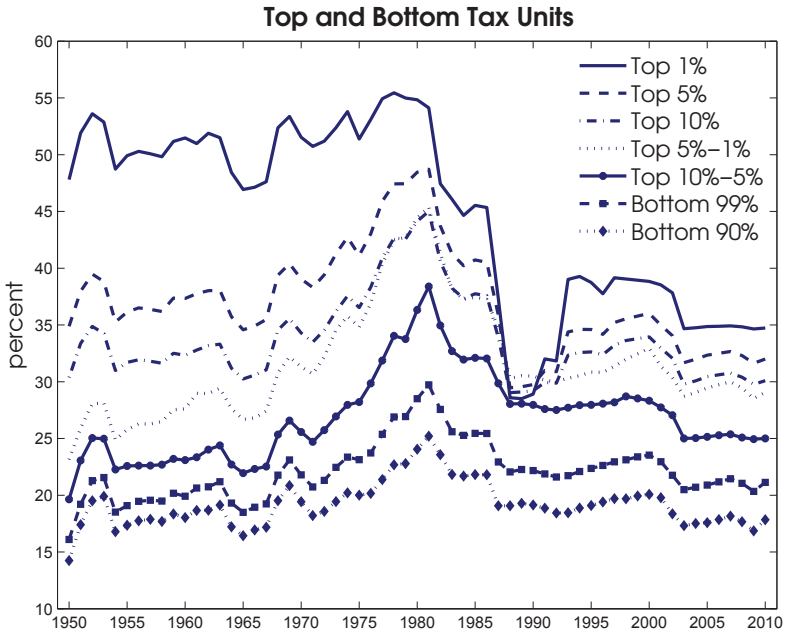
back to the paper...

2nd Problem (main point of paper)

$$\Delta ln(e_t) = \eta \Delta ln(1 - \tau_t) + v_t$$

Endogeneity of MTRs:

- If $Cov(v, \tau) \neq 0$, $\beta^{OLS} \neq \eta$
- Easy example: reverse causality. Tax changes respond to economy and hence to income.
- Nice job discussing many other examples:
  - GE effects
  - bracket creep
  - anticipated tax changes
  - endogenous income distribution
- Shows that in most settings
  - OLS is biased downward and
  - bias is less severe at the top (because more volatile τ)



- high correlation of AMTR changes across income distribution $\Rightarrow$ **GE effects**
- shows **endogeneity of AMTR** (e.g. financing wars, bracket creep,...)

Looking at statutory rate changes solves "mechanical" endogeneity (e.g. bracket creep), but not endogenous tax policy.

Karel's solution

- Borrow "exogenous" narrative tax shocks from Romer-Romer (2009)
- "Exogenous" = uncorrelated with current GDP
- Use insight from Stock-Watson (2008) and Mertens-Ravn (2013) that narrative shocks are not structural in a SVAR sense, but only a proxy (possibly a noisy subset of str'l shock)
- Treat them as IVs to identify SVAR
  $\Rightarrow$ **"proxy SVAR" approach**

SVAR

$$\begin{pmatrix} 1-\tau_t \\ e_t \\ X_t \end{pmatrix} = \begin{bmatrix} 0 & \xi_e & \xi_x \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} 1-\tau_t \\ e_t \\ X_t \end{pmatrix} + \begin{bmatrix} 1 & 0 \\ \eta & \zeta_e \\ \zeta_x & \theta \end{bmatrix} \begin{pmatrix} \nu_t^\tau \\ \nu_t^o \end{pmatrix} + A(L)\nu_{t-1}$$

ν are the structural shocks.

Question It wasn't clear to me what restrictions you are imposing here (if any).

- Seems you're restricting the contemporaneous feedback on e and X but not on τ

VAR

$$y_t = B(L)y_{t-1} + \begin{pmatrix} u_t^\tau \\ u_t^e \\ u_t^x \end{pmatrix}$$

Step 1: Get $\hat{u}$ by OLS.

