ARE TAXES TOO HIGH? A MACHINE-LEARNING APPROACH TO LAFFER CURVE ESTIMATION

Hermes Morgavi

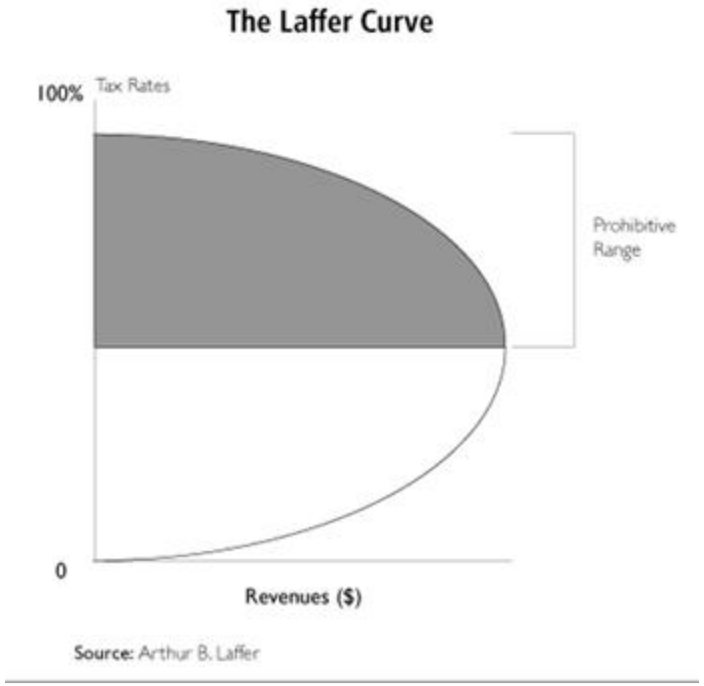
OECD
BETTER POLICIES FOR BETTER LIVES

The Laffer curve

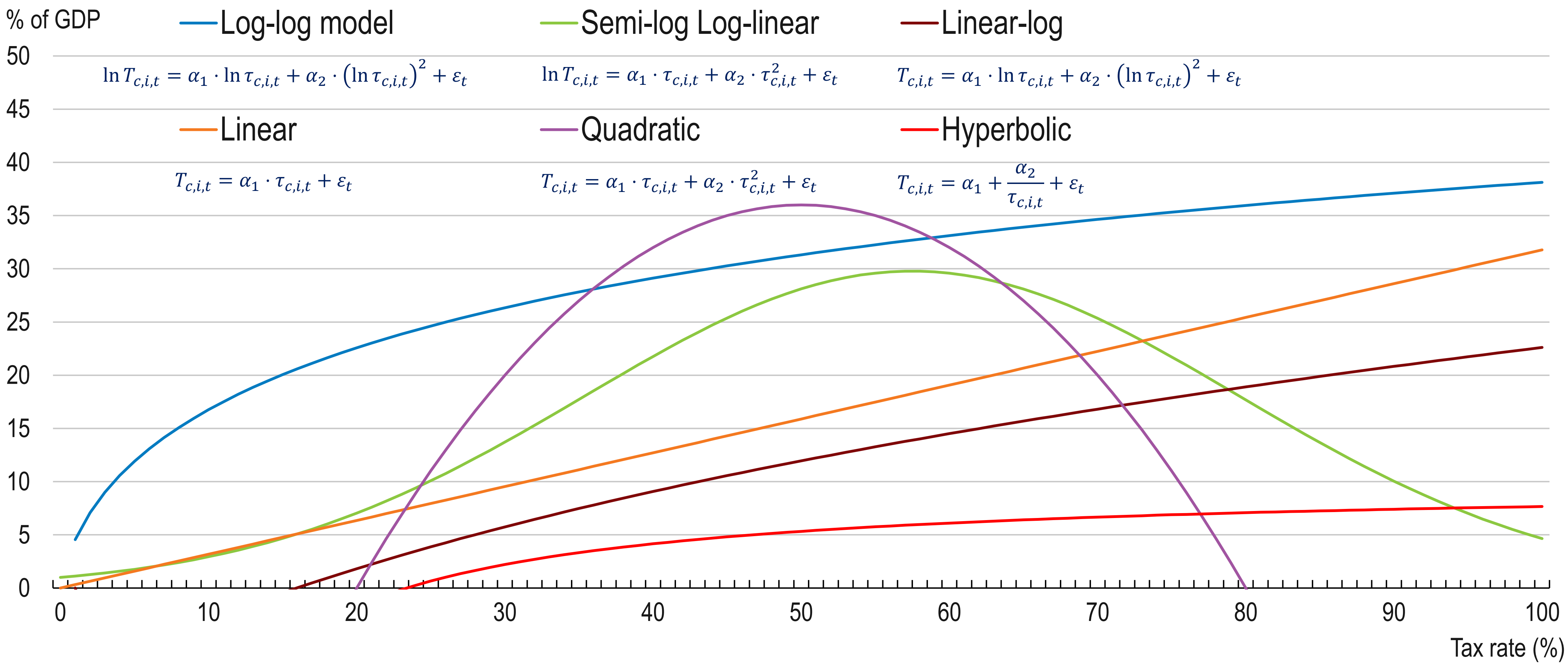
"Trade-off between tax rates and tax revenues"

