

Lecture 5: Taxable Income Elasticities and Broader Effects of Taxation

Stefanie Stantcheva

EC2450b – Spring 2019

TAXABLE INCOME ELASTICITIES

Modern public finance literature focuses on taxable income elasticities instead of hours/participation elasticities

Two main reasons:

- 1) What matters for policy is the total behavioral response to tax rates (not only hours of work but also occupational choices, avoidance, etc.)
- 2) Data availability: taxable income is precisely measured in tax return data

Recent overview of this literature: Saez-Slemrod-Giertz JEL'12

FEDERAL US INCOME TAX CHANGES

Tax rates change frequently over time

Biggest tax rate changes have happened at the top:

Reagan I: ERTA'81: top rate ↓ 70% to 50% (1981-1982)

Reagan II: TRA'86: top rate ↓ 50% to 28% (1986-1988)

Clinton: OBRA'93: top rate ↑ 31% to 39.6% (1992-1993)

Bush: EGTRRA '01: top rate ↓ 39.6% to 35% (2001-2003)

Obama '13: top rate ↑ 35% to 39.6%+3.8% (2012-2013)

Trump '17: top rate ↓ 37%+3.8% (2017-2018)

Taxable Income = Ordinary Income + Realized Capital Gains - Deductions
⇒ Each component can respond to *MTRs*

Historically, a 70 percent marginal tax rate is not unusual

The top marginal income tax rates from 1913 to 2018

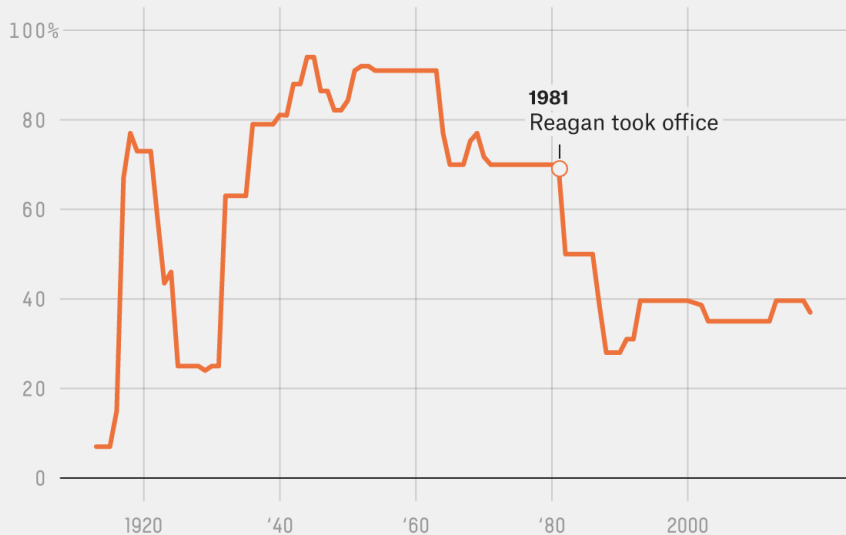


Table A1.
Top Federal Marginal Tax Rates

Year	Ordinary Income	Earned Income	Capital Gains	Corporate Income
	(1)	(2)	(3)	(4)
1952-1963	91.0	91.0	25.0	52
1964	77.0	77.0	25.0	50
1965-1967	70.0	70.0	25.0	48
1968	75.3	75.3	26.9	53
1969	77.0	77.0	27.9	53
1970	71.8	71.8	32.3	49
1971	70.0	60.0	34.3	48
1972-1975	70.0	50.0	36.5	48
1976-1978	70.0	50.0	39.9	48
1979-1980	70.0	50.0	28.0	46
1981	68.8	50.0	23.7	46
1982-1986	50.0	50.0	20.0	46
1987	38.5	38.5	28.0	40
1988-1990	28.0	28.0	28.0	34
1991-1992	31.0	31.0	28.0	34
1993	39.6	39.6	28.0	35
1994-2000	39.6	42.5	28.0	35
2001	39.1	42.0	20.0	35
2002	38.6	41.5	20.0	35
2003-2009	35.0	37.9	15.0	35

Notes: MTRs apply to top incomes. In some instances, lower income taxpayers may face higher MTRs because of income caps on payroll taxes or the so-called 33 percent "bubble" bracket following TRA 86. From 1952 to 1962, a 87% maximum average tax rate provision made the top marginal tax rate 87% instead of 91% for many very top income earners. From 1968 to 1970, rates include surtaxes. For earned income, MTRs include the Health Insurance portion of the payroll tax beginning with year 1994. Rates exclude the effect of phaseouts, which effectively raise top MTRs for many high-income filers. MTRs on realized capital gains are adjusted to reflect that, for some years, a fraction of realized gains were excluded from taxation. Since 2003, dividends are also tax favored with a maximum tax rate of 15%.

LONG-RUN EVIDENCE IN THE US

Goal: evaluate whether top **pre-tax** incomes respond to changes in one minus the marginal tax rate (=net-of-tax rate)

Focus is on pre-tax income before deductions and excluding realized capital gains

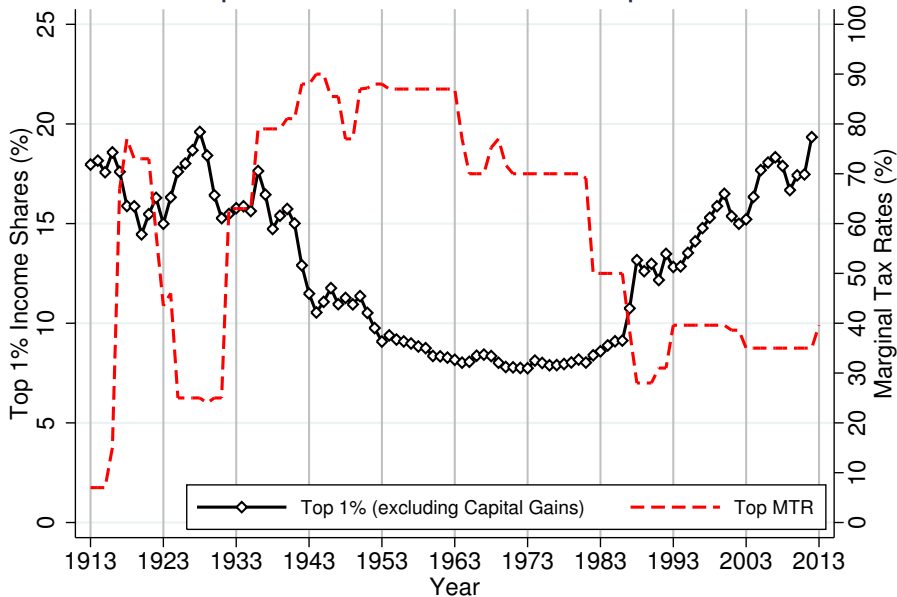
Pioneered by Feenberg-Poterba TPE'93 for period 1951-1990

Piketty-Saez QJE'03 estimate top income shares since 1913 [IRS tabulations for 1913-1959, IRS micro-files since 1960]

Saez TPE'04 proposes detailed analysis for 1960-2000 period using TAXSIM calculator at NBER linked to IRS micro-files

Piketty-Saez-Stantcheva AEJ'14 look at 1913-2010 period for the US

Top 1% Income Share and Top MTR



INCOME SHARE BASED ELASTICITY ESTIMATION

1) **Tax Reform Episode:** Compare top **pre-tax** income shares at t_0 (before reform) and t_1 (after reform)

$$e = \frac{\log sh_{t_1} - \log sh_{t_0}}{\log(1 - \tau_{t_1}) - \log(1 - \tau_{t_0})}$$

where sh_t is top income share and τ_t is the average MTR for top group

Identification assumption: absent tax change, $sh_{t_0} = sh_{t_1}$

2) **Full Time Series:** Run regression:

$$\log sh_t = \alpha + e \cdot \log(1 - \tau_t) + \varepsilon_t$$

and adding time controls to capture non-tax related top income share trends

ID assumption: non-tax related changes in $sh_t \perp \tau_t$

Table 1.

Elasticity estimates using top income share time series

	Top 1%	Next 9%
	(1)	(2)
A. Tax Reform Episodes		
1981 vs. 1984 (ERTA 1981)	0.60	0.21
1986 vs. 1988 (TRA 1986)	1.36	-0.20
1992 vs. 1993 (OBRA 1993)	0.45	
1991 vs. 1994 (OBRA 1993)	-0.39	
B. Full Time Series 1960-2006		
No time trends	1.71 (0.31)	0.01 (0.13)
Linear time trend	0.82 (0.20)	-0.02 (0.02)
Linear and square time trends	0.74 (0.06)	-0.05 (0.03)
Linear, square, and cube time trends	0.58 (0.11)	-0.02 (0.02)

Notes: Estimates in panel A are obtained using series from Figure 1 and using the formula $e = [\log(\text{income share after reform}) - \log(\text{income share before reform})] / [\log(1 - \text{MTR after reform}) - \log(1 - \text{MTR before reform})]$

LONG-RUN EVIDENCE IN THE US

- 1) Clear correlation between top incomes and top income rates both in several short-run tax reform episodes and in the long-run [but hard to assess long-run tax causality]
 - 2) Correlation largely absent below the top 1% (such as the next 9%)
 - 3) Top income shares sometimes do not respond to large tax rate cuts [e.g., Kennedy Tax Cuts of early 1960s]
- 2) and 3) suggest that context matters (such as opportunities to respond / avoid taxes matter), response not due to a universal labor supply elasticity

SPECIFIC TAX REFORM STUDIES

Literature initially developed by analyzing specific tax reforms (instead of full time series)

Lindsey JpubE'87 analyzes ERTA'81 using **repeated cross-section** tax data and finds large elasticities

Feldstein JPE'95 uses **panel** tax data to study TRA'86

Goolsbee JPE'00 uses **executive compensation** data to study OBRA'93

Gruber-Saez JpubE'02 uses 1979-1990 **panel** tax data

Many other studies in the US and abroad (survey by Saez-Slemrod-Giertz JEL'12)

FELDSTEIN JPE'95: METHODOLOGY

Feldstein (1995) estimates the effect of TRA86 on taxable income for top earners using **panel** tax data

1) Constructs three income groups M (Medium), H (High), HH (Highest) based on before reform income in 1985

2) Looks at how incomes and MTRs evolve from 1985 to 1988 for individuals in each group using panel: forms DD estimates

$$\hat{\epsilon} = \frac{\Delta \log(z^H) - \Delta \log(z^M)}{\Delta \log(1 - \tau^H) - \Delta \log(1 - \tau^M)}$$

where z^H , z^M and τ^H , τ^M are income and MTRs of the H and M groups

TABLE 1

RESPONSE OF TAXABLE INCOME OF NONAGED MARRIED TAXPAYERS TO CHANGES IN MARGINAL TAX RATES BETWEEN 1985 AND 1988

1985 MARGINAL TAX RATE	1985 AGI (\$000) (1)	OBSERVATIONS (2)	PERCENTAGE CHANGES OF				
			Net of Tax Rate (3)	Adjusted Full AGI (4)	Adjusted AGI Excluding Capital Gains (5)	Adjusted Taxable Income (6)	Adjusted Taxable Income Plus Gross Loss (7)
22	30.7	800	9.0	9.4	8.4	13.6	13.4
25	36.1	909	13.3	4.5	2.4	3.5	3.7
28	42.7	713	16.3	3.9	4.7	6.0	5.0
33	51.5	771	8.7	2.2	2.2	2.5	2.5
38	67.5	345	16.1	8.0	8.1	9.6	8.8
42	94.3	152	24.1	18.8	14.7	22.0	22.3
45	126.9	45	30.9	12.4	14.8	18.5	15.3
49	177.7	35	41.2	27.1	29.6	42.7	33.9
50	479.0	22	44.0	18.4	70.6	92.4	51.1
22-38		3,538	12.2	5.1	4.6	6.2	6.4
42-45		197	25.6	17.0	14.7	21.0	20.3
49-50		57	42.2	21.3	53.7	71.6	44.8

NOTE.—All observations pertain to married taxpayers under age 65 who filed joint tax returns for 1985 and 1988 with no age exemption in 1988. Taxpayers who created a subchapter S corporation between 1985 and 1988 are eliminated from the sample.

TABLE 2
ESTIMATED ELASTICITIES OF TAXABLE INCOME WITH RESPECT TO NET-OF-TAX RATES

Taxpayer Groups Classified by 1985 Marginal Rate	Net of Tax Rate (1)	Adjusted Taxable Income (2)	Adjusted Taxable Income Plus Gross Loss (3)
	Percentage Changes, 1985–88		
1. Medium (22–38)	12.2	6.2	6.4
2. High (42–45)	25.6	21.0	20.3
3. Highest (49–50)	42.2	71.6	44.8
	Differences of Differences		
4. High minus medium	13.4	14.8	13.9
5. Highest minus high	16.6	50.6	24.5
6. Highest minus medium	30.0	65.4	38.4
	Implied Elasticity Estimates		
7. High minus medium		1.10	1.04
8. Highest minus high		3.05	1.48
9. Highest minus medium		2.14	1.25

NOTE.—The calculations in this table are based on observations for married taxpayers under age 65 who filed joint tax returns for 1985 and 1988 with no age exemption in 1988. Taxpayers who created a subchapter S corporation between 1985 and 1988 are eliminated from the sample.

FELDSTEIN JPE'95: RESULTS

Results: Feldstein obtains very high elasticities (above 1) for top earners

⇒ US was on the wrong side of the Laffer curve for the rich

⇒ Laffer rate $\tau = 1/(1 + a \cdot e) = 1/(1 + 2 \cdot 1) = 33\%$ Cutting top tax rate from 50% to 28% raised revenue

FELDSTEIN JPE'95: ISSUES

- 1) Non-tax related **changes in inequality** [same criticism as top share analysis]: panel helps only if inequality changes due to arrival of new people
- 2) Short-term vs. Long-term response [same criticism as top share analysis]
- 3) **Mean reversion**: rich people in year t tend to revert to the mean in year $t + 1 \Rightarrow$ Panel analysis introduces downward bias in e [when $\tau \downarrow$ for rich]
- 4) **Very small sample** in panel data [57 tax filers in HH group]
[Auten-Carroll RESTAT'99 use larger Treasury panel data and find smaller elasticity 0.65]

In net, not clear panel data adds value relative to repeated-cross-section

FELDSTEIN JPE'95: ISSUES

5) DD can give very biased results when elasticity differs across groups:

Example: (a) M group has $e^M = 0$ so that $\Delta \log(z^M) = 0$ and that H group has $e^H = e > 0$ so that $\Delta \log(z^H) = e \Delta \log(1 - \tau^H)$.

