Effects of Taxes, Part II: Taxable Income Elasticities and Broader Effects of Taxation

Stefanie Stantcheva

Fall 2022
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 $MTR$s
Historically, a 70 percent marginal tax rate is not unusual

The top marginal income tax rates from 1913 to 2018

1981
Reagan took office

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

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 \textbf{pre-tax} incomes respond to changes in one minus the marginal tax rate (\textit{=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
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 $\tau_t$ 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>
<thead>
<tr>
<th></th>
<th>Top 1%</th>
<th>Next 9%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>A. Tax Reform Episodes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1981 vs. 1984 (ERTA 1981)</td>
<td>0.60</td>
<td>0.21</td>
</tr>
<tr>
<td>1986 vs. 1988 (TRA 1986)</td>
<td>1.36</td>
<td>-0.20</td>
</tr>
<tr>
<td>1992 vs. 1993 (OBRA 1993)</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td>1991 vs. 1994 (OBRA 1993)</td>
<td>-0.39</td>
<td></td>
</tr>
<tr>
<td><strong>B. Full Time Series 1960-2006</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No time trends</td>
<td>1.71</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Linear time trend</td>
<td>0.82</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Linear and square time trends</td>
<td>0.74</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Linear, square, and cube time trends</td>
<td>0.58</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>

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

Estimates in Panel B are obtained by time-series regression of log(top 1% income share) on a constant, log (1 - average marginal tax rate), and polynomials time controls from 1960 to 2006 (44 observations). OLS regression. Standard Errors from Newey-West with 8 lags.

Source: Saez et al. (2010)
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

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

where $z^H$, $z^M$ and $\tau^H$, $\tau^M$ 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**

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

**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

<table>
<thead>
<tr>
<th>Taxpayer Groups Classified by 1985 Marginal Rate</th>
<th>Net of Tax Rate (1)</th>
<th>Adjusted Taxable Income (2)</th>
<th>Adjusted Taxable Income Plus Gross Loss (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Medium (22–38)</td>
<td>12.2</td>
<td>6.2</td>
<td>6.4</td>
</tr>
<tr>
<td>2. High (42–45)</td>
<td>25.6</td>
<td>21.0</td>
<td>20.3</td>
</tr>
<tr>
<td>3. Highest (49–50)</td>
<td>42.2</td>
<td>71.6</td>
<td>44.8</td>
</tr>
</tbody>
</table>

Percentage Changes, 1985–88

<table>
<thead>
<tr>
<th>Differences of Differences</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4. High minus medium</td>
<td>13.4</td>
<td>14.8</td>
</tr>
<tr>
<td>5. Highest minus high</td>
<td>16.6</td>
<td>50.6</td>
</tr>
<tr>
<td>6. Highest minus medium</td>
<td>30.0</td>
<td>65.4</td>
</tr>
</tbody>
</table>

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.
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 \( \tau = \frac{1}{1 + a \cdot e} = \frac{1}{1 + 2 \cdot 1} = 33\% \) 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 \( t + 1 \) ⇒ 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
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

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

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

$	au_{it+3}$ and $\varepsilon_{it}$ are correlated [because $\tau_{it+3} = T_{t+3}'(z_{it+3})$]

Instrument: predicted change in MTR assuming income stays constant:

$\log \left[ (1 - \tau_{it+3}^P) / (1 - \tau_{it}) \right]$ where $\tau_{it+3}^P = T_{t+3}'(z_{it})$

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


Table 4  
Basic elasticity results

<table>
<thead>
<tr>
<th>Income controls</th>
<th>None</th>
<th>Log income</th>
<th>Log income 10-piece spline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Broad income</td>
<td>Taxable income</td>
<td>Broad income</td>
</tr>
<tr>
<td>Elasticity</td>
<td>-0.300</td>
<td>-0.462</td>
<td>0.170</td>
</tr>
<tr>
<td></td>
<td>(0.120)</td>
<td>(0.194)</td>
<td>(0.106)</td>
</tr>
<tr>
<td>Dummy for marrieds</td>
<td>-0.008</td>
<td>-0.062</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.018)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Dummy for singles</td>
<td>-0.037</td>
<td>-0.053</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.019)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>Log(income) control</td>
<td>-0.083</td>
<td>-0.167</td>
<td>-0.093</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.021)</td>
<td>(0.015)</td>
</tr>
</tbody>
</table>

Source: Gruber and Saez 2002
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]
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)
IV. Empirical Results

A. Graphical Evidence

This section presents graphical evidence on taxable income responses to the large 1987 reform. Figure 4 shows the evolution of labor income (panels A–B) and capital income (panel C) between 1982–1993 for groups that were affected differently by the 1987 reform, demarcated by a vertical line. The figure is based on a balanced panel of individuals who are observed throughout the period. Panel A shows the effect on labor income using a simple treatment-control assignment based on the reform-induced tax variation shown in Figure 3: the treatment group includes those who experience an increase in the marginal net-of-tax rate on labor income due to the reform.