Step 2: Estimate parameters and shocks

$$
\begin{aligned}
u^{\tau} &= \xi_e u^e + \xi_x' u^x + \nu^{\tau} && (1)\\
u^e &= \eta \nu^{\tau} + \zeta_e' \nu^o && (2)\\
u^x &= \theta \nu^{\tau} + \zeta_x \nu^o && (3)
\end{aligned}
$$

Problem: Cannot estimate ν^{τ} with OLS residuals from (1) since it is endogenous due to (2) and (3).

Solution: Use Romer-Romer narrative shock m_t as IV to estimate parameters and structural shocks.

Nice since only requires instrument to be correlated:

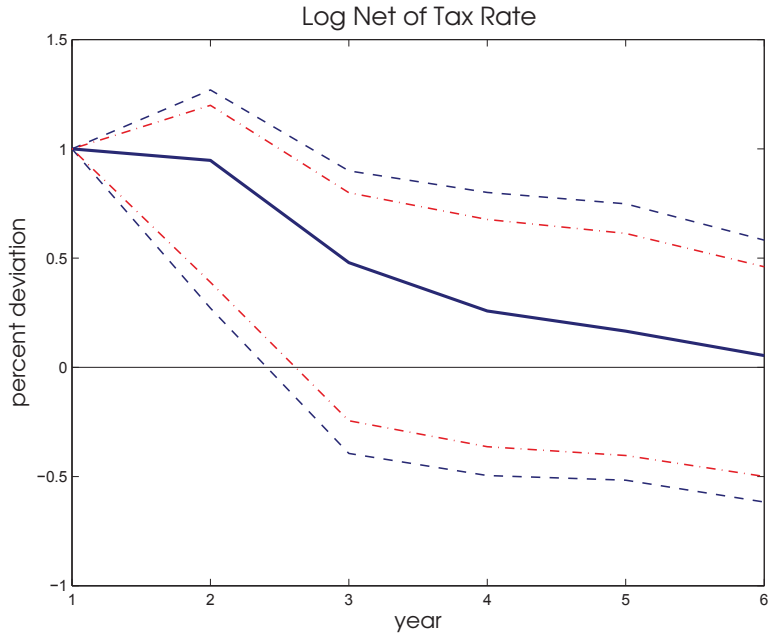
- can use Romer-Romer ATR to proxy for AMTR.
- only requires no contemporaneous correlation, hence RR deficit-driven shocks are indeed "exogenous"

Figure 4a vs. Romer-Romer Figure 6a

Question:

- Why is the response of taxes to tax shock so different between in SVAR and Romers' VAR?
  - only subset of personal income tax shocks?
  - only subset of those PI shocks?
  - richer VAR than Romers?
  - AMTR vs. ATR?

- How much of the persistence (and the hump shape) of your results is driven by the more persistent dynamics of your shocks?