Suppose that $\Delta \log(1 - \tau^M) = 0.5 \cdot \Delta \log(1 - \tau^H)$.

Then, the estimated elasticity

$$\hat{e}^{DD} = e \Delta \log(1 - \tau^H) / [\Delta \log(1 - \tau^H) - \Delta \log(1 - \tau^M)] = 2e$$

In Feldstein JPE'95: Simple Difference $\Delta \log(z) / \Delta \log(1 - \tau)$ uniformly smaller than DD

⇒ Better to focus on a single group as in top share analysis than on the comparison with lower income group control

GRUBER AND SAEZ JPUBE'02

Generalization of Feldstein JPE'95 using IV regression analysis

Use panel data from 1979-1990 on all tax changes available rather than a single reform

Model: $z_{it} = z_{it}^0 \cdot (1 - \tau_{it})^e$ where z_{it}^0 is potential income (if MTR=0), e is elasticity

$$\log \left(\frac{z_{it+3}}{z_{it}} \right) = \alpha + e \cdot \log \left(\frac{1 - \tau_{it+3}}{1 - \tau_{it}} \right) + \varepsilon_{it}$$

τ_{it+3} and ε_{it} are correlated [because $\tau_{it+3} = T'_{t+3}(z_{it+3})$]

Instrument: predicted change in MTR assuming income stays constant:
 $\log[(1 - \tau_{it+3}^p)/(1 - \tau_{it})]$ where $\tau_{it+3}^p = T'_{t+3}(z_{it})$

Isolates changes in tax law ($T_t(\cdot)$) as the only source of variation in tax rates

Table 4
Basic elasticity results^a

Income controls	None		Log income		Log income 10-piece spline	
	Broad income (1)	Taxable income (2)	Broad income (3)	Taxable income (4)	Broad income (5)	Taxable income (6)
Elasticity	-0.300 (0.120)	-0.462 (0.194)	0.170 (0.106)	0.611 (0.144)	0.120 (0.106)	0.400 (0.144)
Dummy for marrieds	-0.008 (0.010)	-0.062 (0.018)	0.045 (0.014)	0.049 (0.023)	0.050 (0.012)	0.055 (0.021)
Dummy for singles	-0.037 (0.012)	-0.053 (0.019)	-0.034 (0.013)	-0.032 (0.022)	-0.036 (0.013)	-0.027 (0.021)
Log(income) control			-0.083 (0.015)	-0.167 (0.021)		

Source: Gruber and Saez 2002

GRUBER AND SAEZ JPUBE'02

Find an elasticity of roughly 0.3-0.4 BUT results are very fragile
[Saez-Slemrod-Giertz JEL'12]

- 1) Sensitive to exclusion of low incomes
- 2) Sensitive to controls for mean reversion
- 3) Subsequent studies find smaller elasticities using data from other countries [Kleven-Schultz AEJ-EP'14 for Denmark]
- 4) Bundles together small tax changes and large tax changes: if individuals respond only to large changes in short-medium run, then estimated elasticity is too low [Chetty et al. QJE'11]

KLEVEN AND SCHULTZ AEJ-EP'14

Key Advantages:

a) Use full population of tax returns in Denmark since 1980 (large sample size, panel structure, many demographic variables, stable inequality)

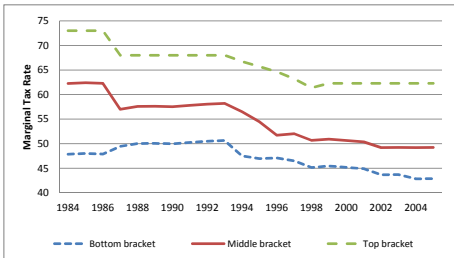
b) A number of reforms changing tax rates differentially across three income brackets and across tax bases (capital income taxed separately from labor income)

c) Show compelling visual DD-evidence of tax responses around the 1986 large reform:

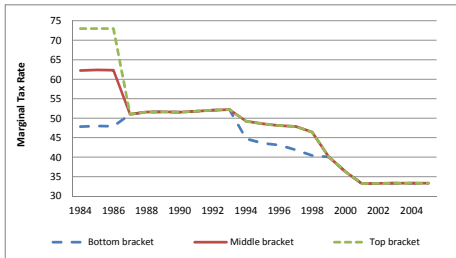
Define treatment and control group in year 1986 (pre-reform), follow the same group in years before and years after the reform (panel analysis)

Figure 2. Two Decades of Danish Tax Reform

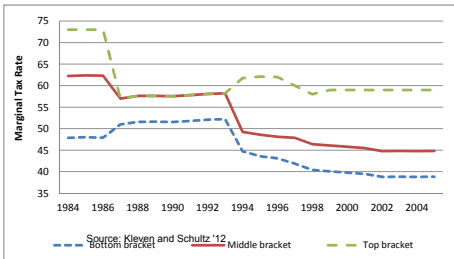
Panel A. Marginal Tax Rate on Labor Income



Panel B. Marginal Tax Rate on Negative Capital Income



Panel C. Marginal Tax Rate on Positive Capital Income



Panel D. Share of Taxpayers in the Three Tax Brackets

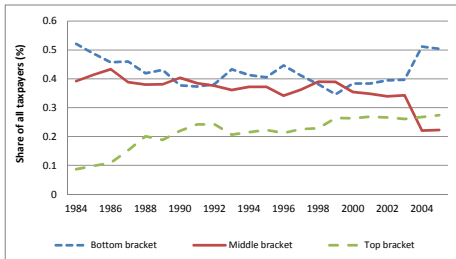
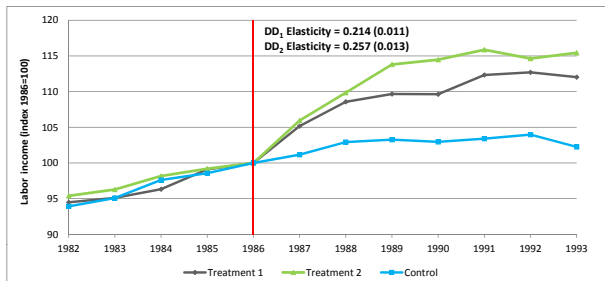


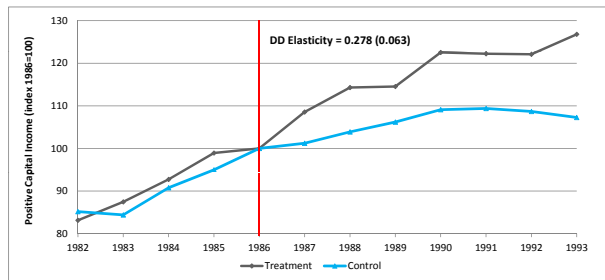
Figure 6. Graphical Evidence on the Effects of the 1987-reform on Taxable Income

Source: Kleven and Schultz '12

Panel A. Labor Income



Panel B. Positive Capital Income



KLEVEN AND SCHULTZ AEJ-EP'14

Key Findings:

- a) Small labor income elasticity (.1)
 - b) bigger capital income elasticities (.2-.3)
 - c) bigger elasticities for large reforms
 - d) modest income shifting between labor and capital in Denmark (likely because top rates on labor and capital are carefully aligned)
- ⇒ Danish tax system optimized to have broad base and few avoidance opportunities

FISCAL EXTERNALITIES

Tax changes due to tax avoidance often generate **fiscal externalities**

A **Fiscal externality** is a change in tax revenue that occurs in any tax base z^B other than z due to the behavioral response to the tax change in the initial base z

(1) z^B can be a different tax base in the same time period (such as corporate income tax base) \Rightarrow **Income shifting**

(2) z^B can be the same tax base in a different time period (such as future income) \Rightarrow **Inter-temporal Substitution**

Efficiency and optimal tax analysis depend on effect on **total** tax revenue so critical to identify fiscal externalities

Inter-Temporal Substitution: Realized Capital Gains

Realized capital gains occur when individual sells asset at a higher price than buying price

Individuals have flexibility in the timing of asset sales and capital gains realizations

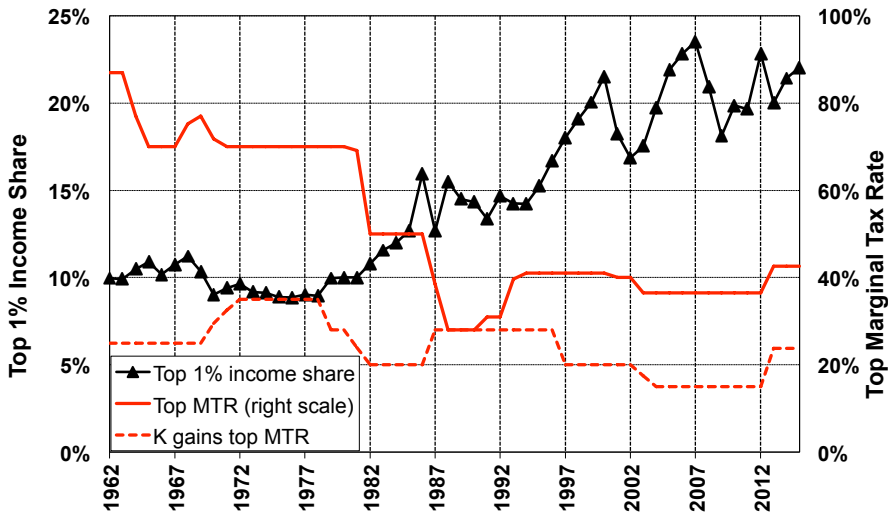
TRA'86 lowered the top tax rate on ordinary income from 50% to 28% but increased the top tax rate on realized capital gains from 20% to 28%

2013: tax rate on KG increased from 15% to 20%+3.8% (Saez TPE'17 proposes simple analysis)

⇒ Surge in capital gains realizations in 1986 and 2012 [and depressed capital gains in 1987 and 2013]

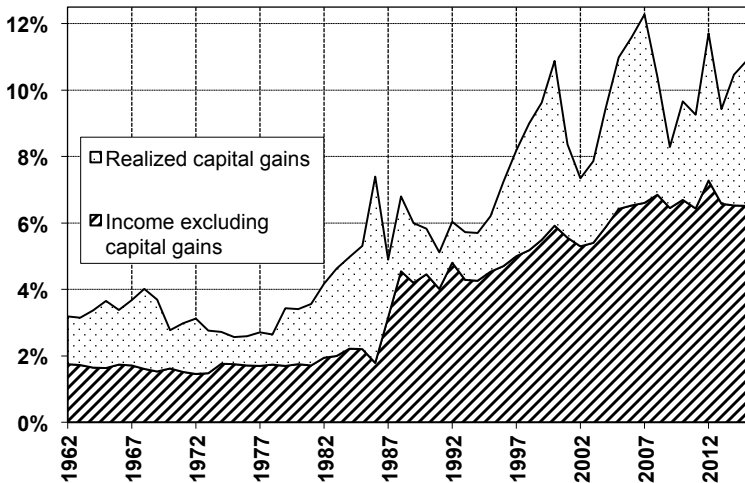
⇒ Short-term elasticity is very large but long-term elasticity is certainly much smaller

Top 1% pre-tax income share and top tax rates



Source: Top 1% income share: Piketty and Saez, 2003 updated to 2015, series including realized capital gains. Top MTR include Federal individual tax + uncapped FICA payroll tax.

US Top 0.1% Income Share and Composition



Source: Piketty and Saez, 2003 updated to 2015. Series based on pre-tax cash market income including realized capital gains, and always excluding government transfers.

INTER-TEMPORAL SUBSTITUTION: STOCK-OPTIONS

Goolsbee JPE'00 analyzes CEO pay around the 1993 Clinton top tax rate increase ↑ [from 31% in 1992 to 39.6% in 1993 announced in late 1992] on executive pay

Finds a strong re-timing response through stock-option exercise (executive can choose the timing of their stock-option exercises)

⇒ Large short-term response due to re-timing, small long-term response

STOCK OPTIONS

Major form of compensation of US top executives. Theoretical goal is to motivate executives to increase the value of the company (stock price $P(t)$)

Stock-option is granted at date t_0 allow executives to buy N company shares at price $P(t_0)$ on or after t_1 (in general $t_1 - t_0 \simeq 3 - 5$ years = vesting period)

Executive exercise option at (chosen) time $t_2 \geq t_1$: pays $N \cdot P(t_0)$ to get shares valued $N \cdot P(t_2)$. Exercise profit $N[P(t_2) - P(t_0)]$ (considered and taxed as wage income in the US)

After t_2 , executive owns N shares, eventually sold at time $t_3 \geq t_2$: realized capital gain $N[P(t_3) - P(t_2)]$ (taxed as capital gains)

TABLE 2
 AVERAGE COMPENSATION BY TYPE FOR HIGH-INCOME EXECUTIVES
 (in Thousands)

	1991	1992	1993	1994	1995
Taxable income	911	1,153	974	965	1,173
Salary	347	336	336	351	373
Bonus	198	207	241	284	330
LTIP payout	57	72	57	64	89
Options exercised	268	496	293	235	381
Other income (nontaxed)	36	37	66	54	78

SOURCE.—Author's calculations for executives with permanent income greater than \$275,000 per year.