13 The vertical line demarcates 1986, which is the last pre-reform year (as the reform was passed in parliament during 1986 and changed tax rates starting from 1987). Income levels in 1986 are normalized to 100 for all groups.

Panel A. Labor income

Panel B. Labor income: large versus small tax cuts

**Figure 4. Graphical Evidence on Taxable Income Responses to the Danish 1987 Reform (continued)**

**Results:**

- **Panel A:**
  - Labor income (index 1986 = 100)
  - **Treatment (tax cuts)**: Red line
  - **Control (tax increases)**: Blue dashed line
  - DD elasticity = 0.214 (0.011)

- **Panel B:**
  - Labor income (index 1986 = 100)
  - **Treatment L (large tax cuts)**: Red line
  - **Control (tax increases)**: Blue dashed line
  - **Treatment S (small tax cuts)**: Green line
  - DD\_L elasticity = 0.257 (0.013)
  - DD\_S elasticity = 0.186 (0.012)
The following main findings emerge from the figure. First, the income trends of labor and capital income, comparing treatment and control groups over the three-year interval from 1986 to 1989. The DD estimates in all panels are based on 2SLS regressions of log income on an after-reform time dummy, a treatment-group dummy, and the log marginal net-of-tax rate, the latter variable being instrumented by the interaction between the after-reform and treatment-group dummies.

Most of the effect of the tax reform materializes within three years. The income trends are very parallel in the years prior to the reform and then start to diverge precisely in 1987, the first post-reform year. The tax reform was passed in parliament during 1986 and changed tax rates starting from 1987.

The following panels further explore the effects of the tax reform on labor and capital income.

Panel A. Labor income

Panel B. Labor income: large versus small tax cuts

Panel C. Positive capital income

The DD elasticity of the difference in taxable labor and capital income, comparing treatment and control groups over the three-year interval 1986–1989. The estimates are very similar to those in Panel A, with the difference between labor and capital income being positive and small net-of-tax rate increases.

The figure shows the evolution of labor income, with the treatment group defined as those who experience the largest net-of-tax rate increases, and the control group includes those who experience a reduction in the marginal net-of-tax rate on labor income due to the reform. Panel C also shows effects on labor income, but splitting the treatment sample into those experiencing at least 15 percent increase in the marginal net-of-tax rate on labor income due to the reform and those experiencing smaller net-of-tax rate increases. Panel C shows the effect of the tax reform on positive capital income, with the treatment group defined as those who experience an increase in the marginal net-of-tax rate on capital income due to the reform. The untreated individuals are those who experience a reduction in the marginal net-of-tax rate on capital income due to the reform. The untreated individuals are denoted as “stay middle” group in Figure 3.

The figure also reports difference-in-differences estimates of the elasticities of taxable labor and capital income, comparing treatment and control groups over the three-year interval 1986–1989. The estimates are very similar to those in Panel A, with the difference between labor and capital income being positive and small net-of-tax rate increases.

Notes:
14 The difference-in-differences estimates are based on 2SLS (two-stage least-squares) regressions of log income on an after-reform time dummy, a treatment-group dummy, and the log marginal net-of-tax rate, the latter variable being an aggregation of groups who experience an increase in the marginal net-of-tax rate on labor income due to the reform. The untreated individuals are those who experience a reduction in the marginal net-of-tax rate on labor income due to the reform. The untreated individuals are denoted as “stay middle” group in Figure 3.

The DD elasticity of the difference in taxable labor and capital income, comparing treatment and control groups over the three-year interval 1986–1989. The estimates are very similar to those in Panel A, with the difference between labor and capital income being positive and small net-of-tax rate increases.

The following panels further explore the effects of the tax reform on labor and capital income.