Panel A. Response of tax to tax

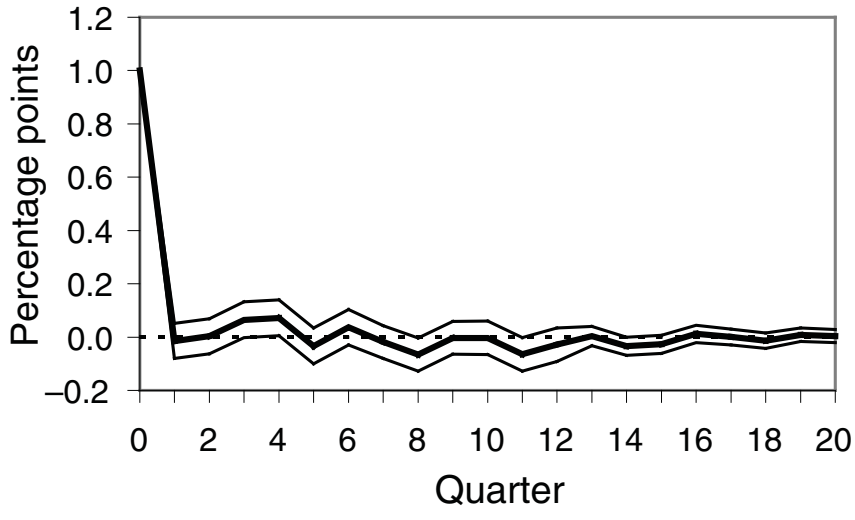


Figure 4c

- Figure 4c shows large and persistent response of total (economy-wide) income to AMTR shocks
- I was surprised about the precision of the estimates given only 12 shocks (and 7 for the top 1%)!

Suggestion

- Could include corporate income (earnings before taxes) and realized capital gains to **assess income shifting** (although not avoidance or evasion).
- Could include CI and cap gain tax shock (although they are not uncorrelated based on Romer-Romer), but could orthogonalize

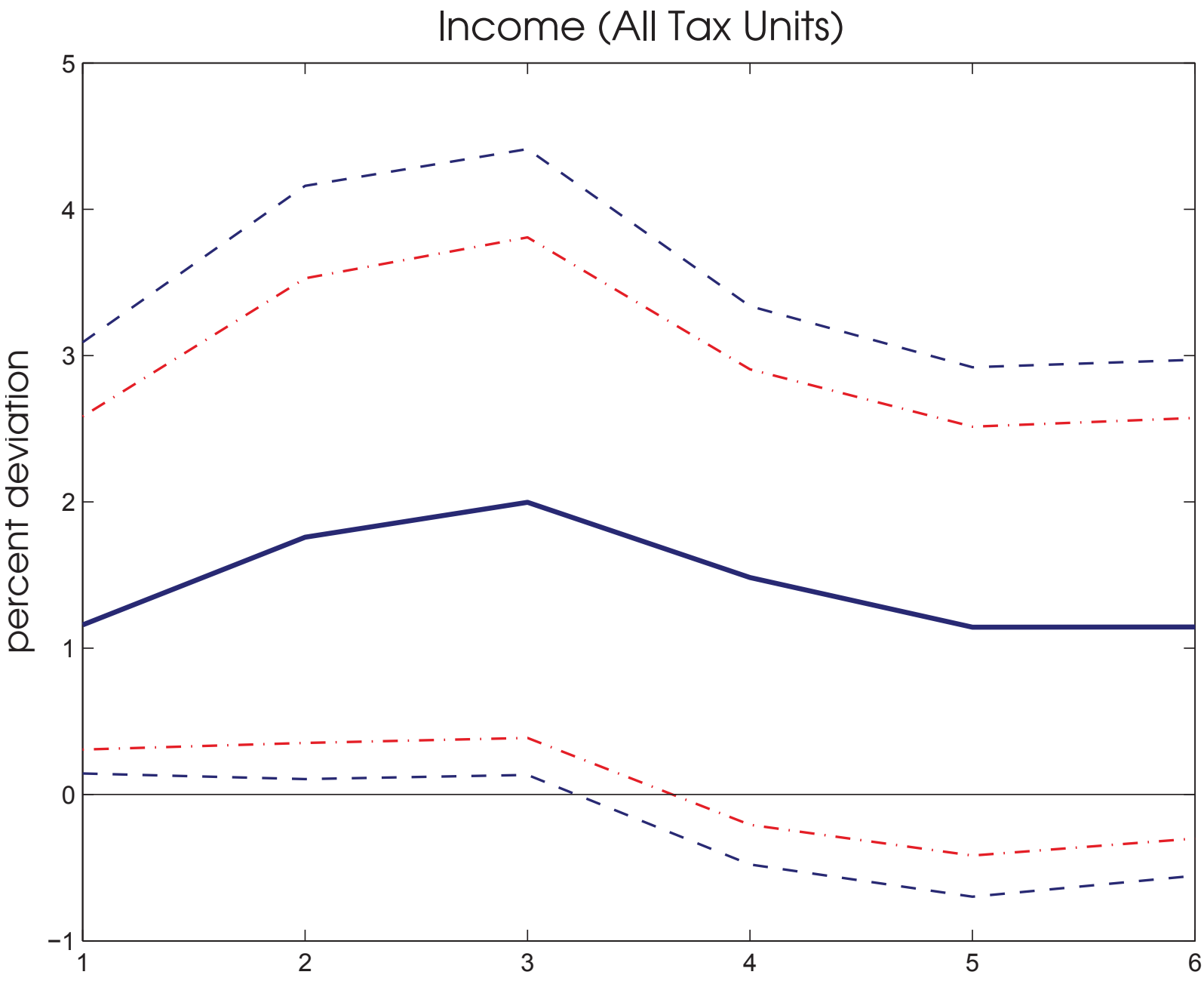
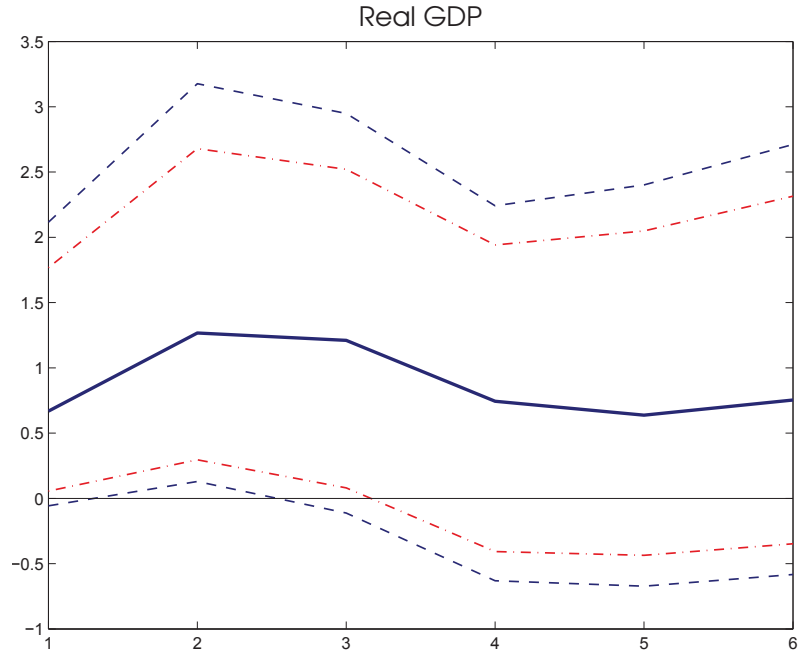