$$R(\tau) = \tau \cdot B(\tau)$$

$$\Delta R(\Delta\tau) = \underbrace{\Delta\tau_1 \cdot B(\tau_0)}_{\substack{Arithmetic\ effect \\ (tax\ rate\ effect)}} + \underbrace{\tau_1 \cdot \Delta B(\tau_1)}_{\substack{Economic\ effect \\ (tax\ base\ effect)}}$$



Common used functional forms for the Laffer curve



Empirical challenges

- **Endogeneity:** Tax rates and taxable income jointly influence each other, making causal identification difficult.
- **Measurement error:** Tax avoidance, underreporting, and income definitions can bias elasticity estimates.
- **Heterogeneity:** Tax responsiveness differs across income groups, occupations, sectors, and taxpayer types.
- **Simultaneous reforms:** Multiple policy changes (rates, brackets, deductions, credits) complicate identification.
- **Data limitations:** Limited microdata, confidentiality, complex tax systems, and infrequent reforms hinder estimation.
- **Dynamic effects:** Behavioural responses often occur gradually rather than immediately.
- **General equilibrium effects:** Tax changes affect employment, investment, and consumption, indirectly influencing taxable income.
- **Fiscal externalities:** Income shifting, timing effects, institutional factors, and redistribution may bias estimated elasticities.