Panel A. Labor income

Panel B. Labor income: large versus small tax cuts

Panel C. Positive capital income

The DD elasticity of the difference in taxable labor and capital income, comparing treatment and control groups over the three-year interval 1986–1989. The estimates are very similar to those in Panel A, with the difference between labor and capital income being positive and small net-of-tax rate increases.
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) ⇒ Income shifting

(2) $z^B$ can be the same tax base in a different time period (such as future income) ⇒ 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.
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 \( \uparrow \) [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)

\( \Rightarrow \) Large short-term response due to re-timing, small long-term response

Some response but smaller around the 2013 tax increase
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)
**EXECUTIVE COMPENSATION**

**TABLE 2**

**Average Compensation by Type for High-Income Executives**  
(in Thousands)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Taxable income</td>
<td>911</td>
<td>1,153</td>
<td>974</td>
<td>965</td>
<td>1,173</td>
</tr>
<tr>
<td>Salary</td>
<td>347</td>
<td>336</td>
<td>336</td>
<td>351</td>
<td>373</td>
</tr>
<tr>
<td>Bonus</td>
<td>198</td>
<td>207</td>
<td>241</td>
<td>284</td>
<td>330</td>
</tr>
<tr>
<td>LTIP payout</td>
<td>57</td>
<td>72</td>
<td>57</td>
<td>64</td>
<td>89</td>
</tr>
<tr>
<td>Options exercised</td>
<td>268</td>
<td>496</td>
<td>293</td>
<td>235</td>
<td>381</td>
</tr>
<tr>
<td>Other income (nontaxed)</td>
<td>36</td>
<td>37</td>
<td>66</td>
<td>54</td>
<td>78</td>
</tr>
</tbody>
</table>

**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 $\tau_{\text{distrib}}$ 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 $\tau_{\text{cg}} = 20\%$]

But distributions can be deferred so that $\tau_{\text{distrib}} \ll \tau_d, \tau_{\text{cg}}$

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 \((1 - \tau_c) \cdot (1 - \tau_{distrib}) > 1 - \tau_i\)

US fed taxes in 2018: \(\tau_c = 21\%\), \(\tau_{cg} = \tau_d = 20\%\), (but \(\tau_{distrib} << 20\%\) 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 \(\tau_i = 39.6\%\) ⇒ individual form better

⇒ wealthy people likely to incorporate their businesses in ’18+

Before TRA’86 (and especially before ERTA’81), top individual rate \(\tau_i\) 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
income also rose dramatically immediately after 1986, mostly because of the disappearance of partnership losses. The sudden jump in S-corporation income, exactly at the time of the tax reform, strongly suggests that this was the consequence of a one-time shift of taxable income from the corporate sector, and the one-time closing of the partnership loss tax shelters (Slemrod 1996, Gordon and Slemrod 2000). The surge in business income reported on individual returns in the 1980s thus cannot be interpreted as a “supply-side” success, as most of those individual income gains came either at the expense of taxable corporate income, or were obtained from the closing of tax shelters by imposing stricter rules on losses from passive businesses.

Figure 2 displays a dramatic shift in the composition of very top incomes away from dividends (which represented more than 60 percent of top incomes in the early 1960s) toward wage income and S-corporation and partnership income. In the early 1960s, when the top 0.01 percent incomes were facing marginal tax rates of about 80 percent on average, tax rates on long-term capital gains were around 25 percent. Thus, dividends were a very disadvantaged form of income for the rich, suggesting that those top income earners had little control over the form of payment, and thus might have been in large part passive investors. The Kennedy tax cuts apparently did not reduce the top individual rate enough (the top rate became 70 percent).
US Top 0.1% Income Share and Composition (excl. K gains)

Dividends
Other capital income
Business income
Salaries

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

Elasticity = .07 (.15)

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

Elasticity = 1.90 (0.43)

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

### Table 2: International Evidence on Top Income Elasticities

<table>
<thead>
<tr>
<th></th>
<th>All 18 countries and fixed periods</th>
<th>Bootstrapping period and country set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)      (2)          (3)</td>
<td>(4)          (5)          (6)</td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No controls</td>
<td>0.324</td>
<td>0.163</td>
<td>0.803</td>
<td>0.364</td>
<td>0.128</td>
<td>0.821</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.039)</td>
<td>(0.053)</td>
<td>(0.043)</td>
<td>(0.085)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Time trend control</td>
<td>0.375</td>
<td>0.182</td>
<td>0.656</td>
<td>0.425</td>
<td>0.191</td>
<td>0.761</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.030)</td>
<td>(0.056)</td>
<td>(0.045)</td>
<td>(0.091)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Country fixed effects</td>
<td>0.314</td>
<td>0.007</td>
<td>0.626</td>
<td>0.267</td>
<td>0.008</td>
<td>0.595</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.039)</td>
<td>(0.044)</td>
<td>(0.035)</td>
<td>(0.070)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>774</td>
<td>292</td>
<td>482</td>
<td>286</td>
<td>132</td>
<td>516</td>
</tr>
</tbody>
</table>
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?