Income (All Tax Units)
percent deviation
5
4
3
2
1
0
−1
1
2
3
4
5
6

Figure 4b vs. Romer-Romer Figure 6c

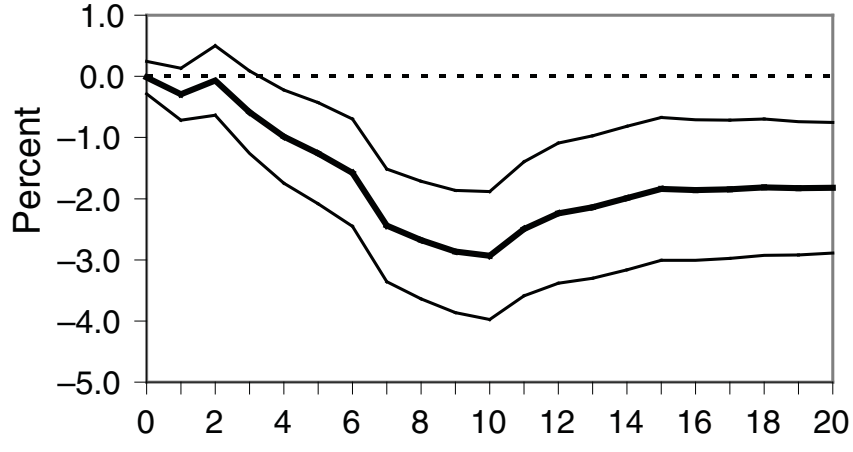
- Figure 4b using GDP shows that part of this is probably income shifting, but there remains a sizable real effect
- Romers' GDP response is much larger (probably too large...)