Derivation of the proposed model

$$R(\tau) = \tau \cdot B(\tau)$$

$$B(\tau_1) = B(\tau_0) \cdot \left(\frac{1-\tau_1}{1-\tau_0}\right)^{\zeta}$$

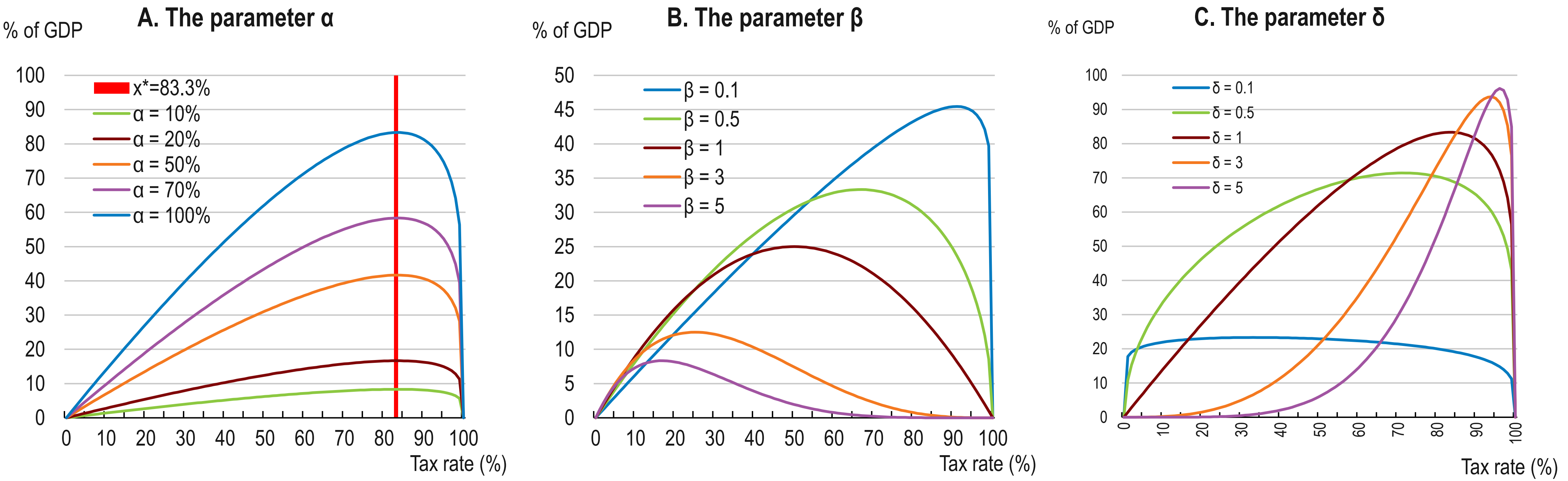
$$R(\tau) = A \cdot \tau \cdot (1-\tau)^{\zeta}$$

$$T_{c,i,t} = \alpha_{c,it} \cdot \left(\frac{\delta_{c,i,t} + \beta_{c,i,t}}{\delta_{c,i,t}}\right)^{\delta_{c,i,t}-1} \cdot \left(\frac{\delta_{c,i,t} + \beta_{c,i,t}}{\beta_{c,i,t}}\right)^{\beta_{c,i,t}} \cdot \tau_{c,i,t}^{\delta_{c,i,t}} \cdot (1-\tau_{c,i,t})^{\beta_{c,i,t}} + \varepsilon_t$$

$$\alpha \in [0,1], \quad \delta \geq 1, \quad \beta > 0$$

$$ln\, T_1 - ln\, T_0 = \underbrace{ln\, \tau_1 - ln\, \tau_0}_{Tax\ rate\ effect} + \underbrace{(\delta - 1) \cdot (ln\, \tau_1 - ln\, \tau_0)}_{GDP\ effect} + \underbrace{\beta \cdot [ln(1-\tau_1) - ln(1-\tau_0)]}_{Tax\ base\ effect}$$

Three parameters to understand



1) Tax revenue remains zero when the tax rate is either 0% or 100%. $(T(\tau = 0\%) = 0;\ T(\tau = 100\%) = 0)$

2) The curve reaches its maximum at $\tau^* = \frac{\delta}{\delta+\beta}$, with maximum revenue as a % of GDP given by $T(\tau^*) = \alpha \cdot \tau^*$

3) The maximum revenue $R(\hat{\tau})$ is equal to $\hat{\tau}_{c,i,t} = \frac{1}{1+\beta_{c,i,t}} < \frac{\delta_{c,i,t}}{\delta_{c,i,t}+\beta_{c,i,t}} = \tau^*_{c,i,t}$ because of the contractionary effect of taxation on GDP

4) The estimated tax-to-GDP ratio does not exceed the corresponding average tax rate $\frac{T(\tau)}{GDP} \leq \tau$

Nonlinear LASSO

$$\hat{\xi}_{LASSO} = \underset{\xi}{argmin}\left\{\sum\left[T_{c,i,t} - \alpha_{c,it}\cdot\left(\frac{\delta_{c,i,t}+\beta_{c,i,t}}{\delta_{c,i,t}}\right)^{\delta_{c,i,t}-1}\cdot\left(\frac{\delta_{c,i,t}+\beta_{c,i,t}}{\beta_{c,i,t}}\right)^{\beta_{c,i,t}}\cdot\tau_{c,i,t}^{\delta_{c,i,t}}\cdot\left(1-\tau_{c,i,t}\right)^{\beta_{c,i,t}}\right]^2\right\}$$