Long-term Correlation between pre-tax top reported incomes and top tax rates

If due solely to tax avoidance, true top income shares were high in the 1950s-1970s but top earners could lower their taxable income (by retaining earnings in businesses and benefit from lower tax rate on capital gains)

But top income share including K gains follows the same U-shape (Piketty, Saez, Stantcheva ’14)

Piketty, Saez, Zucman QJE’18: comprehensive national income estimates are also U-shaped over the century

⇒ Long-run evolution of inequality is not an artifact of tax avoidance or evasion
Emmanuel Saez and Gabriel Zucman

is reported on tax returns. Untaxed capital income includes undistributed corpo-
rate profits, the imputed rents of homeowners, capital income paid to pension
accounts, and dividends and interest retained in trusts, estates, and fiduciaries.
Piketty, Saez, and Zucman (2018) estimate the distribution of 100
percent of
national income by combining national accounts, tax, and survey data. As Figure
3 shows, in both fiscal income and national income statistics, the share of income
earned by the top 1
percent was high before the 1930s and fell from the 1930s to
the 1970s before rising again from the late 1970s on. This U-shaped evolution of
income concentration is a bit less spectacular when one looks at national income
rather than fiscal income, mainly because only the fraction of corporate profits paid
out as dividends are included in fiscal income statistics, while all corporate profits
are included in national income. Accounting for the totality of corporate profits
generally increases the top 1
percent income share, but the effect is stronger in the
post-W
world War II years, a time before the rise of pension plans somewhat broad-
ened equity ownership.

One virtue of distributional national accounts is that they are not affected by
legal changes in business organization. In the United States, a growing number
of businesses have been organized as “pass-through” entities since the late 1980s.
The income of pass-through entities—partnerships, S-corporations, sole proprietor-
ships—is not subject to the corporate income tax; instead, all the income of these
Figure 3
Share of Income Earned by the Top 1 Percent
Source: Saez and Zucman JEP’20

Note: This figure compares the share of fiscal income earned by the top 1 percent tax units (from Piketty
and Saez 2003, updated series including capital gains in income to compute shares but not to define
ranks, to smooth the lumpiness of realized capital gains) to the share of pre-tax national income earned
by the top 1 percent equal-split adults (from Piketty, Saez, and Zucman 2018, updated September 2020,
available on WID.world).
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

Charitable giving of top 1% to mean income

- Mean charitable giving of top 1% divided by mean income [left y-axis]

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

Source: Saez TPE 2017
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).

Source: Saez TPE 2017
Tax Avoidance: Top 1% Income Shares and Top MTR

![Graph showing trends in Top 1% Income Shares and Top MTR over time, with markers for different categories such as Top 1% Share, Top MTR, Top 1% (excl. KG), and MTR K gains.](image_url)
SUPPLY-SIDE OR RENT-SEEKING?
(PIKETTY-SAEZ-STANTCHEVA)

Correlation between \textbf{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 $\Rightarrow$ Consistent with rent-seeking effects

2) Look at cross-country correlation between economic growth and top tax rate cuts $\Rightarrow$ No correlation supports trickle-up

Two micro-tests using CEO pay data (panel in the U.S. + international cross-section).
MICROEVIDENCE ON BARGAINING: EMPIRICAL STRATEGY

- **Effect of general performance on pay (OLS):**

  \[ \text{pay}_{it} = \beta \times \text{p}_{it} + \gamma_i + \chi_t + \alpha X \times X_{it} + \varepsilon_{it} \]

  \( \text{pay}_{it} \): CEO pay in firm \( i \) at time \( t \), \( \text{p}_{it} \): performance measure,
  \( \gamma_i \): firm FE, \( \chi_t \): time FE, \( X_{it} \): CEO controls (age, tenure).