INCOME SHIFTING: CORPORATE AND INDIVIDUAL TAX BASE

Businesses can be organized as **corporations** or **unincorporated businesses** [also called **pass-through** entities]

Corporate profits are first taxed by corporate tax [tax rate $\tau_c = 21\%$]

Net-of-tax profits are taxed again at rate τ_{distrib} when finally distributed to shareholders. Two distribution options:

a) dividends [tax rate $\tau_d = 20\%$ today]

b) retained profits increase stock price: shareholders realize capital gains when finally selling the stock [tax rate $\tau_{cg} = 20\%$]
But distributions can be deferred so that $\tau_{\text{distrib}} \ll \tau_d, \tau_{cg}$

For **unincorporated businesses** (sole proprietorships, partnerships, S-corporations) profits are taxed directly and solely as individual income (tax rate $\tau_i = 37\%$ top MTR or even 31% with 20% business profit deduction since 2018)

CORPORATE AND INDIVIDUAL TAX BASE

Corporate form best if $(1 - \tau_c) \cdot (1 - \tau_{\text{distrib}}) > 1 - \tau_i$

US fed taxes in 2018: $\tau_c = 21\%$, $\tau_{cg} = \tau_d = 20\%$, (but $\tau_{\text{distrib}} \ll 20\%$ if distribution deferred), $\tau_i = 37\%$ or 31%

After 2018 Trump change: corporate form is best, especially if wealthy business owner can defer distribution

Pre 2018, $\tau_c = 35\%$ and $\tau_i = 39.6\% \Rightarrow$ individual form better

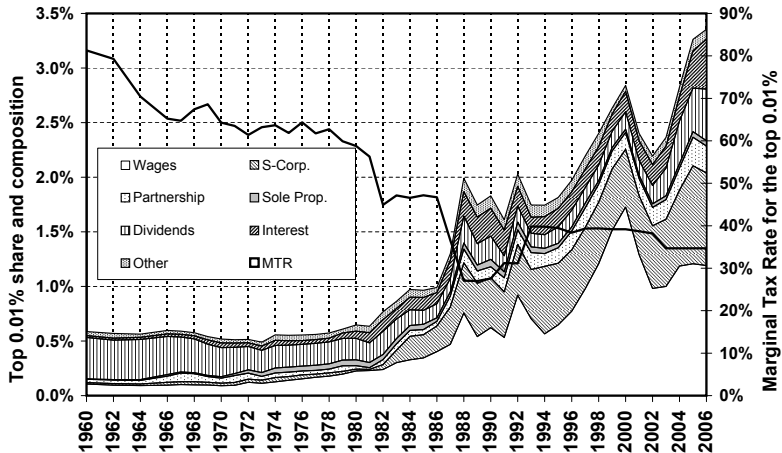
\Rightarrow wealthy people likely to incorporate their businesses in '18+

Before TRA'86 (and especially before ERTA'81), top individual rate τ_i was much higher so corporate form was best

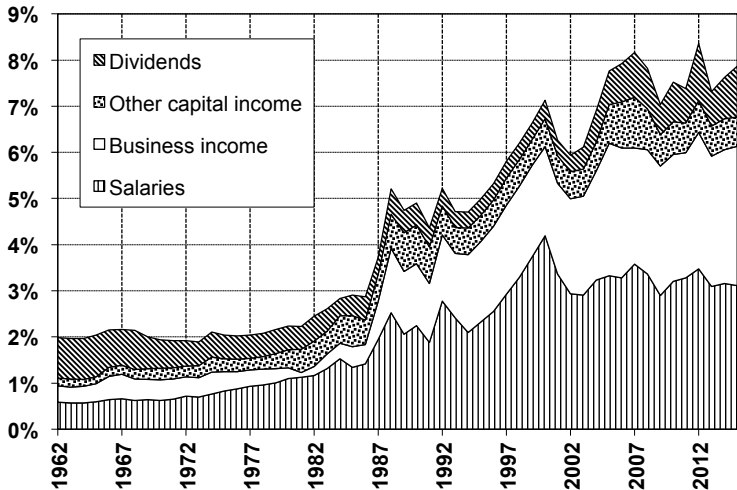
Shifts from corporate to individual base increases business profits at the expense of dividends and realized capital gains

Large part of TRA'86 response is due to such shifting

The Top 0.01% US Income Share, Composition, and MTR



US Top 0.1% Income Share and Composition (excl. K gains)



Source: Piketty and Saez, 2003 updated to 2015. Series based on pre-tax cash market income excluding realized capital gains, and always excluding government transfers.

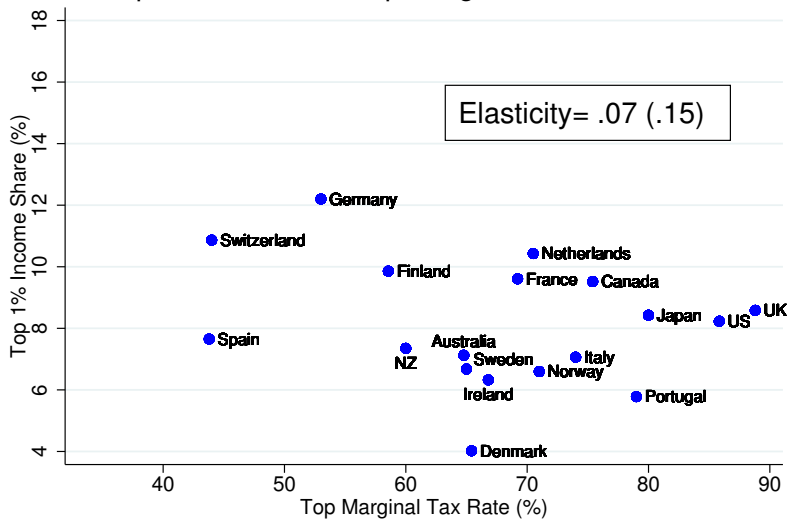
TOP RATES AND TOP INCOMES INTERNATIONAL EVIDENCE

Piketty-Saez-Stantcheva (2014)

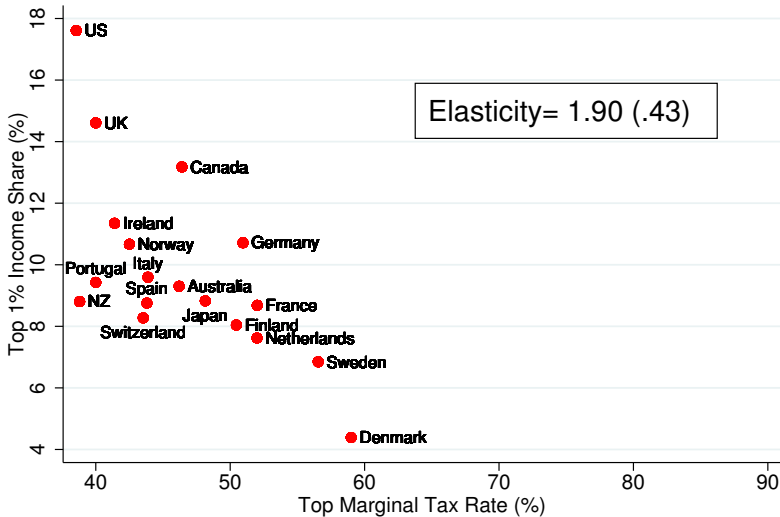
- 1) Use pre-tax top 1% income share data from 18 OECD countries since 1960 using the **World Top Incomes Database**
- 2) Collect data and compute the top (statutory) individual income tax rates using OECD data [including both central and local income taxes].

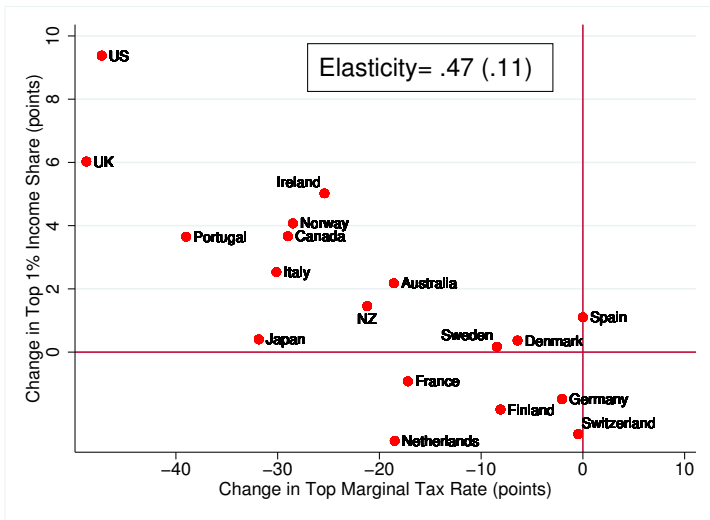
Plot top 1% pre-tax income share against top MTR in 1960-4, in 2005-9, and 1960-4 vs. 2005-9

A. Top 1% Share and Top Marginal Tax Rate in 1960–4



B. Top 1% Share and Top Marginal Tax Rate in 2005–9





Change in Top Tax Rate and Top 1% Share, 1960-4 to 2005-9

Top tax rates and top 1% income share 1960-2009

Table 2: International Evidence on Top Income Elasticities

	All 18 countries and fixed periods			Bootstrapping period and country set		
	1960-2010	1960-1980	1981-2010	Median	5th percentile	95th percentile
	(1)	(2)	(3)	(4)	(5)	(6)

A. Effect of the Top Marginal Income Tax Rate on Top 1% Income Share

Regression: $\log(\text{Top 1\% share}) = a + e \cdot \log(1 - \text{Top MTR}) + \varepsilon$

No controls	0.324 (0.034)	0.163 (0.039)	0.803 (0.053)	0.364 (0.043)	0.128 (0.085)	0.821 (0.032)
Time trend control	0.375 (0.042)	0.182 (0.030)	0.656 (0.056)	0.425 (0.045)	0.191 (0.091)	0.761 (0.032)
Country fixed effects	0.314 (0.025)	0.007 (0.039)	0.626 (0.044)	0.267 (0.035)	0.008 (0.070)	0.595 (0.026)
Number of observations	774	292	482	286	132	516

ECONOMIC EFFECTS OF TAXING THE TOP 1%

Strong empirical evidence that **pre-tax** top incomes are affected by top tax rates

3 potential scenarios with very different policy consequences

1) Supply-Side: Top earners work less and earn less when top tax rate increases ⇒ Top tax rates should not be too high

2) Tax Avoidance/Evasion: Top earners avoid/evade more when top tax rate increases

⇒ a) Eliminate loopholes, b) Then increase top tax rates

3) Rent-seeking: Top earners extract more pay (at the expense of the 99%) when top tax rates are low ⇒ High top tax rates are desirable

Real changes vs. tax Avoidance? Charitable giving

Test using charitable giving behavior of top income earners (Saez TPE '17)

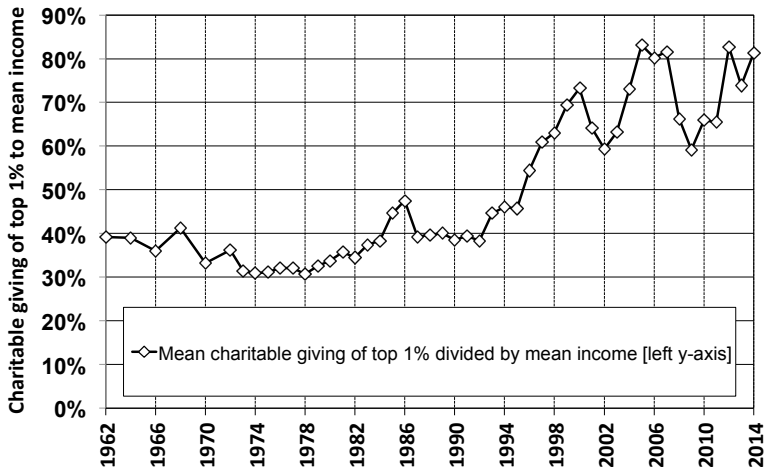
Because charitable is tax deductible, incentives to give are stronger when tax rates are higher

Under the tax avoidance scenario, reported incomes and reported charitable giving should move in opposite directions

Empirically, charitable giving of top income earners has grown in close tandem with top incomes

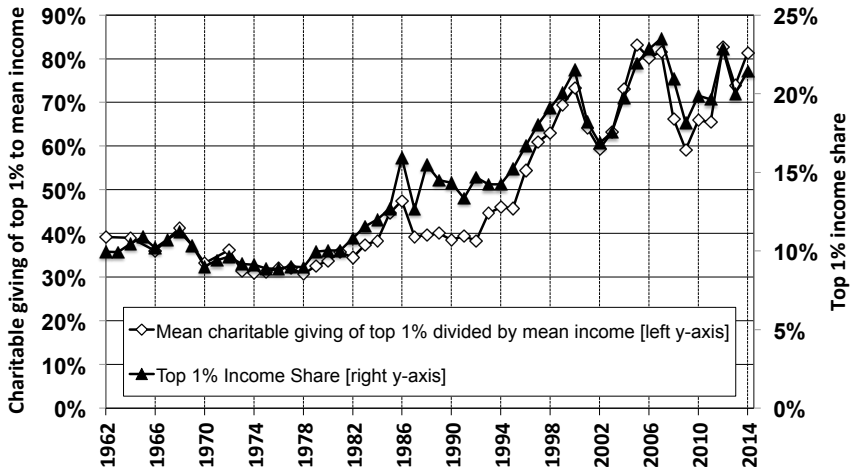
⇒ Incomes at the top have grown for real

Charitable Giving of Top 1% Income Earners



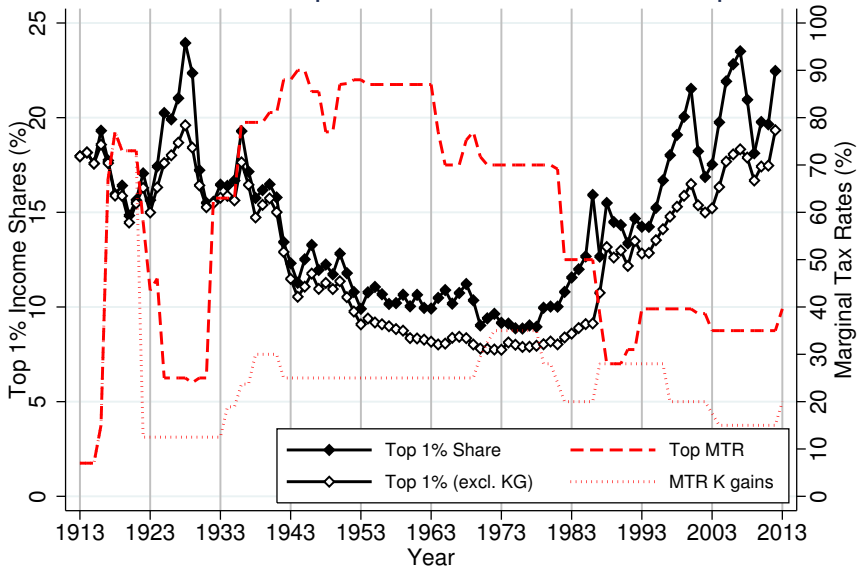
Source: The figure depicts average charitable giving of top 1% incomes (normalized by average income per family) on the left y-axis.

