Using the Alternative Minimum Tax to Identify the Elasticity of Taxable Income For Higher-Income Taxpayers

Paper Presentation - NTA Panel on "Tax Avoidance"

Ali Abbas
Cornell University

November 23, 2019
Introduction

- Taxes are imposed and enforced for a number of reasons
- Significant behavioral changes amongst taxpayers a possibility
- Distortionary, reducing economic welfare
- Reduce expected tax receipt needed to fund planned expenditure; undercut equality restoration
Behavioral Changes Due to Tax Changes

- Under certain conditions, the elasticity of taxable income (with respect to the net-of-tax rate) is sufficient to capture behavioral changes:
  - Labor supply changes
  - Income shifting and more aggressive tax planning (avoidance)
  - Under-declaration (evasion)

- Impact of behavioral changes via changes in taxable income amplified for higher-income individuals
Federal Individual Income Tax Shares by Income Quintile

- Highest Quintile
- Fourth Quintile
- Middle Quintile
- Second Quintile
- Lowest Quintile

<table>
<thead>
<tr>
<th>Year</th>
<th>Highest Quintile</th>
<th>Fourth Quintile</th>
<th>Middle Quintile</th>
<th>Second Quintile</th>
<th>Lowest Quintile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1979</td>
<td>60</td>
<td>20</td>
<td>0</td>
<td>-20</td>
<td>-10</td>
</tr>
<tr>
<td>1981</td>
<td>62</td>
<td>22</td>
<td>1</td>
<td>-16</td>
<td>-12</td>
</tr>
<tr>
<td>1983</td>
<td>64</td>
<td>24</td>
<td>2</td>
<td>-14</td>
<td>-10</td>
</tr>
<tr>
<td>1985</td>
<td>66</td>
<td>26</td>
<td>3</td>
<td>-12</td>
<td>-8</td>
</tr>
<tr>
<td>1987</td>
<td>68</td>
<td>28</td>
<td>4</td>
<td>-10</td>
<td>-6</td>
</tr>
<tr>
<td>1989</td>
<td>70</td>
<td>30</td>
<td>5</td>
<td>-8</td>
<td>-4</td>
</tr>
<tr>
<td>1991</td>
<td>72</td>
<td>32</td>
<td>6</td>
<td>-6</td>
<td>-2</td>
</tr>
<tr>
<td>1993</td>
<td>74</td>
<td>34</td>
<td>7</td>
<td>-4</td>
<td>0</td>
</tr>
<tr>
<td>1995</td>
<td>76</td>
<td>36</td>
<td>8</td>
<td>-2</td>
<td>2</td>
</tr>
<tr>
<td>1997</td>
<td>78</td>
<td>38</td>
<td>9</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>1999</td>
<td>80</td>
<td>40</td>
<td>10</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>2001</td>
<td>82</td>
<td>42</td>
<td>11</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>2003</td>
<td>84</td>
<td>44</td>
<td>12</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>2005</td>
<td>86</td>
<td>46</td>
<td>13</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>2007</td>
<td>88</td>
<td>48</td>
<td>14</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>2009</td>
<td>90</td>
<td>50</td>
<td>15</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>2011</td>
<td>92</td>
<td>52</td>
<td>16</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>2013</td>
<td>94</td>
<td>54</td>
<td>17</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>2015</td>
<td>96</td>
<td>56</td>
<td>18</td>
<td>18</td>
<td>22</td>
</tr>
</tbody>
</table>
Significance of Income Tax Receipts

**Federal Individual Income Tax as a Share of Total Federal Tax Receipt**

- **Individual Income Taxes**
- **Corporate Income Taxes**
- **Social Insurance and Retirement Receipts**
- **Excise Taxes**
- **Other**

Data from 1979 to 2015.
Literature on ETI wrt MTR

- Historically, many designs have been used to estimate ETI wrt NTR:
  - Difference-in-differences
  - Time series analysis

- Studies esp prior to 2000 revealed high overall elasticities of 1 to 3:
Many recent studies have relied on "bunching estimators", popularized by Saez (2010):


Average elasticity estimates in this sub-literature of around 0-0.4

- Chetty et al. (2011) find implied estimates of 0.01 at the top kink using Danish data

- In the US, Saez (2010) found estimates of 0.1-0.3 for lower income levels, but 0.006 for higher-income individuals facing the top MTR
Saez (2002, 2010) considers higher MTRs:
- Estimated ETI of 0.03 for 31% - 36%
- Estimated ETI of 0.006 for top tax kink 36% - 39.6%

Fig. 8. Density distributions around top kink points, 1993–1997
Magnitude of Estimates Low at Top Tax Kink

Why is ETI estimated via bunching methods so low at the top tax kink in the United States?