Subject to $\|\boldsymbol{\xi}\|_{\mathbf{1}} < \bar{\xi}$

$$\alpha_{c,it} = \frac{e^{\boldsymbol{\lambda}\cdot\boldsymbol{X}_{\boldsymbol{c},\boldsymbol{i},\boldsymbol{t}}}}{1+e^{\boldsymbol{\lambda}\cdot\boldsymbol{X}_{\boldsymbol{c},\boldsymbol{i},\boldsymbol{t}}}}$$

$$\delta_{c,i,t} = 1 + exp\left(\boldsymbol{\phi}\cdot\boldsymbol{X}_{\boldsymbol{c},\boldsymbol{i},\boldsymbol{t}}\right)$$

$$\beta_{c,i,t} = exp\left(\boldsymbol{\theta}\cdot\boldsymbol{X}_{\boldsymbol{c},\boldsymbol{i},\boldsymbol{t}}\right)$$

$$\boldsymbol{\xi} = [\boldsymbol{\lambda},\boldsymbol{\phi},\boldsymbol{\theta}]^{\boldsymbol{T}}$$

The data

DEPENDENT VARIABLES

- Personal income taxes revenue as a % of GDP (PIT)
- Corporate income taxes as a % of GDP (CIT)
- Environmental and Consumption taxes as a % of GDP (VAT)