Charitable Giving of Top 1% Income Earners



Source: The figure depicts average charitable giving of top 1% incomes (normalized by average income per family) on the left y-axis. For comparison, the figure reports the top 1% income share (on the right y-axis).

Tax Avoidance: Top 1% Income Shares and Top MTR



Supply-Side or Rent-Seeking? (Piketty-Saez-Stantcheva)

Correlation between **pre-tax** top incomes and top tax rates

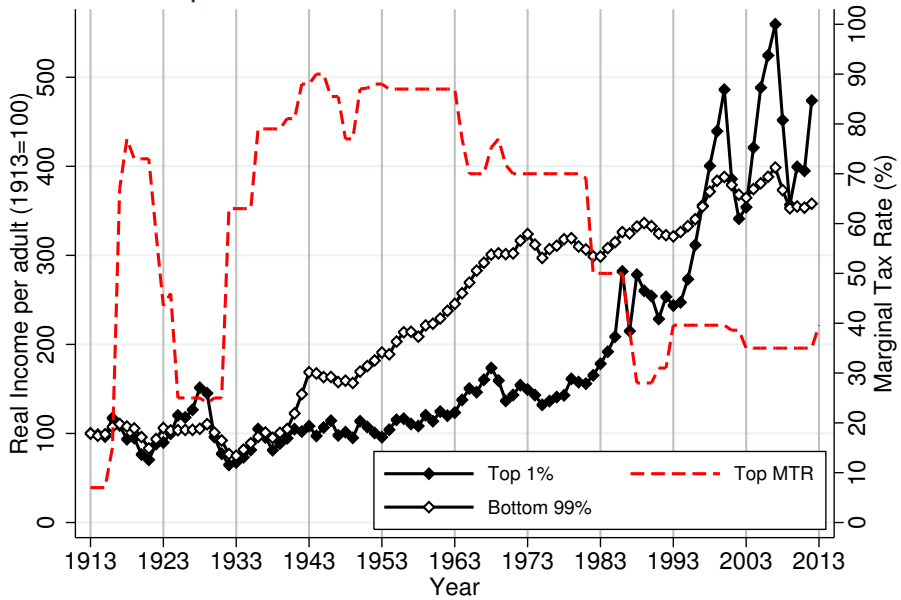
If rent-seeking: growth in top 1% incomes should come at the expense of bottom 99% (and conversely)

Two macro-preliminary tests:

- 1) In the US, top 1% incomes grow slowly from 1933 to 1975 and fast afterwards. Bottom 99% incomes grow fast from 1933 to 1975 and slowly afterwards ⇒ Consistent with rent-seeking effects
- 2) Look at cross-country correlation between economic growth and top tax rate cuts ⇒ No correlation supports trickle-up

Two micro-tests using CEO pay data (panel in the U.S. + international cross-section).

Top 1% and Bottom 99% Income Growth



CEO Pay in the US: Empirical Strategy

- Effect of general performance on pay (OLS):

$$pay_{it} = \beta * p_{it} + \gamma_i + \chi_t + \alpha_X * X_{it} + \varepsilon_{it}$$

pay_{it} : CEO pay in firm i at time t , p_{it} : performance measure, γ_i : firm FE, χ_t : time FE, X_{it} : CEO controls (age, tenure).

- Effect of luck performance on pay (IV):

- Stage:** Effect of luck on performance measure

$$p_{it} = b * p_{luck,it} + g_i + c_t + \alpha_X * X_{it} + e_{it} \quad (1)$$

$p_{luck,it}$: luck measure (asset-weighted average industry performance).
Part of performance due to (observable) luck \hat{p}_{it} = prediction from (1).

- Stage:** Estimate sensitivity of pay to predictable changes in p_{it} :

$$y_{it} = \beta_{luck} * \hat{p}_{it} + \gamma_i + \chi_t + \alpha_X * X_{it} + \varepsilon_{it}$$

If $\beta_{luck} \neq 0$: pay for luck.

If $\beta_{luck} \geq \beta$: no filtering at all of luck component.

CEO Pay in the US: Luck and performance measures

- **Performance measures:**
 1. Net Income
 2. Shareholder Wealth (log)
- **Measure of pay:** Total Pay
- **Measure of luck:** Mean asset-weighted performance of other firms in industry.
- **Data:** Forbes 800 + Execucomp, COMPUSTAT-CRSP.
- **Years:** **1970-2010**
- Analysis repeated for high tax period (pre-1986) and low tax period (post-1987) to study effect of tax rates.

Table 3: US CEO Pay Evidence, 1970-2010

Firm performance measure	Log(net income)			Log(stock-market value)		
	Log(CEO pay)	Log(CEO pay)	Log(industry level workers pay)	Log(CEO pay)	Log(CEO pay)	Log(industry level workers pay)
Outcome (LHS variable)			Industry level			Industry level
OLS vs. IV	OLS	Industry luck IV	OLS regression	OLS	Industry luck IV	OLS regression
	(1)	(2)	(3)	(4)	(5)	(6)
A. Effect of firm performance on log-pay in high-top tax rate period (1970-1986)						
Firm performance (RHS variable)	0.23*** (0.013)	0.34*** (0.072)	0.00 (0.010)	0.28*** (0.022)	0.22* (0.123)	0.00 (0.015)
Number of observations	8,632	8,503	890	9,005	8,865	898
B. Effect of firm performance on log-pay in low-top tax rate period (1987-2010)						
Firm performance (RHS variable)	0.27*** (0.012)	0.70*** (0.148)	-0.02 (0.020)	0.37*** (0.021)	0.95*** (0.309)	-0.02 (0.023)
Number of observations	14,914	14,697	1,422	17,775	17,593	1,443
C. Test for difference between low- and high- top tax rate periods						
Difference Panel B - Panel A	0.04***	0.36*	-0.019	0.09***	0.72**	-0.023
p-value of difference	0.01	0.06	0.440	0.00	0.05	0.46

CEO Pay in the US: Results

- Incomplete filtering of luck component in CEO pay: $\beta_{luck} \neq 0$.
- Pay for luck is large and almost no filtering: $\beta_{luck} \geq \beta$.
- Pay for luck much stronger in low tax period, consistent with bargaining model.

Could pay for luck be consistent with optimal contracting view?

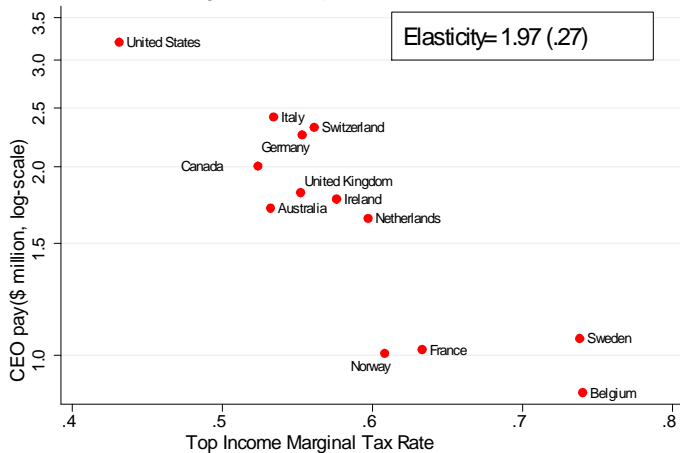
- CEO incentivized to predict luck shocks? But why reward average performance (2SLS uses no between firm variation) and why reward less when MTR higher?
- Maybe not bargaining but impossibility to filter out luck?
 - Badly governed firms exhibit more pay for luck (BM and our results - not shown for sake of time).
 - Still means there is a lot of "non-deserved" pay!
- Most important criticism: CEO human capital value increasing in industry performance?
 - Strikingly, workers' wages show no 'pay for luck' (columns 3 and 6).

- Fernandez et. al. (2012) data:
 - Compensation (BoardEx + Execucomp)
 - Stock ownership (LionShares)
 - Firm Performance (Worldscope and Datastream)
 - Firm governance (various sources)
1. Does controlling for firm performance still leave CEO pay dependent on top tax rates?
 2. Does effect of top tax rate on CEO pay depend on firm governance?

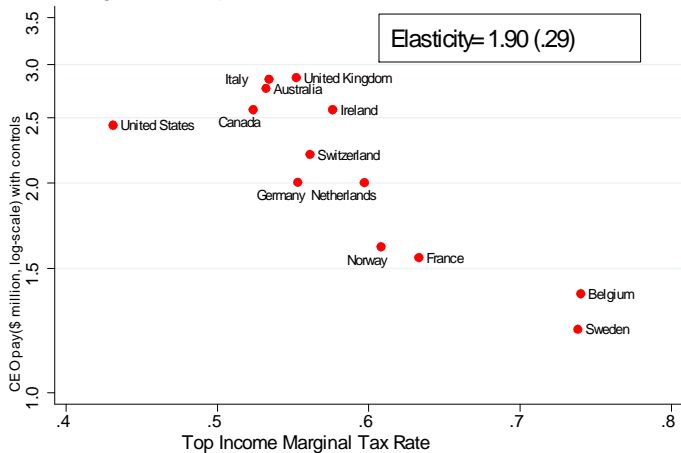
International CEO pay: Reward for Performance

- Does controlling for firm performance still leave CEO pay dependent on top tax rates?
 - In supply side story, should not (increase in labor effort translates into firm performance).
 - In bargaining story, additional negative effect of top tax rate on CEO pay through rent-seeking.
 - Requires very comprehensive set of measures of firm performance (use firm sales, stock market return and std dev, leverage, Tobin's q)
- Result:
 - Without controls for firm performance, elasticity 1.97 of CEO pay to top retention rate
 - With controls: elasticity 1.9.
 - Almost none of the effect of top MTR goes through firm performance (i.e., productive CEO effort?)

A Average CEO compensation



B. Average CEO compensation with controls



- Does effect of top tax rate on CEO pay depend on firm governance?
 - In badly governed firms, pay should react more to tax rates as both real supply side response and bargaining response add up.
- Index of (good) governance :
 - Insider ownership
 - Institutional ownership
 - Whether CEO also chairman of board
 - Average number of outside board positions of board members
 - Fraction of independent board directors.
- Result:
 - Retention rate increases CEO pay, but less so in well-governed firms
 - Huge elasticity of bonuses and equity pay to tax rates, very small one for salaries (extraction easier through discretionary bonuses and equity pay?)

Table 4: International CEO Pay Evidence

Outcome (LHS variable)	Log(CEO pay)	Log(CEO pay)	Log(CEO pay)	Log(CEO pay)	Log(CEO salary)	Log(CEO bonus and equity pay)
	(1)	(2)	(3)	(4)	(5)	(6)
Explanatory variables (RHS variables)						
log(1-Top MTR)	1.97*** (0.27)	1.90*** (0.286)	1.92*** (0.336)	1.90*** (0.328)	0.35* (0.189)	4.68*** (0.782)
Governance index			-0.10*** (0.020)	-0.19*** (0.038)	-0.02 (0.072)	-0.26 (0.201)
log(1-Top MTR)*Governance index				-0.13** (0.057)	0.06 (0.089)	-0.03 (0.281)
Firm and CEO controls	no	yes	yes	yes	yes	yes
Number of observations	2,959	2,844	2,711	2,711	2,691	2,711

INTERNATIONAL MIGRATION

Public debate concern that top skilled individuals move to low tax countries (e.g., in EU context) or low tax states (within US Federation, see Moretti-Wilson AER17, Young et al. 16)

Migration concern bigger in public debate than supply-side concern within a country

Relatively little work on tax induced international migration of top skilled workers

Hard to get data but interesting variation due to proliferation of special low tax schemes for highly paid foreigners in Europe

Kleven-Landais-Saez AER'13 look at **football players** in Europe (highly mobile group, many tax reforms) ⇒ Find significant migration responses to taxes after European football market was de-regulated in '95

Akcigit-Baslandze-Stantcheva AER'16 look at **innovators** (using patent data) mobility and find significant tax effects for top innovators

KLEVEN-LANDAIS-SAEZ-SCHULTZ QJE'14

Exploit the 1991 Danish tax scheme: high earnings immigrants ($\geq 103,000$ Euros/year) taxed at flat 25% rate (instead of regular tax with top 59% rate) for 3 years

Use population wide Danish tax data and DD strategy: compare immigrants above eligibility earnings threshold (treatment) to immigrants below threshold (control)

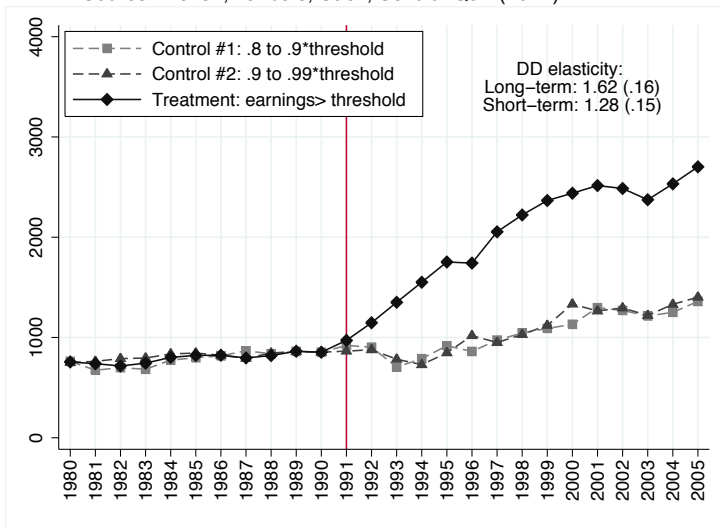
Key Finding: Scheme doubles the number of highly paid foreigners in Denmark relative to controls

⇒ Elasticity of migration with respect to the net-of-tax rate above one (much larger than the within country elasticity of earnings)

⇒ Tax coordination will be key to preserve progressive taxation in the EU

Figure 1 : Total number of foreigners in different income groups

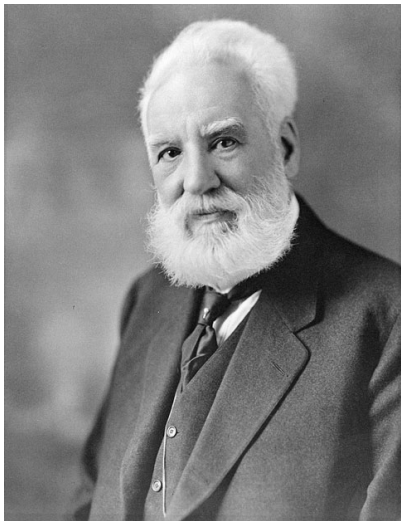
Source: Kleven, Landais, Saez, Schultz QJE (2014)



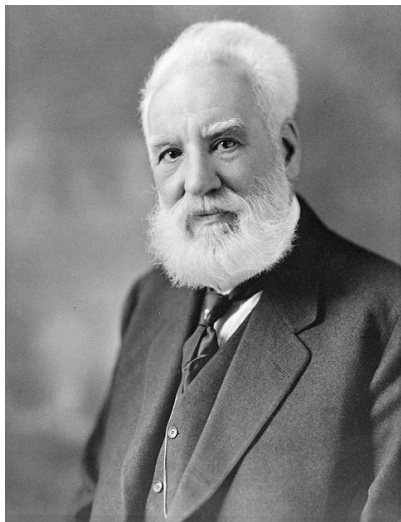
Control 1= annualized income between .8 and .9 of threshold

Control 2= annualized income between .9 and .995 of threshold.