- **Effect of luck performance on pay (IV):**
  
  1. **Stage:** Effect of luck on performance measure

     \[ \text{p}_{it} = b \times \text{p}_{luck, it} + g_i + c_t + \alpha X \times X_{it} + e_{it} \]  

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

  2. **Stage:** Estimate sensitivity of pay to predictable changes in \( \text{p}_{it} \):

     \[ \text{y}_{it} = \beta_{luck} \times \hat{\text{p}}_{it} + \gamma_i + \chi_t + \alpha X \times X_{it} + \varepsilon_{it} \]

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

     If \( \beta_{luck} \geq \beta \): no filtering at all of luck component.
MICROEVIDENCE ON BARGAINING: LUCK AND PERFORMANCE MEASURES

- **Performance measures:**
  1. Net Income
  2. Return on Equity
  3. Shareholder Wealth (log)

- **Measure of pay:** Total Pay and Cash pay (total pay could be driven by options’ value gains with industry performance)

- **Measure of luck:** Mean asset-weighted performance of other firms in industry.

- **Data:** Forbes 800 + Execucomp, COMPUSTAT-CRSP.

- **Years:** 1970-2011 - includes high tax period (pre-1987) and low tax period (post-1987).
Table 3: US CEO Pay Evidence, 1970-2010

<table>
<thead>
<tr>
<th>Firm performance measure</th>
<th>Log(net income)</th>
<th>Log(stock-market value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Log(CEO pay)</td>
<td>Log(CEO pay)</td>
</tr>
<tr>
<td></td>
<td>Log(industry level workers pay)</td>
<td>Log(industry level workers pay)</td>
</tr>
<tr>
<td>Outcome (LHS variable)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Log(CEO pay)</td>
<td>Log(CEO pay)</td>
</tr>
<tr>
<td></td>
<td>Industry luck IV</td>
<td>Industry luck IV</td>
</tr>
<tr>
<td></td>
<td>OLS regression</td>
<td>OLS regression</td>
</tr>
<tr>
<td>OLS vs. IV</td>
<td>CLS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>(5)</td>
<td>(6)</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Firm performance (RHS variable)</th>
<th>0.23***</th>
<th>0.34***</th>
<th>0.00</th>
<th>0.28***</th>
<th>0.22*</th>
<th>0.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of observations</td>
<td>8,632</td>
<td>8,503</td>
<td>890</td>
<td>9,005</td>
<td>8,865</td>
<td>898</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Firm performance (RHS variable)</th>
<th>0.27***</th>
<th>0.70***</th>
<th>-0.02</th>
<th>0.37***</th>
<th>0.95***</th>
<th>-0.02</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of observations</td>
<td>14,914</td>
<td>14,697</td>
<td>1,422</td>
<td>17,775</td>
<td>17,593</td>
<td>1,443</td>
</tr>
</tbody>
</table>

C. Test for difference between low- and high-top tax rate periods

<table>
<thead>
<tr>
<th>Difference Panel B - Panel A</th>
<th>0.04***</th>
<th>0.36*</th>
<th>-0.019</th>
<th>0.09***</th>
<th>0.72**</th>
<th>-0.023</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value of difference</td>
<td>0.01</td>
<td>0.06</td>
<td>0.440</td>
<td>0.00</td>
<td>0.05</td>
<td>0.46</td>
</tr>
</tbody>
</table>
MICROEVIDENCE ON BARGAINING: RESULTS

- Incomplete filtering of luck component in CEO pay: $\beta_{\text{luck}} \neq 0$.
- Pay for luck is large and almost no filtering: $\beta_{\text{luck}} \geq \beta$.
- Top tax rates decrease pay for luck more than general pay for performance, consistent with bargaining model.
MICROEVIDENCE ON BARGAINING: DISCUSSION

Could pay for luck be consistent with optimal contracting view?

- CEO human capital value increasing in industry performance? But pay responds more to increases than decreases in industry performance; CEO turnover countercyclical (BM).
- Is 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!
INTERNATIONAL CEO PAY: DATA

- Fernandez et al. (2012) data:
  i) Compensation (BoardEx + Execucomp)
  ii) Stock Ownership (LionShares)
  iii) Firm Performance (Worldscope and Datastream)
  iv) 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