Potential confounder? I incorporate the interaction of the regular tax schedule and the Alternative Minimum Tax (AMT) schedule for higher-income individuals.

Use publicly available samples of data provided by the IRS Statistics of Income (SOI) Division from 1993-2011
- Repeated annual cross-sections of individual income tax returns
- Oversamples higher-income individuals
- Restrict data to individuals who turn in their AMT form (Form 6251) – total of 634,703 observations.
Preview of Main Results
The estimated ETI at the RT-AMT intersection kink is approximately 0.08
   An order of magnitude higher than 0.006 (Saez, 2010)
   Same order of magnitude as Chetty et al. (2011) who found 0.01 - but 8 times higher
   Bounded between 0.04-0.09, estimated using non-parametric bounds developed by Bertanha, McCallum and Seegert (2018)

Cleanest estimate is for taxpayers not reporting long-term cap gains (27% of sample).
   Estimated elasticity of 0.15.

Estimated ETI for self-employed is 0.07 as compared to 0.11 for wage earners-only
   Self-employed defined as those with nonzero Schedule C, Schedule E (S corp/partnership) or farming income; and zero OR nonzero wage income
Alternative Minimum Tax (AMT)

- AMT ensures that higher-income individuals do not take too many deductions and pay "fair share" of taxes.
- AMT defines taxable income differently, since many regular income tax deductions are fully or partially disallowed.
  - But it does provide a significant, fixed deduction: e.g. $45,000 for MFJ in year 2000 - phases out at higher levels of AMTI
- With fixed deduction phase-out, effective AMT rates are 26%, 32.5%, 35% and 28%
### Regular Income Tax and AMT Schedules (e.g. year 2000)

<table>
<thead>
<tr>
<th>Regular Taxable Income (MFJ)</th>
<th>Tax Rates</th>
<th>AMT Taxable Income (MFJ)</th>
<th>Tax Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 - $43,850</td>
<td>15%</td>
<td>$0 - $105,000</td>
<td>26%</td>
</tr>
<tr>
<td>$43,850 - $105,950</td>
<td>28%</td>
<td>$105,000 - $161,000</td>
<td>32%</td>
</tr>
<tr>
<td>$105,950 - $161,450</td>
<td>31%</td>
<td>$161,000 - $285,000</td>
<td>35%</td>
</tr>
<tr>
<td>$161,450 - $288,350</td>
<td>36%</td>
<td>$285,000 and above</td>
<td>28%</td>
</tr>
<tr>
<td>$288,350 and above</td>
<td>39.6%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Estimate that 42% of taxpayers with regular TI > $200,000 faced effective schedule in 2000
Regular Income Tax and AMT Schedules: Illustration

Regular Tax Schedule

Taxpayer with USD 300K in gross (pre-tax) income
Total exemptions and deductions (regular tax) = USD 80K

AMT Schedule

Taxpayer with USD 300K in gross (pre-tax) income
AMT exemption = USD 45K
Interaction of the Alternative Minimum Tax (AMT) and Regular Income Tax

Combined Schedules

Taxpayer with USD 300K in gross (pre-tax) income
Total exemptions and deductions (regular tax) = USD 80K
AMT exemption = USD 45K

Effective Schedule

Taxpayer with USD 300K in gross (pre-tax) income
Total exemptions and deductions (regular tax) = USD 80K
AMT exemption = USD 45K
Interaction of the Alternative Minimum Tax (AMT) and Regular Income Tax

- Effective "top kink" does not correspond to the regular tax top kink
  - Dispersion of bunching at regular tax top kink
  - Effective kink has a bigger change in gradient: 28% to 39.6% in 2000, compared to 36% to 39.6% on RT
  - Taxpayer-specific location of effective kink provides valuable additional variation

- Additional variation can be exploited to estimate ETI using other non-traditional bunching methods

- Variation in location of kinks increases robustness to endogeneity concerns
Empirical Strategy

- Find intersection kink for each taxpayer:
  - Regular tax and AMT schedules piece-wise linear.
  - Finding diff in deductions allows for solving system of equations for top pieces.
Empirical Strategy contd.