Question Why is Romers' GDP response to ATR much larger than the response to your AMTR shock?

- Income vs. substitution/distortion? If so could you decompose the two using both shocks (but would probably invalidated using ATR as an instrument)?
- Smaller sample of shocks?
- Only PI taxes?



Panel C. Response of GDP to tax



Tables 2 and 4 compare OLS with IV across income distribution

- Estimates much more precise for top incomes
- IV response seem to be somewhat smaller for lower incomes, but not much!
- response of bottom 90% is surprisingly large
- Definitely much higher elasticity than micro TI elasticities (based on relative differences, e.g. using control groups) and LS elasticities, especially at the bottom
- Suggests GE effects and maybe extensive margin effects?

Suggestion

- Might be interesting to look at hours to assess extensive margin (Chetty-Manoli-Weber)

	All Tax Units		Top 1%	Top 5%	Top 10%	Top 5-1%	Top 10-5%	Btm. 99%	Btm. 90%
	Series 1	Series 2							
OLS									
Total Inc.	−0.33	−0.42	0.58*	0.28	0.14	−0.14	−0.27*	−0.60**	−0.80***
	(−0.89, 0.23)	(−0.95, 0.11)	(−0.08, 1.25)	(−0.37, 0.94)	(−0.44, 0.72)	(−0.46, 0.17)	(−0.58, 0.04)	(−1.08, −0.12)	(−1.37, −0.23)
IV									
Total Inc.	1.16**	1.37**	1.27***	1.12**	1.11**	0.65*	0.83*	1.13*	1.33*
	(0.25, 2.08)	(0.12, 2.61)	(0.41, 2.12)	(0.18, 2.07)	(0.13, 2.09)	(−0.12, 1.42)	(−0.14, 1.80)	(−0.03, 2.29)	(−0.01 – 2.67)