Elasticity = 1.97 (.27)
B. Average CEO compensation with controls

Elasticity = 1.90 (.29)
INTERNATIONAL CEO PAY: GOVERNANCE

- 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>
<thead>
<tr>
<th>Outcome (LHS variable)</th>
<th>Log(CEO pay) (1)</th>
<th>Log(CEO pay) (2)</th>
<th>Log(CEO pay) (3)</th>
<th>Log(CEO pay) (4)</th>
<th>Log(CEO salary) (5)</th>
<th>Log(CEO bonus and equity pay) (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(1-Top MTR)</td>
<td>1.97***</td>
<td>1.90***</td>
<td>1.92***</td>
<td>1.90***</td>
<td>0.35*</td>
<td>4.68***</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.286)</td>
<td>(0.336)</td>
<td>(0.328)</td>
<td>(0.189)</td>
<td>(0.782)</td>
</tr>
<tr>
<td>Governance index</td>
<td>-0.10***</td>
<td>-0.19***</td>
<td>-0.02</td>
<td>-0.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.038)</td>
<td>(0.072)</td>
<td>(0.201)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(1-Top MTR)*Governance index</td>
<td>-0.13**</td>
<td>0.06</td>
<td>-0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.089)</td>
<td>(0.281)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm and CEO controls</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Number of observations</td>
<td>2,959</td>
<td>2,844</td>
<td>2,711</td>
<td>2,711</td>
<td>2,691</td>
<td>2,711</td>
</tr>
</tbody>
</table>
Effects of Changes in Taxation on Economic Activity

Tax policies have a significant impact on economic activity. By altering the marginal tax rates, the government can influence the decisions of both individual taxpayers and businesses. Some of the key effects of tax changes on economic activity include:

1. **Haircuts on Capital Gains:**
   - In some countries, capital gains are taxed at a lower rate than ordinary income. This differential treatment can encourage investment and encourage individuals to hold onto assets rather than selling them to take advantage of the tax break.

2. **Marginal Tax Rates:**
   - Higher marginal tax rates can discourage additional income from being earned, as the extra income will be taxed at a higher rate. This can reduce the incentive to work overtime or pursue additional income streams.

3. **Tax Incentives for Business Activities:**
   - Tax policies can be designed to encourage specific economic activities. For example, certain types of business investments might be eligible for tax credits or deductions, which can make them more attractive for entrepreneurs.

4. **Tax Holidays and Incubators:**
   - Some countries offer tax holidays for new businesses or incubators, which can provide a temporary reduction in the tax burden to help businesses get started.

5. **Tax Deductions for R&Dspending:**
   - Incentives for R&D can be provided through deductions or credits for R&D expenses, which can encourage businesses to invest in research and development activities.

These are just a few examples of the ways in which tax policies can influence economic activity. The specific effects will depend on the details of the tax laws and how they are implemented in each country.
Exploit the 1991 Danish tax scheme: high earnings immigrants (≥ 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: .8 to .9*threshold
Control #2: .9 to .99*threshold
Treatment: earnings > threshold

DD elasticities:
Long-term: 1.62 (.16)
Short-term: 1.28 (.15)

Control 1 = annualized income between .8 and .9 of threshold
Control 2 = annualized income between .9 and .995 of threshold.

DD specifications
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

- 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?
STUDY THE EFFECTS OF TAXES ON MIGRATION USING PATENT DATA

- Use a unique international panel data to overcome challenges:
  - 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

Shunpei Yamazaki (3,780 patents)
The most prolific inventor until 2008
Born: Japan
Works: Japan

Salman Akram (713 patents)
Micron Technology
Born: Nigeria
Works: U.S.

Edwin Herbert Land (535 patents)
Founder of Polaroid
Born: U.S.
Worked: U.S.
Patent quality increases inventor income, directly and \textit{indirectly}.

\textbf{Quality measures (dynamic and lagged)}

- Citations-weighted patents (benchmark)
- Patent count
- Average citations per patent
- Max citations per patent
- Patent breadth (claims-weighted patents)
- Impact breadth (\# tech classes citing patent).

\textbf{Inventor Ranking}

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

\textarrow{Dynamic, Persistent, Life-time 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).

→  Dynamic, Persistent, Life-time ranking

**Inventor Ranking**

- Assign inventors to group based on home country.
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 (number of tech classes citing patent).

**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 (number of tech classes citing patent).

→ 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.
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).

→ 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).

→ 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.

Top 1-5%

Quality in region at time t
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).

→ 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.

![Quality in region at time t]
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).

→ 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.

![Quality in region at time t](image-url)
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).