Alexander G. Bell



Alexander G. Bell



- Inventor of the telephone (1876).
- Created Bell Telephone Company (1877).
- By 1886: more than 150,000 people in U.S. own telephones.

James L. Kraft



James L. Kraft



- Invented a pasteurization technique for cheese and established his company.
- Created Kraft Foods Inc.
- His company grew into a conglomerate responsible for creating some of the United States' most popular food products and employing more than 100,000 people.

Ralph Baer



Ralph Baer



- Created TV game unit with paddle controls.
- Today, the video gaming industry is worth \$66 billion.

Introduction

- ... and the list goes on.
- In addition to being very prolific inventors, these innovators had something else in common:
- They were all **immigrants**.
- What determines the patterns of migration of highly skilled people?

Introduction

- ... and the list goes on.
- In addition to being very prolific inventors, these innovators had something else in common:
- They were all immigrants.
- What determines the patterns of migration of highly skilled people?

Introduction

- ... and the list goes on.
- In addition to being very prolific inventors, these innovators had something else in common:
- They were all **immigrants**.
- What determines the patterns of migration of highly skilled people?

Introduction

- ... and the list goes on.
- In addition to being very prolific inventors, these innovators had something else in common:
- They were all **immigrants**.
- What determines the patterns of migration of highly skilled people?

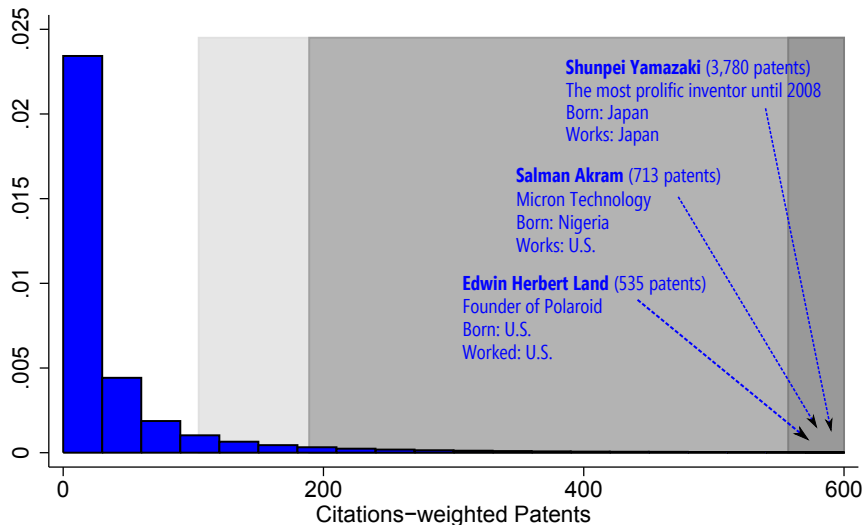
Study the Effects of Taxes on Migration using Patent Data

- Use a **unique international panel data** to overcome challenges:
 - ▶ **Patent data** from the USPTO and EPO, 1977-2000.
 - ▶ Track inventors in 8 big patenting countries: CA, CH, DE, FR, IT, JP, UK, US through residential addresses.
- Study effects of **top tax rates** on “**superstar**” inventors’ locations.
- Patent data gives direct measures of inventor quality.
- Detailed controls for *counterfactual* earnings in each potential location.

Three levels of analysis:

- ① Macro country-year level migration flows (country-by-year variation).
- ② Country case studies (quasi-experimental variation from reforms).
- ③ Micro inventor level location choice model
(differential impact of top MTR within country-year.
Inventor quality → ↑ propensity to be treated).

Superstar Inventors in a Highly Skewed Quality Distribution



Three Data sources: DID, EPO, PCT

- Inventors: employees, researchers, self-employed.
- “Assignee” is legal owner (firm or individual), can be \neq from inventor. Focus on employees.

Main Data: Disambiguated Inventor Data

- USPTO: 4.2 million patent records, 3.1 million inventors in 1975-2010.
- 18% of worldwide direct patent filings (26% of all patents).
- Disambiguated names with residential addresses (Lai *et al.*, 2012).

Additional Data 1: European Patent Office (EPO) data

- Very recent disambiguation, higher representation of EU patents.

Additional Data 2: Patent Cooperation Treaty (PCT) data

▶ USPTO Stats

▶ EPO Stats

▶ Details

Three Data sources: DID, EPO, PCT

- Inventors: employees, researchers, self-employed.
- “Assignee” is legal owner (firm or individual), can be \neq from inventor. Focus on employees.

Main Data: Disambiguated Inventor Data

- USPTO: 4.2 million patent records, 3.1 million inventors in 1975-2010.
- 18% of worldwide direct patent filings (26% of all patents).
- Disambiguated names with residential addresses (Lai *et al.*, 2012).

Additional Data 1: European Patent Office (EPO) data

- Very recent disambiguation, higher representation of EU patents.

Additional Data 2: Patent Cooperation Treaty (PCT) data

▶ USPTO Stats

▶ EPO Stats

▶ Details

Three Data sources: DID, EPO, PCT

- Inventors: employees, researchers, self-employed.
- “Assignee” is legal owner (firm or individual), can be \neq from inventor. Focus on employees.

Main Data: Disambiguated Inventor Data

- USPTO: 4.2 million patent records, 3.1 million inventors in 1975-2010.
- 18% of worldwide direct patent filings (26% of all patents).
- Disambiguated names with residential addresses (Lai *et al.*, 2012).

Additional Data 1: European Patent Office (EPO) data

- Very recent disambiguation, higher representation of EU patents.

Additional Data 2: Patent Cooperation Treaty (PCT) data

Three Data sources: DID, EPO, PCT

- Inventors: employees, researchers, self-employed.
- “Assignee” is legal owner (firm or individual), can be \neq from inventor. Focus on employees.

Main Data: Disambiguated Inventor Data

- USPTO: 4.2 million patent records, 3.1 million inventors in 1975-2010.
- 18% of worldwide direct patent filings (26% of all patents).
- Disambiguated names with residential addresses (Lai *et al.*, 2012).

Additional Data 1: European Patent Office (EPO) data

- Very recent disambiguation, higher representation of EU patents.

Additional Data 2: Patent Cooperation Treaty (PCT) data

Inventor Quality Measures and Ranking

Patent quality increases inventor income, directly and *indirectly*.

Quality measures

(dynamic and lagged)

- 1 Citations-weighted patents (benchmark)
- 2 Patent count
- 3 Average citations per patent
- 4 Max citations per patent
- 5 Patent breadth (claims-weighted patents)
- 6 Impact breadth (# tech classes citing patent).

▸ Correlations

▸ Patent breadth, breadth of impact

Inventor Ranking

- Group countries by patenting intensity (robust):
 1. U.S., 2. JP, 3. EU + CA
- Assign inventors to group based on home country.

→ Dynamic, Persistent, Life-time ranking

Inventor Quality Measures and Ranking

Patent quality increases inventor income, directly and *indirectly*.

Quality measures

(dynamic and lagged)

- 1 Citations-weighted patents (benchmark)
- 2 Patent count
- 3 Average citations per patent
- 4 Max citations per patent
- 5 Patent breadth (claims-weighted patents)
- 6 Impact breadth (# tech classes citing patent).

▸ Correlations

▸ Patent breadth, breadth of impact

Inventor Ranking

- Group countries by patenting intensity (robust):
 1. U.S., 2. JP, 3. EU + CA
- Assign inventors to group based on home country.

→ Dynamic, Persistent, Life-time ranking

Inventor Quality Measures and Ranking

Patent quality increases inventor income, directly and *indirectly*.

Quality measures

(dynamic and lagged)

- 1 Citations-weighted patents (benchmark)
- 2 Patent count
- 3 Average citations per patent
- 4 Max citations per patent
- 5 Patent breadth (claims-weighted patents)
- 6 Impact breadth (# tech classes citing patent).

▶ Correlations

▶ Patent breadth, breadth of impact

Inventor Ranking

- Group countries by patenting intensity (robust):
 1. U.S., 2. JP, 3. EU + CA
- Assign inventors to group based on home country.

→ Dynamic, Persistent, Life-time ranking

Inventor Quality Measures and Ranking

Patent quality increases inventor income, directly and *indirectly*.

Quality measures

(dynamic and lagged)

- 1 Citations-weighted patents (benchmark)
- 2 Patent count
- 3 Average citations per patent
- 4 Max citations per patent
- 5 Patent breadth (claims-weighted patents)
- 6 Impact breadth (# tech classes citing patent).

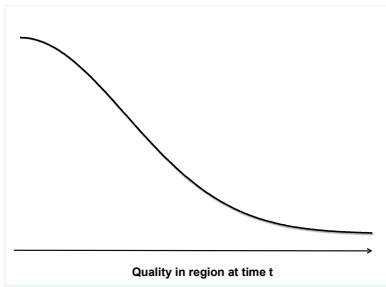
▸ Correlations

▸ Patent breadth, breadth of impact

→ Dynamic, Persistent, Life-time ranking

Inventor Ranking

- Group countries by patenting intensity (robust):
 1. U.S., 2. JP, 3. EU + CA
- Assign inventors to group based on home country.



Inventor Quality Measures and Ranking

Patent quality increases inventor income, directly and *indirectly*.

Quality measures

(dynamic and lagged)

- 1 Citations-weighted patents (benchmark)
- 2 Patent count
- 3 Average citations per patent
- 4 Max citations per patent
- 5 Patent breadth (claims-weighted patents)
- 6 Impact breadth (# tech classes citing patent).

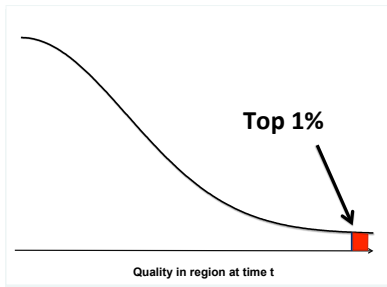
▸ Correlations

▸ Patent breadth, breadth of impact

→ Dynamic, Persistent, Life-time ranking

Inventor Ranking

- Group countries by patenting intensity (robust):
 1. U.S., 2. JP, 3. EU + CA
- Assign inventors to group based on home country.



Inventor Quality Measures and Ranking

Patent quality increases inventor income, directly and *indirectly*.

Quality measures

(dynamic and lagged)

- 1 Citations-weighted patents (benchmark)
- 2 Patent count
- 3 Average citations per patent
- 4 Max citations per patent
- 5 Patent breadth (claims-weighted patents)
- 6 Impact breadth (# tech classes citing patent).

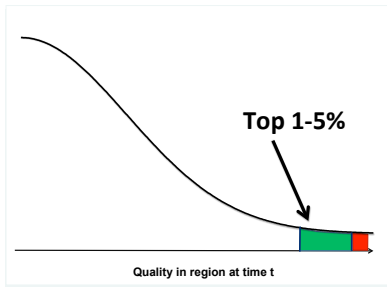
▸ Correlations

▸ Patent breadth, breadth of impact

→ Dynamic, Persistent, Life-time ranking

Inventor Ranking

- Group countries by patenting intensity (robust):
 1. U.S., 2. JP, 3. EU + CA
- Assign inventors to group based on home country.



Inventor Quality Measures and Ranking

Patent quality increases inventor income, directly and *indirectly*.

Quality measures

(dynamic and lagged)

- 1 Citations-weighted patents (benchmark)
- 2 Patent count
- 3 Average citations per patent
- 4 Max citations per patent
- 5 Patent breadth (claims-weighted patents)
- 6 Impact breadth (# tech classes citing patent).

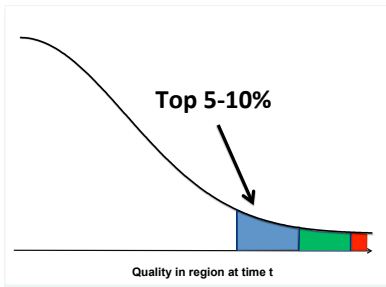
▸ Correlations

▸ Patent breadth, breadth of impact

→ Dynamic, Persistent, Life-time ranking

Inventor Ranking

- Group countries by patenting intensity (robust):
 1. U.S., 2. JP, 3. EU + CA
- Assign inventors to group based on home country.



Inventor Quality Measures and Ranking

Patent quality increases inventor income, directly and *indirectly*.

Quality measures

(dynamic and lagged)

- 1 Citations-weighted patents (benchmark)
- 2 Patent count
- 3 Average citations per patent
- 4 Max citations per patent
- 5 Patent breadth (claims-weighted patents)
- 6 Impact breadth (# tech classes citing patent).