- Using taxpayer taxable income, I find the distance to the intersection kink.
- Recenter all individual taxpayer kinks.
"Excess Bunching" and Estimated Elasticity

For first cut, I use the parametric, local polynomial approach (Chetty et al., 2011) with uniform dist. assumption to estimate counterfactual density:

\[ C_j = \sum_{i=0}^{p} \beta_i z^i_j + \sum_{i=c-l}^{c+l} \phi_i D_j + \epsilon_j \]

The counterfactual frequency of observations \( \hat{C}_j^{cf} \) is predicted

"Excess bunching" is then:

\[ \hat{b} = \frac{\sum_{j=c-l}^{c+l} C_j - \hat{C}_j}{\sum_{j=c-l}^{c+l} C_j/(2l+1)} \]

Use the traditional Saez (2010) estimator to estimate the ETI wrt the net-of-tax rate:

\[ \hat{e} = \frac{\hat{b}}{K \cdot \frac{\Delta \tau}{W \cdot 1 - \tau_1}} \]
Graphical Evidence: Observed and Counterfactual Densities

1993-2011

Excess mass $b = 0.75(0.27)$
Est. elasticity $e = 0.08(0.03)$

Observation Density

Counterfactual Density

Bunching Region: -30,000 to +10,000
Degree of polynomial = 8
Binwidth = $10,000$
I re-estimate the ETI with assumptions weaker than those used by the traditional Saez estimator, which assumes known heterogeneity distribution across the kink (Blomquist et al., 2018).

For now, exploit Bertanha, McCallum, and Seegert (2016) as first cut:

- Unobserved distribution must be bounded above and below by some amount $M$

\[
\bar{\varepsilon} = \frac{2 \left[ f(\eta^+)^2/2 + f(\eta^-)^2/2 + M \cdot B \right]^{1/2} - (f(\eta^+) + f(\eta^-))}{M(\ln(1 - \tau_0) - \ln(1 - \tau_1))}
\]

\[
\underline{\varepsilon} = \frac{-2 \left[ f(\eta^+)^2/2 + f(\eta^-)^2/2 - M \cdot B \right]^{1/2} + (f(\eta^+) + f(\eta^-))}{M(\ln(1 - \tau_0) - \ln(1 - \tau_1))}.
\]

Estimated non-parametric bounds on elasticity estimates: bounded below at 0.04 and above at 0.09.
## Main Results

<table>
<thead>
<tr>
<th>Years</th>
<th>MTR Change</th>
<th>Bin Width</th>
<th>Bunching region</th>
<th>All Filers</th>
<th>Self-employment Income</th>
<th>Wage earners only</th>
<th>Positive Long-Term Gains</th>
<th>Non-positive Long-Term Gains</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993-2011</td>
<td>28% - 37.5%</td>
<td>5,000</td>
<td>-30,000 – 10,000</td>
<td>0.07***</td>
<td>0.07***</td>
<td>0.08</td>
<td>0.04</td>
<td>0.12***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1,000</td>
<td>-30,000 – 10,000</td>
<td>0.07***</td>
<td>0.06***</td>
<td>0.07</td>
<td>0.03</td>
<td>0.12***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,000</td>
<td>-30,000 – 10,000</td>
<td>0.08***</td>
<td></td>
<td>0.11*</td>
<td>0.04</td>
<td>0.15***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,000</td>
<td>-30,000 – 10,000</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.06)</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,000</td>
<td>-30,000 – 10,000</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.06)</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10,000</td>
<td>-30,000 – 10,000</td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.05)</td>
<td>(0.02)</td>
<td>(0.03)</td>
</tr>
</tbody>
</table>
Channels of Manipulation Across Wage Earners

IRA Account Manipulation Around Intersection Kink

Non-Ded IRA Contributions

Taxable Income Relative to Intersection Kink ($)

-200000  -100000  0  100000  200000

-5000  0  5000  10000  15000  20000

Ali Abbas (Cornell University)
### Main Results: Period-wise

<table>
<thead>
<tr>
<th>Years</th>
<th>Description</th>
<th>MTR Change</th>
<th>Binwidth</th>
<th>Bunching region</th>
<th>All Filers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993-2011</td>
<td>Entire sample period</td>
<td>28% - 37.5%</td>
<td>10,000</td>
<td>-30,000 – 10,000</td>
<td>0.08***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.03)</td>
</tr>
<tr>
<td>1993-2002</td>
<td>Two acts (OBRA and EGTRRA): High MTR</td>
<td>28% - 39.4%</td>
<td>10,000</td>
<td>-30,000 – 10,000</td>
<td>0.12**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.06)</td>
</tr>
<tr>
<td>2003 - 2011</td>
<td>JGTRRA of 2003</td>
<td>28% - 35%</td>
<td>10,000</td>
<td>-30,000 – 10,000</td>
<td>0.07***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.02)</td>
</tr>
</tbody>
</table>
Some Next Steps

- Variation in kink points allows for estimation under various non-traditional bunching estimators:
  - Already implemented Bertanha et al. (2018) non-parametric bounds

- For taxpayers not facing the AMT-RT effective schedule, assess whether the standard bunching estimation approach generates higher elasticity estimates

- More robust sensitivity analysis using different bandwidths