→ 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

\[ \text{income} = 200,000 + 1,400 \times \text{citations} \]

Source: Bell et al. (2015).
Link between Inventor Quality and Income in IRS data

Source: Bell et al. (2015).
Link between Inventor Quality and Income in IRS data

Source: Bell et al. (2015).
Link between Inventor Quality and Income in IRS data

Source: Bell et al. (2015).
Link between Inventor Quality and Income in IRS data

Source: Bell et al. (2015).
Link between Inventor Quality and Income in IRS data

Source: Bell et al. (2015).
Link between Inventor Quality and Income in Swedish and Finnish Admin data

Source: Olof Ejermo and Otto Toivaannen.
Top $\left( 1 - \tau \right)$ and % of Domestic Inventors in Home Country

**Elasticity** = 0.08 (0.009)

- **Log fraction of top quality domestic inventors**
- **Log Top retention rate** $\left( 1 - \tau \right)$

**Elasticity** = -0.01 (0.022)

- **Log fraction of low quality domestic inventors**
- **Log Top retention rate** $\left( 1 - \tau \right)$

---

(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.
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

The chart illustrates the top tax rate differential and the number of foreign inventors. The top tax rate differential shows a sharp decrease starting in 1988, possibly due to the Tax Reform Act of 1986. The number of foreign inventors also experienced a significant drop around this time.
### Case Study: U.S. TRA 1986

<table>
<thead>
<tr>
<th>Year</th>
<th>U.S. Top tax rate differential</th>
<th>Foreign Top 1% Inventors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1982</td>
<td>-0.4</td>
<td>26</td>
</tr>
<tr>
<td>1984</td>
<td>-0.3</td>
<td>44</td>
</tr>
<tr>
<td>1986</td>
<td>-0.2</td>
<td>2</td>
</tr>
<tr>
<td>1988</td>
<td>-0.1</td>
<td>1</td>
</tr>
<tr>
<td>1989</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td>0.6</td>
<td></td>
</tr>
</tbody>
</table>

The graph shows the trend of U.S. top tax rate differential and foreign top 1% inventors over years 1982 to 1992. The top tax rate differential decreased significantly in 1988, aligning with a decrease in foreign top 1% inventors.
Case Study: U.S. TRA 1986

The graph illustrates the changes in Foreign Top 1% Inventors and the Top tax rate differential from 1982 to 1992. The U.S. and Synthetic U.S. data are represented by different lines, with the Top tax rate differential marked by red dots.

Key observations:
- A significant drop in Foreign Top 1% Inventors occurred in 1986, coinciding with changes in tax rates.
- The top tax rate differential shows a notable decrease around 1988, which may be associated with tax policy changes.

The data suggests a correlation between tax rate changes and foreign investment trends within the U.S. economy during this period.
Case Study: U.S. TRA 1986

The graph shows the changes in the foreign top 1% inventors and the top tax rate differential in the U.S. and Synthetic U.S. over the years 1982 to 1992. The top tax rate differential is indicated by the red line, with a significant drop in 1988. The foreign top 1% inventors are shown by the black line, which fluctuates over the years. The Synthetic U.S. data is represented by the dashed black line.
Case Study: U.S. TRA 1986

Elasticity = 3.42 (0.654)
Case Study: U.S. TRA 1986

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 1%</td>
<td>6.8%</td>
<td>16.4%</td>
</tr>
<tr>
<td>Top 10-25%</td>
<td>13%</td>
<td>11.4%</td>
</tr>
</tbody>
</table>
Case Study: U.S. TRA 1986

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 1%</td>
<td>6.8%</td>
<td>16.4%</td>
</tr>
<tr>
<td>Top 10-25%</td>
<td>13%</td>
<td>11.4%</td>
</tr>
</tbody>
</table>
Case Study: Denmark’s 1992 Preferential Tax Reform

Elasticity = 0.71 (0.242)

<table>
<thead>
<tr>
<th>Year</th>
<th>Denmark</th>
<th>Synthetic Denmark</th>
<th>Top tax rate differential</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Elasticity: 0.71 (0.242)
- Top tax rate differential
- Normal share of foreign inventors
- Year

---

The graph illustrates the changes in the normal share of foreign inventors and the top tax rate differential before and after the 1992 preferential tax reform in Denmark. The elasticity value of 0.71 suggests a significant response to the tax reform, indicating a positive relationship between the top tax rate differential and the normal share of foreign inventors.
\[ Pr(y_{it} = c) = f(\alpha_{rit} \log (1 - \text{top MTR}_{ct}^i) + \beta_c x_{ti} + \eta x_{cti} + \zeta x_{ct}) \]

\( x_{ti} \): 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.