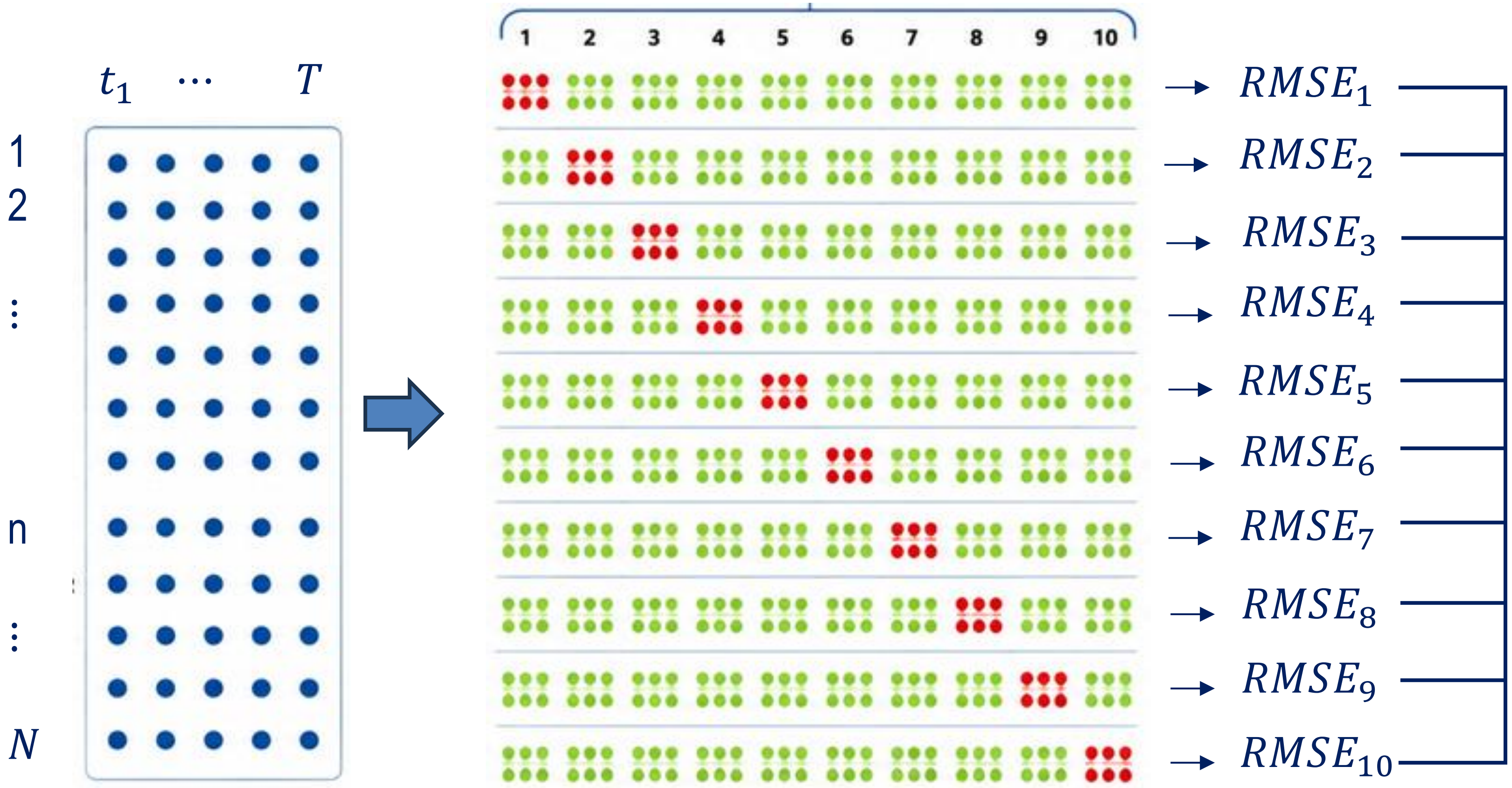
TAX RATES

- Average tax wedge personal income tax rate
- Corporate income tax on distributed profit
- Average value added tax rate

What determines the parameters?

1. Tax system
2. Macroeconomic conditions
3. Economic activity composition
4. Quality of institutions
5. Demography
6. Public expenditure composition
7. Public revenue composition
8. Geography