Figure 8b,c,d

Since results suggest

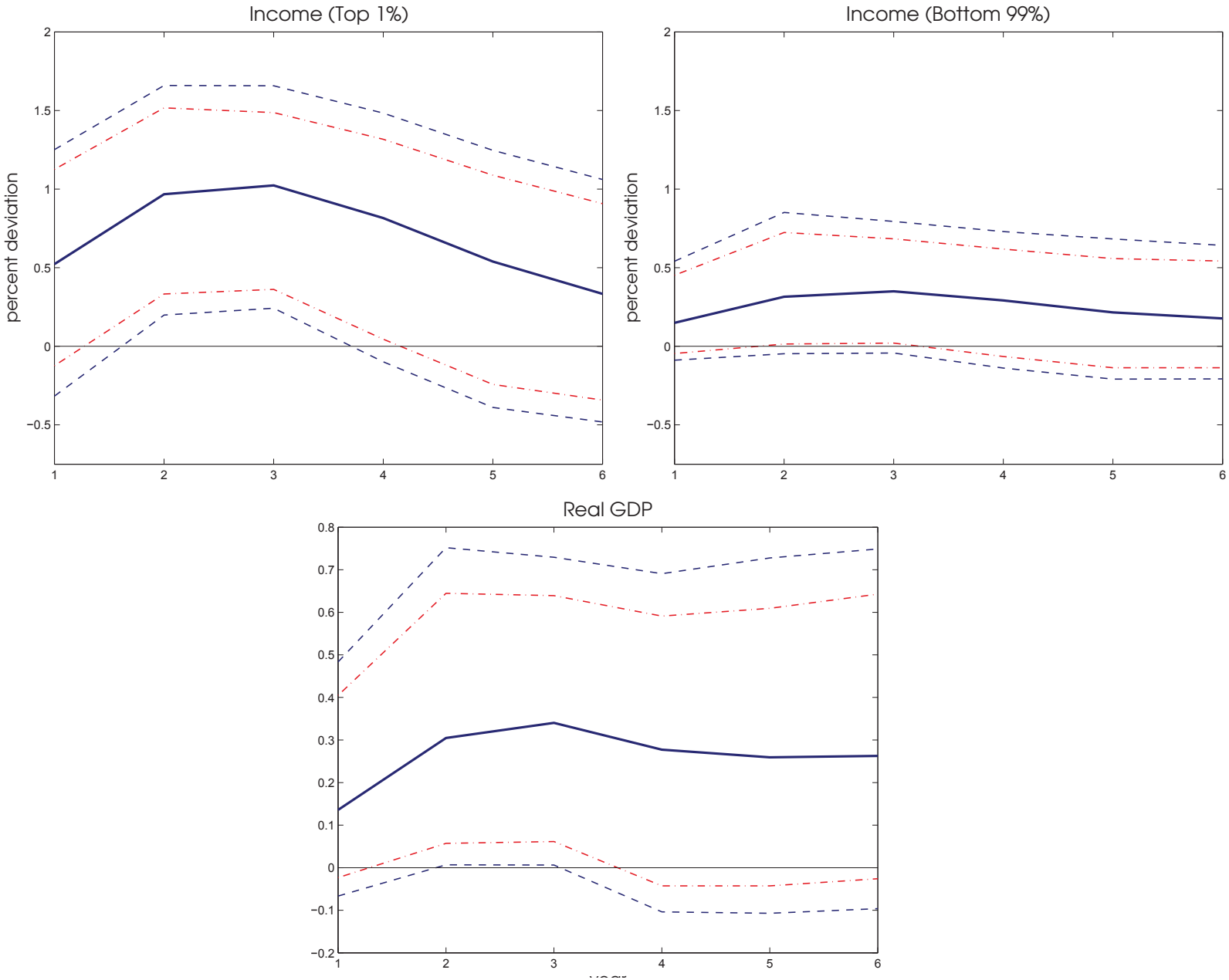
1. top 1% drives a lot of the aggregate response and
2. top 1% is also more precisely estimated

Karel looks at shocks to top MTR.

- Indeed he finds significant response of total income of top 1% controlling for bottom 99%
- even marginal effect on GDP
- also marginal effect on bottom 99% (!), suggesting spill-overs

  $\Rightarrow$ **trickle-down economics resurrected?**
- however effect is smaller than using all tax rates since he includes (controls) bottom 99%

  $\Rightarrow$ again, suggests taking GE into account is important


Income (Top 1%)
percent deviation
2
1.5
1
0.5
0
−0.5
1
2
3
4
5
6
Income (Bottom 99%)
percent deviation
2
1.5
1
0.5
0
−0.5
1
2
3
4
5
6
Real GDP
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
−0.2
1
2
3
4
5
6

Suggestion

- Since top 1% response very precisely estimated, could be interesting to further disaggregate a la Piketty-Saez (i.e. 0.1% etc.)
- Is there any variation in AMTR within the top 1%?
  - maybe earlier in the sample (before TRA 1986)
  - before 1960s when there was no max tax on earned income
  - ...

What I've learned from Karel's paper

- GE effects seem larger than I expected (although not as large as RR)
- Proxy SVAR seems to be a nice tool (I need to study up...)
- Confirms that it is hard to go from micro to macro "elasticities"
- I believe micro identification approaches are still our best hope to estimate ϵ, but not for η
- didn't even go into implications for inequality!

This is a great paper!