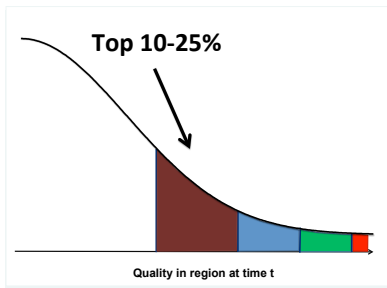
▸ Correlations

▸ Patent breadth, breadth of impact

→ Dynamic, Persistent, Life-time ranking

Inventor Ranking

- Group countries by patenting intensity (robust):
 1. U.S., 2. JP, 3. EU + CA
- Assign inventors to group based on home country.



Inventor Quality Measures and Ranking

Patent quality increases inventor income, directly and *indirectly*.

Quality measures

(dynamic and lagged)

- 1 Citations-weighted patents (benchmark)
- 2 Patent count
- 3 Average citations per patent
- 4 Max citations per patent
- 5 Patent breadth (claims-weighted patents)
- 6 Impact breadth (# tech classes citing patent).

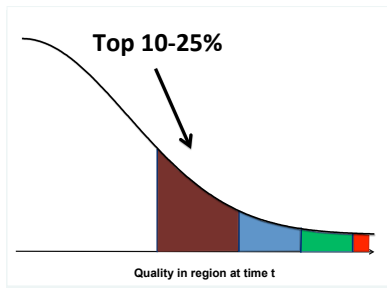
▸ Correlations

▸ Patent breadth, breadth of impact

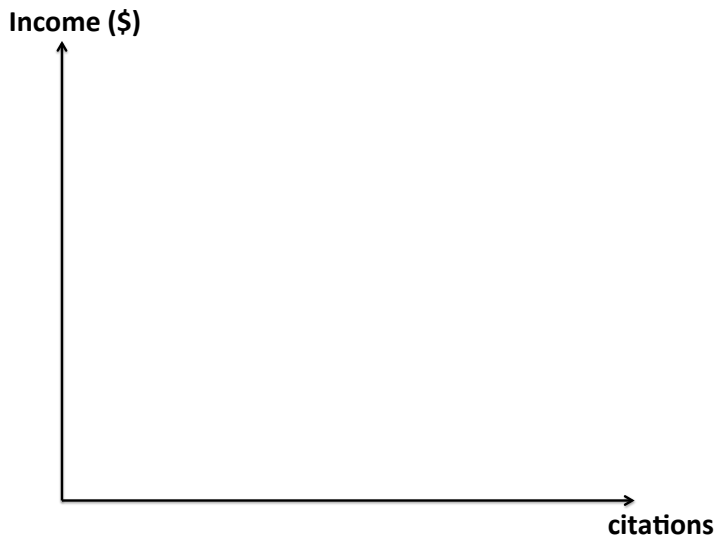
→ Dynamic, Persistent, Life-time ranking

Inventor Ranking

- Group countries by patenting intensity (robust):
 1. U.S., 2. JP, 3. EU + CA
- Assign inventors to group based on home country.

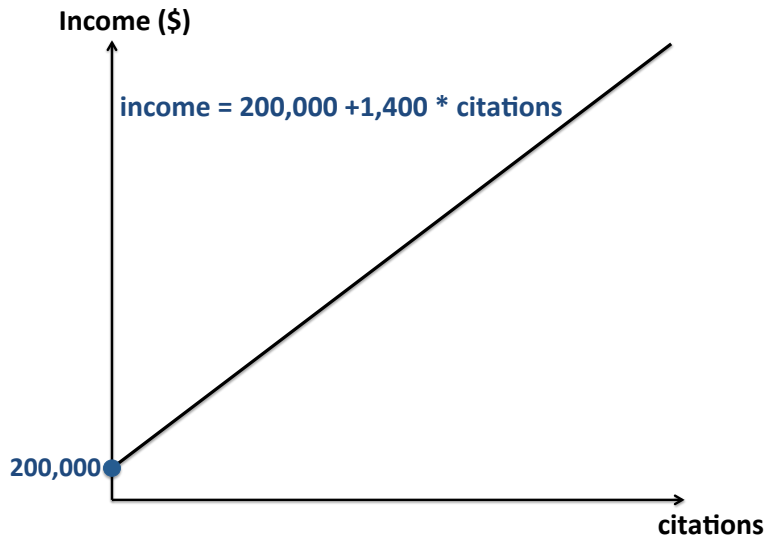


Link between Inventor Quality and Income in IRS data

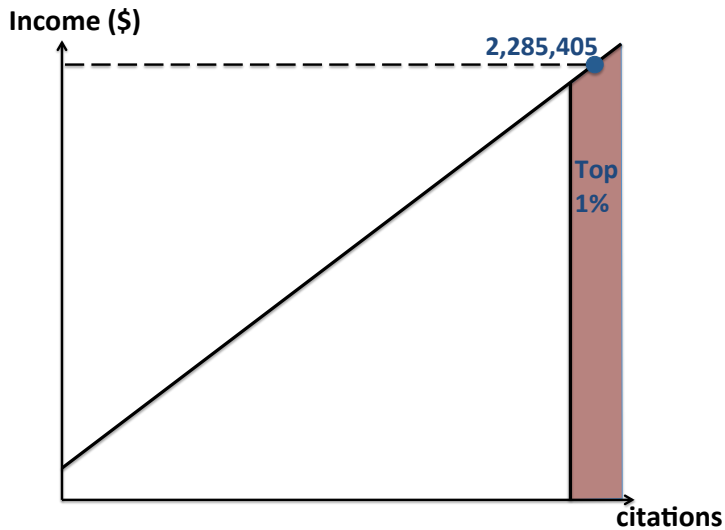


Source: Bell *et al.* (2015).

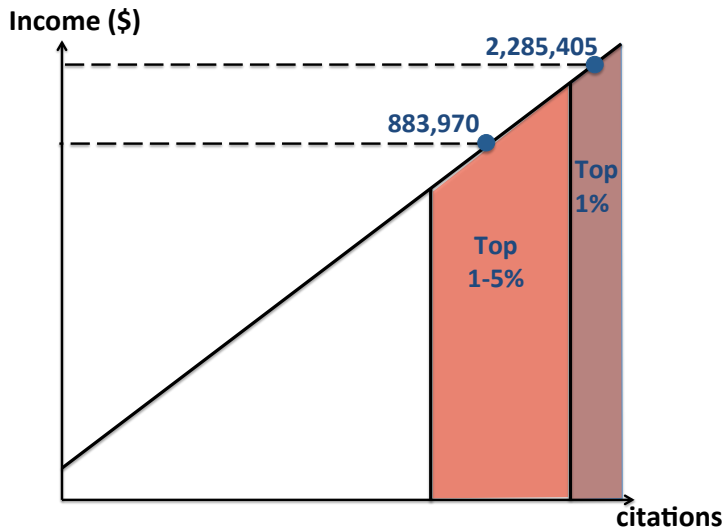
Link between Inventor Quality and Income in IRS data



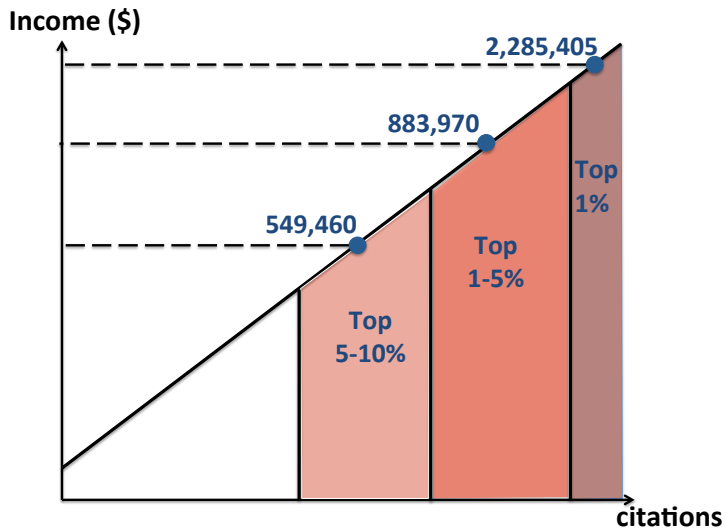
Link between Inventor Quality and Income in IRS data



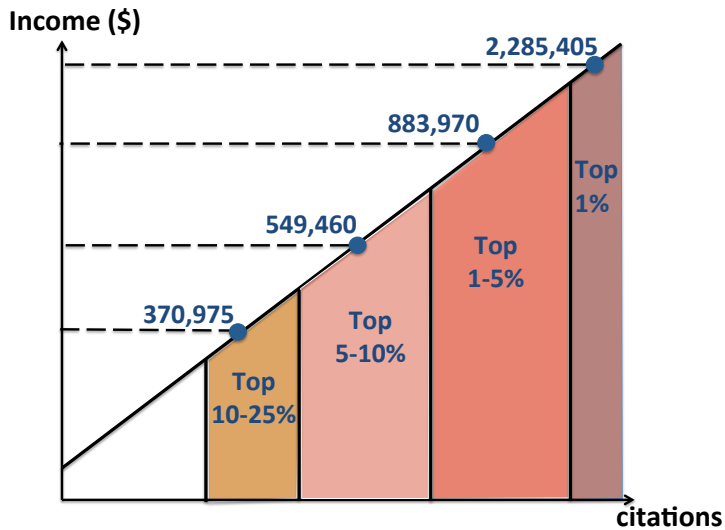
Link between Inventor Quality and Income in IRS data



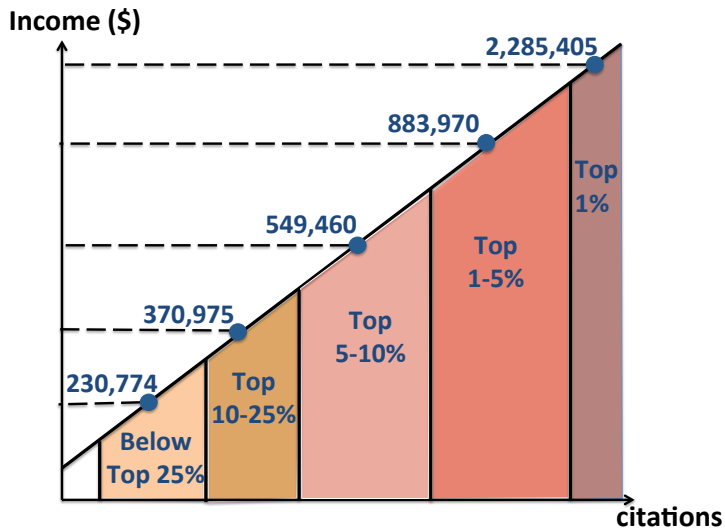
Link between Inventor Quality and Income in IRS data



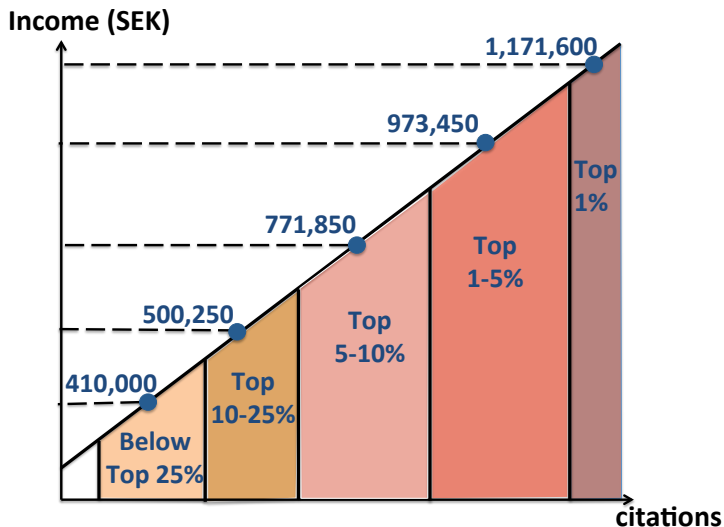
Link between Inventor Quality and Income in IRS data



Link between Inventor Quality and Income in IRS data

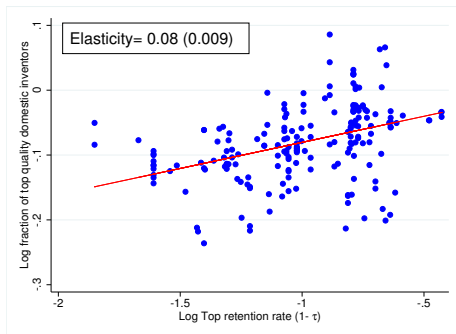


Link between Inventor Quality and Income in Swedish and Finnish Admin data

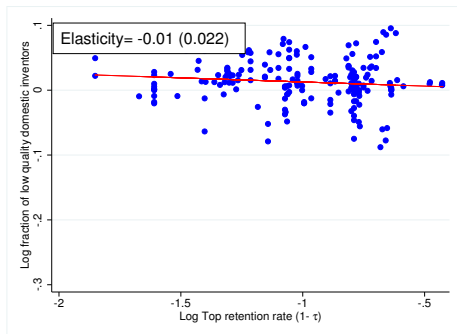


Source: Olof Ejermo and Otto Toivaanen.

Top $(1 - \tau)$ and % of Domestic Inventors in Home Country



(a) Top quality inventors

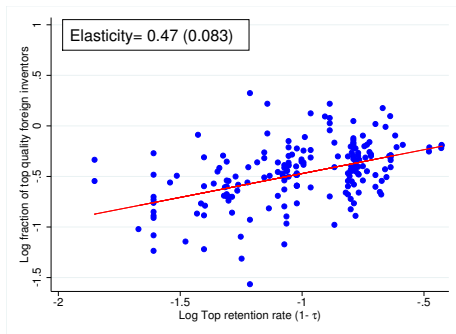


(b) Low quality inventors

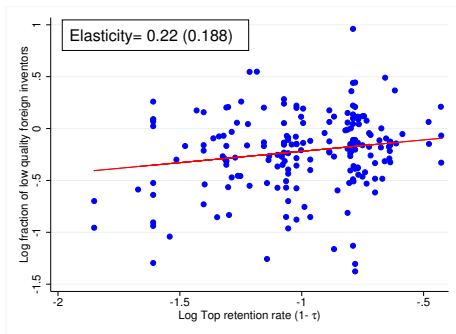
Additional macro level results in the paper:

- Domestic and Foreign inventors.
- For different quality levels, in different datasets.
- With leads and lags.