Cross-validation

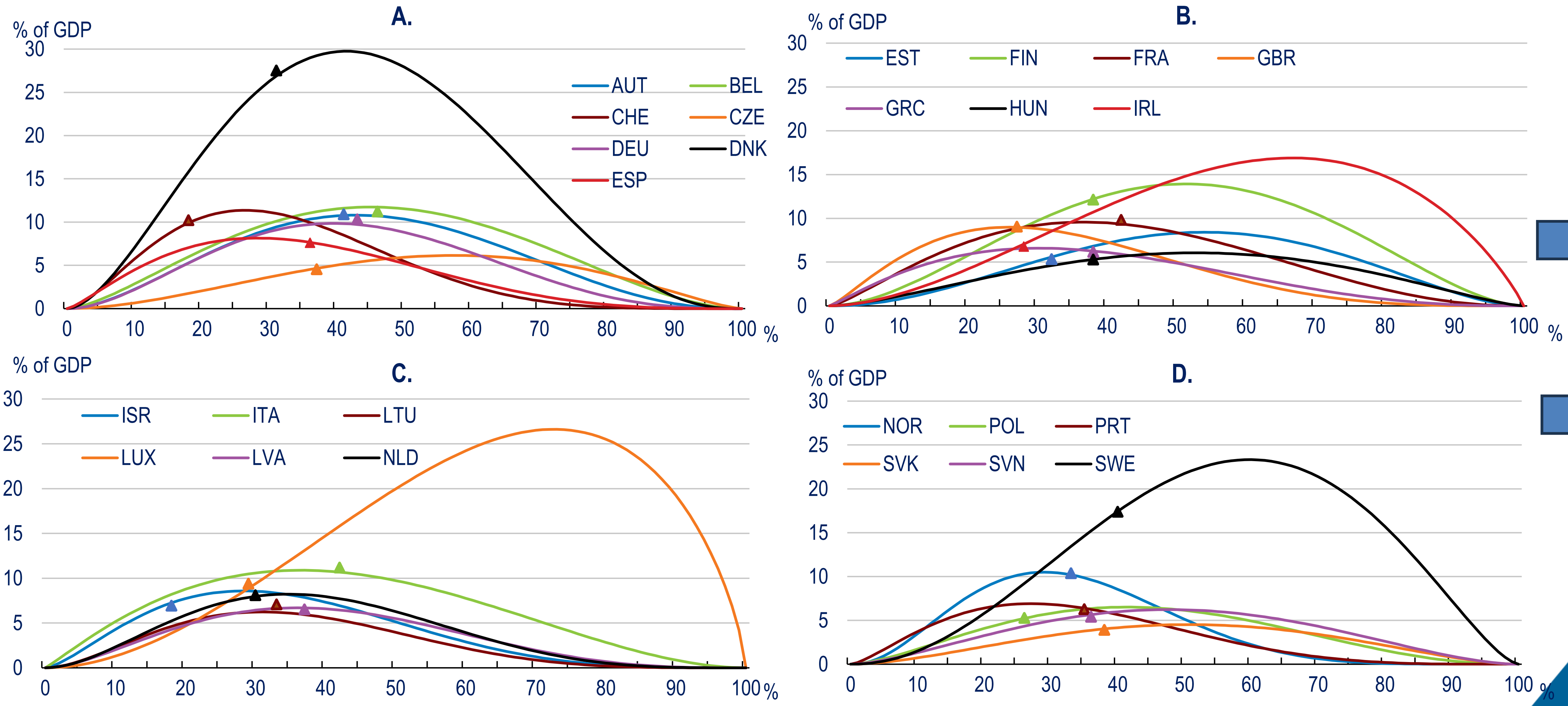


$$\overline{RMSE} = \sum_{k=1}^{10} RMSE_k$$

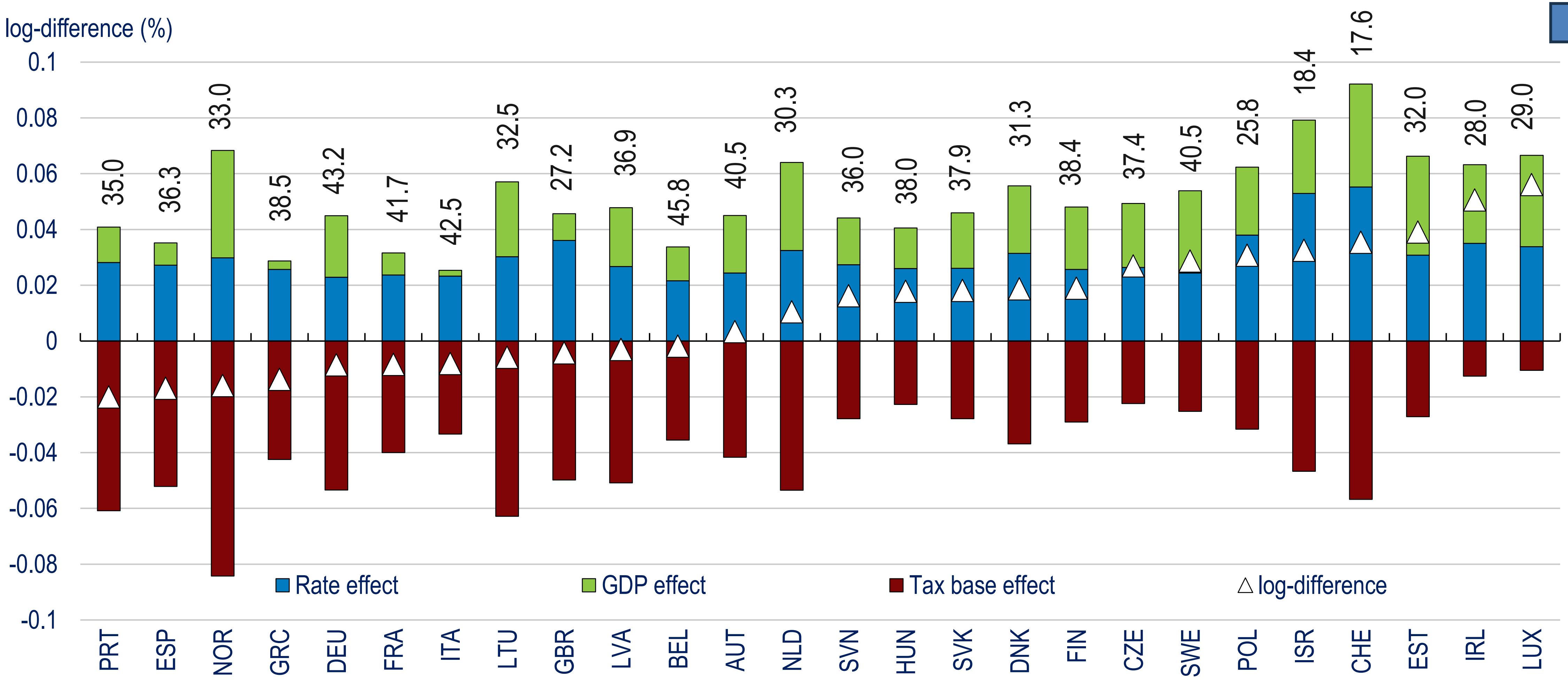
529 observations
26 countries
Years 2000-21