\( x_{cti} \): individual-country pair covariates: home dummy, patent stock in inventor’s tech field, distance, common language.

- \( x_{ct} \): country covariates.
\[ Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}^i_{ct}) + \beta_{c}x_{ti} + \eta_{cti} + \zeta_{ct}) \]

\(x_{ti}\): 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.

\(x_{cti}\): individual-country pair covariates: home dummy, patent stock in inventor’s tech field, distance, common language.

- \(x_{ct}\): country covariates.
\[ Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}^i_{ct}) + \beta_c x_{ti} + \eta x_{cti} + \zeta x_{ct}) \]

\(x_{ti}\): 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.

\(x_{cti}\): individual-country pair covariates: home dummy, patent stock in inventor’s tech field, distance, common language.

\(x_{ct}\): country covariates.
$Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}_i^{ct}) + \beta_c x_{ti} + \eta x_{cti} + \zeta x_{ct})$

$x_{ti}$: 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.

$x_{cti}$: individual-country pair covariates: home dummy, patent stock in inventor’s tech field, distance, common language.

- $x_{ct}$: country covariates.
Pr(y_{it} = c) = f(\alpha_{rit} \log (1 - \text{top MTR}^i_{ct}) + \beta_c x_{ti} + \eta x_{cti} + \zeta x_{ct})

\(x_{ti}\): individual covariates (× country FE), control for counterfactual earnings. Age, tech field, works for multinational, ranking + quality × country FE + quality × country FE × trend + quality × country FE × trend × tech field.

\(x_{cti}\): individual-country pair covariates: home dummy, patent stock in inventor’s tech field, distance, common language.

\(x_{ct}\): 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).
\[ Pr(y_{it} = c) = f( \alpha_{rit} \log (1 - \text{top MTR}_{ct}^i) + \beta_c x_{ti} + \eta x_{cti} + \zeta x_{ct} ) \]

\( x_{ti} \): individual covariates (\( \times \) country FE), control for \textit{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.

\( x_{cti} \): individual-country pair covariates: home dummy, patent stock in inventor’s tech field, distance, common language.

- \( x_{ct} \): country covariates.

- \textbf{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).
\[ Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}_{ct}^i) + \beta_c x_{ti} + \eta x_{cti} + \zeta x_{ct}) \]

\(x_{ti}\): 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.

\(x_{cti}\): individual-country pair covariates: home dummy, patent stock in inventor’s tech field, distance, common language.

- \(x_{ct}\): 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 \(g \in \{\text{top 5-10\%}, \text{top 10-25\%}, \text{below top 25\%}\}\).
## Country-by-year Variation and General Equilibrium Effects

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

Note: **p < 0.01, *p < 0.05, *p < 0.10.
## Superstars vs. Non-Superstars

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Retention Rate × Top 1</td>
<td>1.328**</td>
<td>1.456**</td>
<td>1.399**</td>
<td>1.352**</td>
</tr>
<tr>
<td></td>
<td>(0.644)</td>
<td>(0.642)</td>
<td>(0.667)</td>
<td>(0.669)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 1-5</td>
<td>0.885*</td>
<td>1.022**</td>
<td>0.961*</td>
<td>0.907*</td>
</tr>
<tr>
<td></td>
<td>(0.514)</td>
<td>(0.514)</td>
<td>(0.532)</td>
<td>(0.536)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 5-10</td>
<td>0.576</td>
<td>0.719</td>
<td>0.658</td>
<td>0.599</td>
</tr>
<tr>
<td></td>
<td>(0.495)</td>
<td>(0.483)</td>
<td>(0.501)</td>
<td>(0.506)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 10-25</td>
<td>0.303</td>
<td>0.456</td>
<td>0.398</td>
<td>0.341</td>
</tr>
<tr>
<td></td>
<td>(0.486)</td>
<td>(0.466)</td>
<td>(0.481)</td>
<td>(0.484)</td>
</tr>
<tr>
<td>Log Retention Rate × Below Top 25</td>
<td>0.022</td>
<td>0.207</td>
<td>0.153</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>(0.493)</td>
<td>(0.471)</td>
<td>(0.478)</td>
<td>(0.482)</td>
</tr>
</tbody>
</table>

| 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

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality × Country FE</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Quality × Country FE × Year</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Quality × Country FE × Year × Field FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

### 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

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

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