Top $(1 - \tau)$ and % of Foreign Inventors



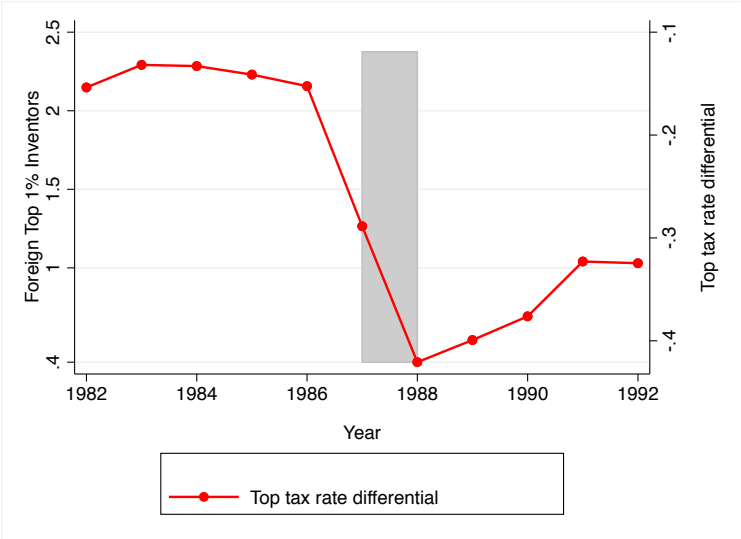
(a) Top quality inventors



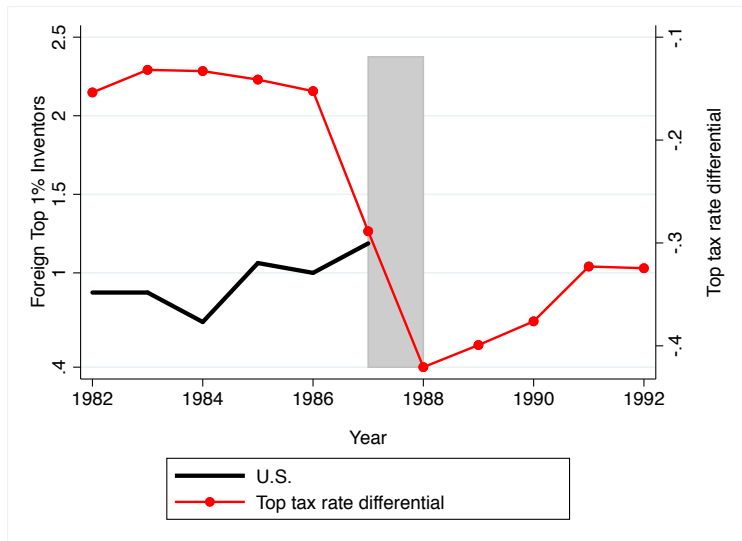
(b) Low quality inventors

Log outcomes at the country-year level. Partial residual plots controlling for country's patent stock, GDP per capita, country fixed effects, year fixed effects. Elasticities reported (standard errors clustered at the country level).

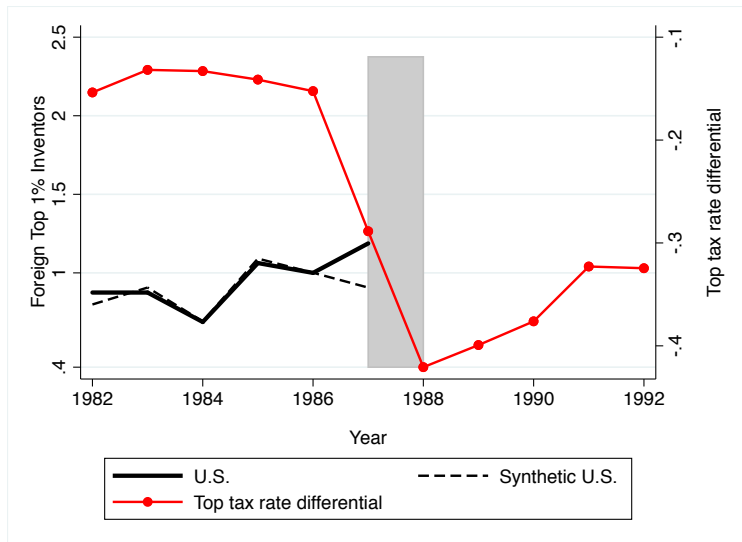
Case Study: U.S. TRA 1986



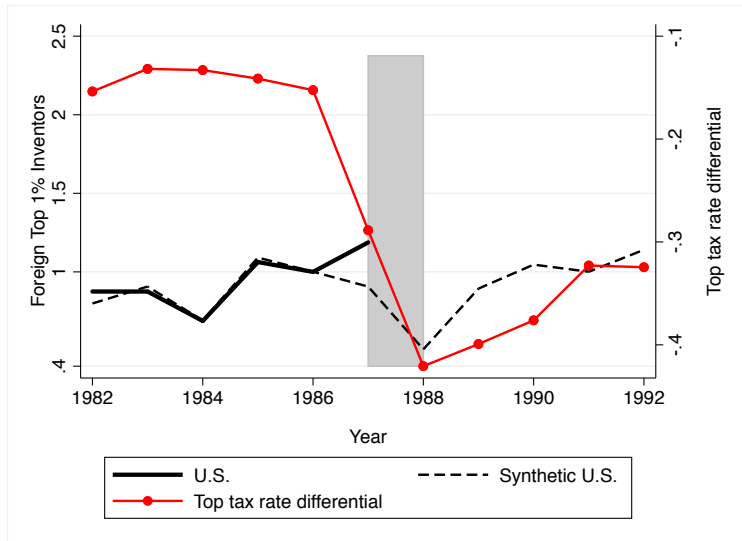
Case Study: U.S. TRA 1986



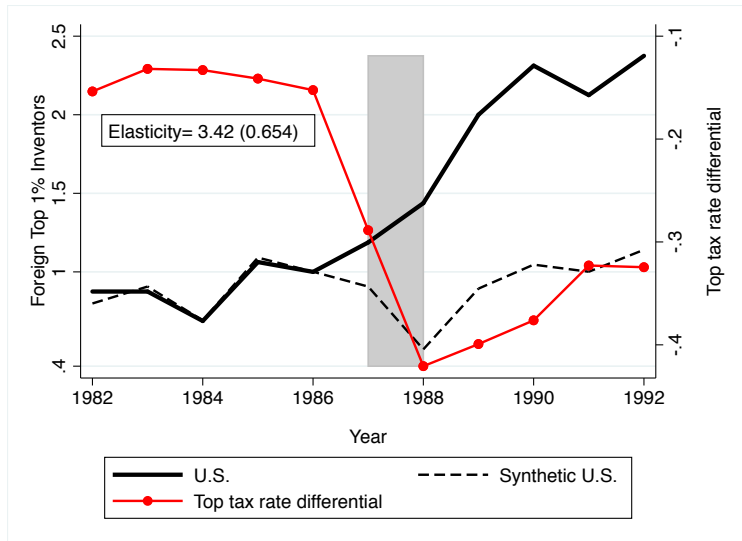
Case Study: U.S. TRA 1986



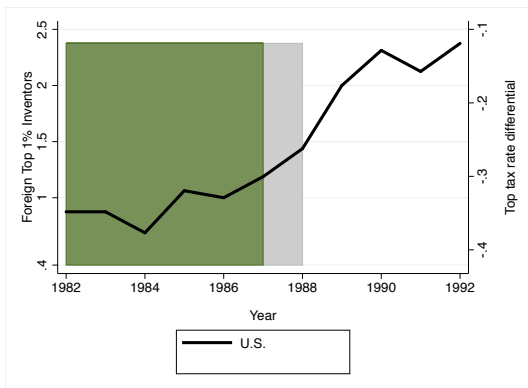
Case Study: U.S. TRA 1986



Case Study: U.S. TRA 1986



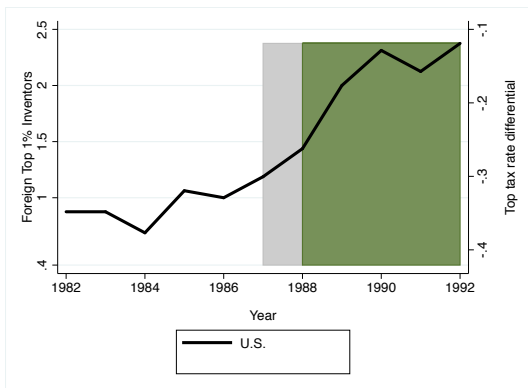
Case Study: U.S. TRA 1986



Structural break in growth of foreign top 1% relative to lower quality inventors.

Inventor quality	Pre T.R.A 1986	Post T.R.A 1986
Top 1%	6.8%	16.4%
Top 10-25%	13%	11.4%

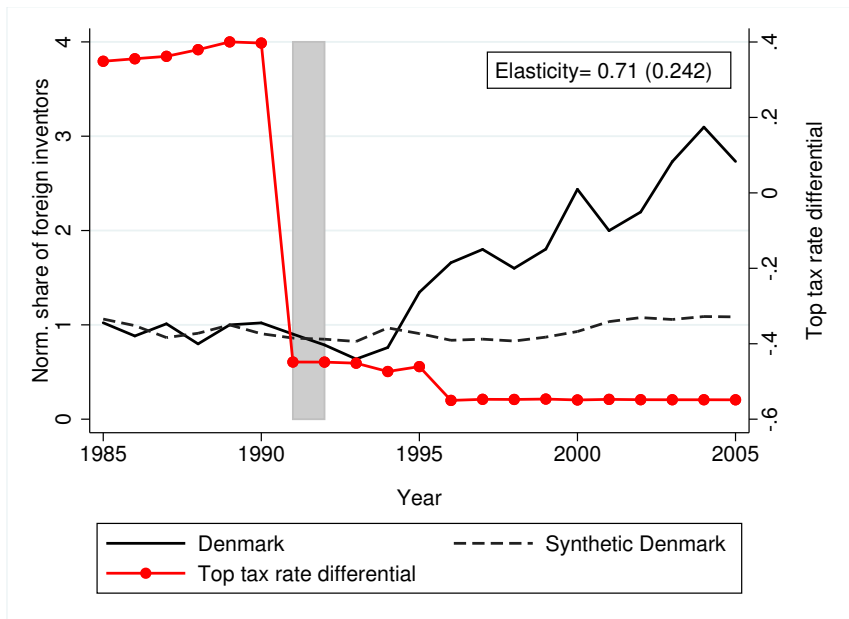
Case Study: U.S. TRA 1986



Structural break in growth of foreign top 1% relative to lower quality inventors.

Inventor quality	Pre T.R.A 1986	Post T.R.A 1986
Top 1%	6.8%	16.4%
Top 10-25%	13%	11.4%

Case Study: Denmark's 1992 Preferential Tax Reform



$$Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}_{ct}^i) + \beta_c \mathbf{x}_{ti} + \eta \mathbf{x}_{cti} + \zeta \mathbf{x}_{ct})$$

\mathbf{x}_{ti} : individual covariates (\times country FE), control for *counterfactual* earnings. Age, tech field, works for multinational, ranking

+ quality \times country FE

+ quality \times country FE \times trend

+ quality \times country FE \times trend \times tech field.

\mathbf{x}_{cti} : individual-country pair covariates: home dummy, patent stock in inventor's tech field, distance, common language.

- \mathbf{x}_{ct} : country covariates.

$$Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}_{ct}^i) + \beta_c \mathbf{x}_{ti} + \eta \mathbf{x}_{cti} + \zeta \mathbf{x}_{ct})$$

\mathbf{x}_{ti} : individual covariates (\times country FE), control for *counterfactual* earnings. Age, tech field, works for multinational, ranking

+ quality \times country FE

+ quality \times country FE \times trend

+ quality \times country FE \times trend \times tech field.

\mathbf{x}_{cti} : individual-country pair covariates: home dummy, patent stock in inventor's tech field, distance, common language.

- \mathbf{x}_{ct} : country covariates.

$$Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}_{ct}^i) + \beta_c \mathbf{x}_{ti} + \eta \mathbf{x}_{cti} + \zeta \mathbf{x}_{ct})$$

\mathbf{x}_{ti} : individual covariates (\times country FE), control for *counterfactual* earnings. Age, tech field, works for multinational, ranking

+ quality \times country FE

+ quality \times country FE \times trend

+ quality \times country FE \times trend \times tech field.

\mathbf{x}_{cti} : individual-country pair covariates: home dummy, patent stock in inventor's tech field, distance, common language.

- \mathbf{x}_{ct} : country covariates.

$$Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}_{ct}^i) + \beta_c \mathbf{x}_{ti} + \eta \mathbf{x}_{cti} + \zeta \mathbf{x}_{ct})$$

\mathbf{x}_{ti} : individual covariates (\times country FE), control for *counterfactual* earnings. Age, tech field, works for multinational, ranking

+ quality \times country FE

+ quality \times country FE \times trend

+ quality \times country FE \times trend \times tech field.

\mathbf{x}_{cti} : individual-country pair covariates: home dummy, patent stock in inventor's tech field, distance, common language.

- \mathbf{x}_{ct} : country covariates.

$$Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}_{ct}^i) + \beta_c \mathbf{x}_{ti} + \eta \mathbf{x}_{cti} + \zeta \mathbf{x}_{ct})$$

\mathbf{x}_{ti} : individual covariates (\times country FE), control for *counterfactual* earnings. Age, tech field, works for multinational, ranking

+ quality \times country FE

+ quality \times country FE \times trend

+ quality \times country FE \times trend \times tech field.

\mathbf{x}_{cti} : individual-country pair covariates: home dummy, patent stock in inventor's tech field, distance, common language.

- \mathbf{x}_{ct} : country covariates.
- **Country-by-year variation:** patent stock, GDP per capita, country FEs, year FEs, country-specific time trends.
 - ▶ Contemporaneous country-specific policies?
 - ▶ Loads general equilibrium effects and sorting on coefficient of top tax (e.g.: inflow of higher ability inventors could displace low ability inventors if rigid demand).

$$Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}_{ct}^i) + \beta_c \mathbf{x}_{ti} + \eta \mathbf{x}_{cti} + \zeta \mathbf{x}_{ct})$$

\mathbf{x}_{ti} : individual covariates (\times country FE), control for *counterfactual* earnings. Age, tech field, works for multinational, ranking

+ quality \times country FE

+ quality \times country FE \times trend

+ quality \times country FE \times trend \times tech field.