476 training
53 validation

PIT: heterogeneous Laffer curves



PIT: different marginal revenue effects

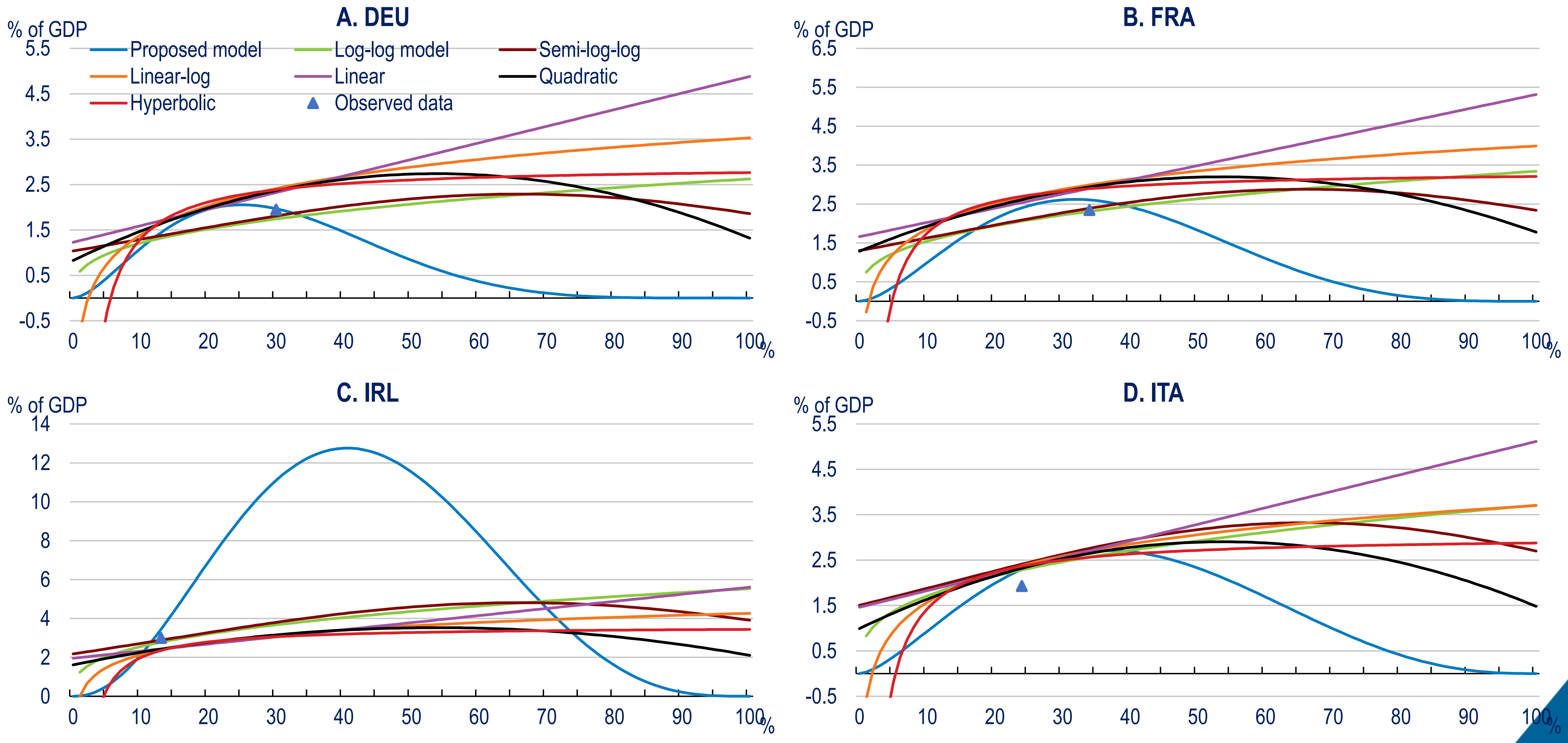


CIT: proposed versus conventional models

Estimated effect of a 1-pp tax rate increase across models

	CIT revenue as a % of GDP	Tax rate	Proposed model	Log-log model	Semi log-log model	Linear-log model	Linear model	Quadratic model	Hyperbolic model
LUX	6.0%	24.9%	-0.35	0.07	0.07	0.03	0.04	0.04	0.02
DEU	2.0%	29.9%	-0.03	0.02	0.02	0.03	0.04	0.03	0.02
NLD	3.5%	25.0%	-0.03	0.04	0.04	0.03	0.04	0.04	0.02
NOR	6.0%	22.0%	-0.03	0.06	0.07	0.04	0.04	0.04	0.03
FRA	2.3%	34.4%	-0.02	0.02	0.03	0.02	0.04	0.02	0.01
BEL	3.7%	29.6%	-0.01	0.04	0.04	0.03	0.04	0.03	0.02
DNK	3.2%	22.0%	0.00	0.04	0.05	0.04	0.04	0.04	0.03
AUT	3.0%	25.0%	0.03	0.03	0.04	0.03	0.04	0.04	0.02
ITA	1.9%	24.0%	0.06	0.03	0.03	0.03	0.04	0.04	0.03
PRT	3.1%	31.5%	0.07	0.03	0.04	0.03	0.04	0.03	0.02
GRC	2.2%	24.0%	0.09	0.02	0.03	0.03	0.04	0.04	0.03
ESP	2.0%	25.0%	0.10	0.03	0.03	0.03	0.04	0.04	0.02
FIN	2.5%	20.0%	0.11	0.04	0.04	0.04	0.04	0.04	0.04
GBR	2.4%	19.0%	0.12	0.04	0.04	0.04	0.04	0.05	0.04
ISR	3.1%	23.0%	0.16	0.05	0.05	0.04	0.04	0.04	0.03
LVA	0.2%	20.0%	0.16	0.02	0.02	0.04	0.04	0.04	0.04
SVK	3.0%	21.0%	0.16	0.04	0.05	0.04	0.04	0.04	0.03
SWE	4.3%	21.4%	0.16	0.06	0.06	0.04	0.04	0.04	0.03
CHE	3.1%	21.1%	0.17	0.04	0.04	0.04	0.04	0.04	0.03
LTU	1.5%	15.0%	0.17	0.03	0.03	0.06	0.04	0.05	0.07
HUN	1.3%	9.0%	0.17	0.06	0.04	0.09	0.04	0.06	0.18
EST	1.8%	20.0%	0.19	0.02	0.02	0.04	0.04	0.04	0.04
SVN	2.0%	19.0%	0.21	0.03	0.03	0.04	0.04	0.05	0.04
POL	2.2%	19.0%	0.21	0.04	0.04	0.04	0.04	0.05	0.04
CZE	3.6%	19.0%	0.26	0.06	0.06	0.04	0.04	0.05	0.04
IRL	3.0%	12.5%	0.47	0.08	0.06	0.07	0.04	0.05	0.10
RMSE		0.37	0.42	0.42	0.42	0.42	0.42	0.41	0.41

CIT: it is not only a matter of better data fit



Main findings

- The proposed model imposes the theoretical properties of the Laffer curve while allowing its shape to vary with countries' economic, demographic and institutional characteristics.
- This produces more heterogeneous estimates and improves out-of-sample performance.
- Structural reforms can increase revenue capacity and tax-base and GDP responsiveness.
- Large tax changes may themselves alter the determinants of the curve.
- In several OECD countries there is little scope to raise additional revenue through higher income tax rates, while there is comparatively more space in consumption taxation.

Thank you!