\mathbf{x}_{cti} : individual-country pair covariates: home dummy, patent stock in inventor's tech field, distance, common language.

- \mathbf{x}_{ct} : country covariates.

- **Superstars vs. Non-superstars**: include country \times year FE.

- ▶ Logic: Top 1% and slightly lower quality inventors very comparable.

- ▶ Only inventors actually in top tax bracket are directly affected by top tax.

- ▶ Higher quality \rightarrow Higher income \rightarrow higher propensity to be treated by top MTR (MTR \approx ATR).

$$Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}_{ct}^i) + \beta_c \mathbf{x}_{ti} + \eta \mathbf{x}_{cti} + \zeta \mathbf{x}_{ct})$$

\mathbf{x}_{ti} : individual covariates (\times country FE), control for *counterfactual* earnings. Age, tech field, works for multinational, ranking

+ quality \times country FE

+ quality \times country FE \times trend

+ quality \times country FE \times trend \times tech field.

\mathbf{x}_{cti} : individual-country pair covariates: home dummy, patent stock in inventor's tech field, distance, common language.

- \mathbf{x}_{ct} : country covariates.

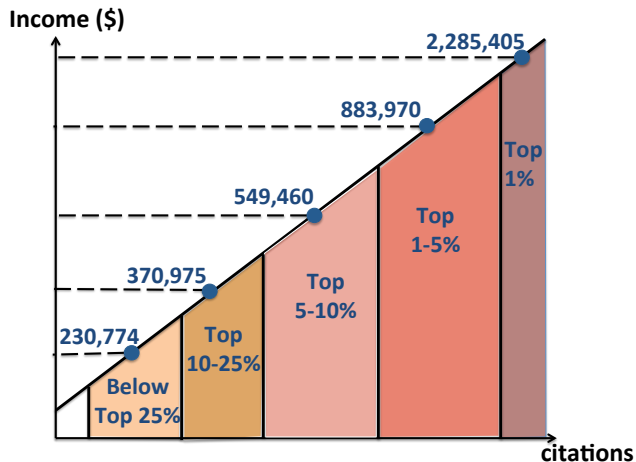
- **Superstars vs. Non-superstars**: include country \times year FE.

- ▶ Logic: Top 1% and slightly lower quality inventors very comparable.

- ▶ Only inventors actually in top tax bracket are directly affected by top tax.

- ▶ Higher quality \rightarrow Higher income \rightarrow higher propensity to be treated by top MTR (MTR \approx ATR).

Choice of the Control Group?



Trade-off in the choice of the control group.

→ Provide set of effects of $(1 - MTR)$ on all quality groups.

→ Provide elasticity of top 1% relative to several control groups

$g \in \{\text{top 5-10\%, top 10-25\%, below top 25\%}\}$.

Country-by-year Variation and General Equilibrium Effects

	(1)	(2)	(3)	(4)
Log Retention Rate \times Top 1	0.890** (0.365)	0.891** (0.377)	0.965** (0.384)	0.951** (0.383)
Log Retention Rate \times Top 1-5	0.447** (0.182)	0.456** (0.197)	0.527*** (0.199)	0.507** (0.203)
Log Retention Rate \times Top 5-10	0.141 (0.142)	0.155 (0.148)	0.227 (0.147)	0.202 (0.148)
Log Retention Rate \times Top 10-25	-0.131 (0.113)	-0.107 (0.114)	-0.0296 (0.108)	-0.0533 (0.106)
Log Retention Rate \times Below Top 25	-0.415*** (0.150)	-0.358** (0.171)	-0.275 (0.176)	-0.285 (0.176)
Quality \times Country FE	NO	YES	YES	YES
Quality \times Country FE \times Year	NO	NO	YES	YES
Quality \times Country FE \times Year \times Field FE	NO	NO	NO	YES
Domestic elasticity s.e	0.02 (0.009)	0.02 (0.009)	0.024 (0.009)	0.023 (0.009)
Foreign elasticity s.e	0.75 (0.305)	0.751 (0.319)	0.807 (0.324)	0.798 (0.322)
Observations	8,645,464	8,617,464	8,617,464	8,617,464

Superstars vs. Non-Superstars

	(1)	(2)	(3)	(4)
Log Retention Rate × Top 1	1.328** (0.644)	1.456** (0.642)	1.399** (0.667)	1.352** (0.669)
Log Retention Rate × Top 1-5	0.885* (0.514)	1.022** (0.514)	0.961* (0.532)	0.907* (0.536)
Log Retention Rate × Top 5-10	0.576 (0.495)	0.719 (0.483)	0.658 (0.501)	0.599 (0.506)
Log Retention Rate × Top 10-25	0.303 (0.486)	0.456 (0.466)	0.398 (0.481)	0.341 (0.484)
Log Retention Rate × Below Top 25	0.022 (0.493)	0.207 (0.471)	0.153 (0.478)	0.110 (0.482)
Quality × Country FE	NO	YES	YES	YES
Quality × Country FE × Year	NO	NO	YES	YES
Quality × Country FE × Year × Field FE	NO	NO	NO	YES
Control: Top 5-10				
Domestic elasticity	0.02	0.02	0.02	0.02
s.e	(0.009)	(0.009)	(0.009)	(0.009)
Foreign elasticity	0.63	0.62	0.62	0.63
s.e	(0.314)	(0.321)	(0.318)	(0.319)
Control: Top 10-25				
Domestic elasticity	0.03	0.02	0.02	0.02
s.e	(0.009)	(0.009)	(0.009)	(0.009)
Foreign elasticity	0.86	0.84	0.84	0.85
s.e	(0.323)	(0.334)	(0.335)	(0.334)
Control: Below Top 25				
Domestic elasticity	0.03	0.03	0.03	0.03
s.e	(0.009)	(0.010)	(0.011)	(0.011)
Foreign elasticity	1.09	1.05	1.04	1.04
s.e	(0.340)	(0.376)	(0.382)	(0.381)
Observations	8,645,464	8,617,464	8,617,464	8,617,464

Superstars vs. Non-Superstars

	(1)	(2)	(3)	(4)
Log Retention Rate × Top 1	1.328** (0.644)	1.456** (0.642)	1.399** (0.667)	1.352** (0.669)
Log Retention Rate × Top 1-5	0.885* (0.514)	1.022** (0.514)	0.961* (0.532)	0.907* (0.536)
Log Retention Rate × Top 5-10	0.576 (0.495)	0.719 (0.483)	0.658 (0.501)	0.599 (0.506)
Log Retention Rate × Top 10-25	0.303 (0.486)	0.456 (0.466)	0.398 (0.481)	0.341 (0.484)
Log Retention Rate × Below Top 25	0.022 (0.493)	0.207 (0.471)	0.153 (0.478)	0.110 (0.482)
Quality × Country FE	NO	YES	YES	YES
Quality × Country FE × Year	NO	NO	YES	YES
Quality × Country FE × Year × Field FE	NO	NO	NO	YES
Control: Top 5-10				
Domestic elasticity	0.02	0.02	0.02	0.02
s.e	(0.009)	(0.009)	(0.009)	(0.009)
Foreign elasticity	0.63	0.62	0.62	0.63
s.e	(0.314)	(0.321)	(0.318)	(0.319)
Control: Top 10-25				
Domestic elasticity	0.03	0.02	0.02	0.02
s.e	(0.009)	(0.009)	(0.009)	(0.009)
Foreign elasticity	0.86	0.84	0.84	0.85
s.e	(0.323)	(0.334)	(0.335)	(0.334)
Control: Below Top 25				
Domestic elasticity	0.03	0.03	0.03	0.03
s.e	(0.009)	(0.010)	(0.011)	(0.011)
Foreign elasticity	1.09	1.05	1.04	1.04
s.e	(0.340)	(0.376)	(0.382)	(0.381)
Observations	8,645,464	8,617,464	8,617,464	8,617,464

Superstars vs. Non-Superstars

	(1)	(2)	(3)	(4)
Log Retention Rate × Top 1	1.328** (0.644)	1.456** (0.642)	1.399** (0.667)	1.352** (0.669)
Log Retention Rate × Top 1-5	0.885* (0.514)	1.022** (0.514)	0.961* (0.532)	0.907* (0.536)
Log Retention Rate × Top 5-10	0.576 (0.495)	0.719 (0.483)	0.658 (0.501)	0.599 (0.506)
Log Retention Rate × Top 10-25	0.303 (0.486)	0.456 (0.466)	0.398 (0.481)	0.341 (0.484)
Log Retention Rate × Below Top 25	0.022 (0.493)	0.207 (0.471)	0.153 (0.478)	0.110 (0.482)
Quality × Country FE	NO	YES	YES	YES
Quality × Country FE × Year	NO	NO	YES	YES
Quality × Country FE × Year × Field FE	NO	NO	NO	YES
Control: Top 5-10				
Domestic elasticity	0.02	0.02	0.02	0.02
s.e	(0.009)	(0.009)	(0.009)	(0.009)
Foreign elasticity	0.63	0.62	0.62	0.63
s.e	(0.314)	(0.321)	(0.318)	(0.319)
Control: Top 10-25				
Domestic elasticity	0.03	0.02	0.02	0.02
s.e	(0.009)	(0.009)	(0.009)	(0.009)
Foreign elasticity	0.86	0.84	0.84	0.85
s.e	(0.323)	(0.334)	(0.335)	(0.334)
Control: Below Top 25				
Domestic elasticity	0.03	0.03	0.03	0.03
s.e	(0.009)	(0.010)	(0.011)	(0.011)
Foreign elasticity	1.09	1.05	1.04	1.04
s.e	(0.340)	(0.376)	(0.382)	(0.381)
Observations	8,645,464	8,617,464	8,617,464	8,617,464

Superstars vs. Non-Superstars

	(1)	(2)	(3)	(4)
Log Retention Rate × Top 1	1.328** (0.644)	1.456** (0.642)	1.399** (0.667)	1.352** (0.669)
Log Retention Rate × Top 1-5	0.885* (0.514)	1.022** (0.514)	0.961* (0.532)	0.907* (0.536)
Log Retention Rate × Top 5-10	0.576 (0.495)	0.719 (0.483)	0.658 (0.501)	0.599 (0.506)
Log Retention Rate × Top 10-25	0.303 (0.486)	0.456 (0.466)	0.398 (0.481)	0.341 (0.484)
Log Retention Rate × Below Top 25	0.022 (0.493)	0.207 (0.471)	0.153 (0.478)	0.110 (0.482)
Quality × Country FE	NO	YES	YES	YES
Quality × Country FE × Year	NO	NO	YES	YES
Quality × Country FE × Year × Field FE	NO	NO	NO	YES
Control: Top 5-10				
Domestic elasticity	0.02	0.02	0.02	0.02
s.e	(0.009)	(0.009)	(0.009)	(0.009)
Foreign elasticity	0.63	0.62	0.62	0.63
s.e	(0.314)	(0.321)	(0.318)	(0.319)
Control: Top 10-25				
Domestic elasticity	0.03	0.02	0.02	0.02
s.e	(0.009)	(0.009)	(0.009)	(0.009)
Foreign elasticity	0.86	0.84	0.84	0.85
s.e	(0.323)	(0.334)	(0.335)	(0.334)
Control: Below Top 25				
Domestic elasticity	0.03	0.03	0.03	0.03
s.e	(0.009)	(0.010)	(0.011)	(0.011)
Foreign elasticity	1.09	1.05	1.04	1.04
s.e	(0.340)	(0.376)	(0.382)	(0.381)
Observations	8,645,464	8,617,464	8,617,464	8,617,464

The Role of Companies

	(1)	(2)
Log Retention Rate × Top 1	1.345** (0.676)	1.366** (0.692)
Log Retention Rate × Top 1-5	0.819 (0.550)	0.649 (0.593)
Log Retention Rate × Top 5-10	0.453 (0.516)	0.313 (0.581)
Log Retention Rate × Top 10-25	0.122 (0.509)	0.0350 (0.550)
Log Retention Rate × Below Top 25	-0.314 (0.524)	-0.430 (0.565)
Log Retention Rate × Not Multinational	-0.219* (0.124)	
Log Retention Rate × Activity abroad		-1.506*** (0.151)
Quality × Country FE	YES	YES
Quality × Country FE × Year	YES	YES
Quality × Country FE × Year × Field FE	YES	YES
Control: Top 5-10		
Domestic elasticity	0.022	0.288
s.e	(0.009)	(0.083)
Foreign elasticity	0.756	1.038
s.e	(0.327)	(0.301)
Control: Top 10-25		
Domestic elasticity	0.030	0.363
s.e	(0.009)	(0.089)
Foreign elasticity	1.038	1.313
s.e	(0.330)	(0.322)
Control: Below Top 25		
Domestic elasticity	0.041	0.492
s.e	(0.010)	(0.095)
Foreign elasticity	1.407	1.771
s.e	(0.342)	(0.341)
Observations	7,060,896	6,169,624

The Role of Companies

	(1)	(2)
Log Retention Rate × Top 1	1.345** (0.676)	1.366** (0.692)
Log Retention Rate × Top 1-5	0.819 (0.550)	0.649 (0.593)
Log Retention Rate × Top 5-10	0.453 (0.516)	0.313 (0.581)
Log Retention Rate × Top 10-25	0.122 (0.509)	0.0350 (0.550)
Log Retention Rate × Below Top 25	-0.314 (0.524)	-0.430 (0.565)
Log Retention Rate × Not Multinational	-0.219* (0.124)	
Log Retention Rate × Activity abroad		-1.506** (0.151)
Quality × Country FE	YES	YES
Quality × Country FE × Year	YES	YES
Quality × Country FE × Year × Field FE	YES	YES
Control: Top 5-10		
Domestic elasticity	0.022	0.288
s.e	(0.009)	(0.083)
Foreign elasticity	0.756	1.038
s.e	(0.327)	(0.301)
Control: Top 10-25		
Domestic elasticity	0.030	0.363
s.e	(0.009)	(0.089)
Foreign elasticity	1.038	1.313
s.e	(0.330)	(0.322)
Control: Below Top 25		
Domestic elasticity	0.041	0.492
s.e	(0.010)	(0.095)
Foreign elasticity	1.407	1.771
s.e	(0.342)	(0.341)
Observations	7,060,896	6,169,624