Lecture 6: Optimal Transfers

Stefanie Stantcheva

Fall 2016
GOALS OF THIS LECTURE

1) Wrap up the study of taxable income responses: migration responses to taxes.

2) Study the optimal design of transfer programs and introduce extensive margin responses into the model.

3) Empirical evidence on responses to transfers.
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 15, Young et al. 14)

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

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 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
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
Control 1 = annualized income between .8 and .9 of threshold  
Control 2 = annualized income between .9 and .995 of threshold.
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 elasticity:
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.
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 \( \rightarrow \uparrow \) 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.
Three Data sources: DID, EPO, PCT

- Inventors: employees, researchers, self-employed.
- "Assignee" is legal owner (firm or individual), can be ≠ 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 ≠ 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 ≠ 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 ≠ 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 \textit{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)

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 (number of tech classes citing patent).

Correlations

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).

Inventor Ranking

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

Correlations

- Patent breadth, breadth of impact

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).

**Inventor Ranking**

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

---

![Graph showing quality in region at time t]

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).

**Inventor Ranking**

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

![Graph showing Top 1% quality in region at time t]
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).

Inventor Ranking

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

Correlations: Patent breadth, breadth of impact

Dynamic, Persistent, Life-time ranking

Top 1-5%
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

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

Top 5-10%

Quality in region at time t
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).

Inventor Ranking

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

Correlations

- Patent breadth, breadth of impact

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).

Inventor Ranking

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

Correlations

Correlations
Patent breadth, breadth of impact

Dynamic, Persistent, Life-time ranking
Link between Inventor Quality and Income in IRS data

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

Income ($)

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

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

\[
\text{Elasticity} = 0.08 \ (0.009)
\]

-2 -1.5 -1 -0.5

Log fraction of top quality domestic inventors

(a) Top quality inventors

\[
\text{Elasticity} = -0.01 \ (0.022)
\]

-2 -1.5 -1 -0.5

Log fraction of low quality domestic 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

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

![Graph showing the relationship between Foreign Top 1% Inventors and Top tax rate differential from 1982 to 1992. The graph illustrates a significant decrease in the top tax rate differential around 1988, which corresponds to a drop in Foreign Top 1% Inventors.]
Case Study: U.S. TRA 1986

![Graph showing the relationship between Foreign Top 1% Inventors and Top tax rate differential from 1982 to 1992.]
Case Study: U.S. TRA 1986

The graph illustrates the change in Foreign Top 1% Inventors and the Top tax rate differential from 1982 to 1992. The data shows a significant drop in Foreign Top 1% Inventors and an increase in the top tax rate differential in 1988, possibly due to the implementation of the U.S. Tax Reform Act of 1986.
Case Study: U.S. TRA 1986

![Graph](image-url)

- **Foreign Top 1% Inventors**:
  - U.S.
  - Synthetic U.S.
- **Top tax rate differential**
Case Study: U.S. TRA 1986

Elasticity = 2.99 (.649)

Foreign Top 1% Inventors

Year

U.S.

Synthetic U.S.

Top tax rate differential


2.5

2

1.5

1

.5


Year

U.S.

Synthetic U.S.

Top tax rate differential
Case Study: U.S. TRA 1986

![Graph showing the structural break in growth of foreign top 1% relative to lower quality inventors.]

**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>Below Top 1%</td>
<td>10.5%</td>
<td>11.3%</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>Below Top 1%</td>
<td>10.5%</td>
<td>11.3%</td>
</tr>
</tbody>
</table>
Case Study: Denmark’s 1992 Preferential Tax Reform

Elasticity = 0.7 (±0.242)

- Determination of the Elasticity: The elasticity measures the responsiveness of the normative share of foreign inventors to changes in the top tax rate differential.

- Trends in Top Tax Rate Differential and Norm. Share of Foreign Inventors:
  - The graph illustrates the top tax rate differential and the normative share of foreign inventors from 1985 to 2005.
  - The elasticity value indicates a moderate response to changes in the top tax rate differential.

- Comparative Analysis:
  - The synthetic Denmark line shows a trend similar to Denmark but with a slight adjustment.
  - The graph highlights the year 1990, noting a significant change in the tax reform.

- Conclusion:
  - The tax reform had a notable impact on the number of foreign inventors in Denmark, as evidenced by the elasticity value and the trend shifts.

Graphical Elements:
- X-axis: Year (1985 to 2005)
- Y-axis: Top tax rate differential
- Line colors and styles:
  - Denmark: solid black line
  - Synthetic Denmark: dashed black line
  - Top tax rate differential: red circular markers
\[ Pr(y_{it} = c) = f(\alpha_{rit} \log(1 - \text{top MTR}^i_{ct}) + \beta_c x_{ti} + \eta x_{cti} + \zeta x_{ct}) \]

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

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

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

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

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

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

- \textbf{Superstars vs. Non-superstars}: include country ∗ 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 → Higher income → higher propensity to be treated by top MTR (MTR ∼ ATR).
**Choice of the Control Group?**

<table>
<thead>
<tr>
<th>Income ($)</th>
<th>Top 1%</th>
<th>Top 1-5%</th>
<th>Top 5-10%</th>
<th>Below Top 25%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,285,405</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>883,970</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>549,460</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>370,975</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>230,774</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Trade-off in the choice of the control group.

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

→ Provide elasticity of top 1% relative to several control groups $g \in \{\text{top 5-10\%}, \text{top10-25\%}, \text{below top 25\%}\}$. 
## Country-by-year Variation and General Equilibrium Effects

<table>
<thead>
<tr>
<th>Term</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>0.894***</td>
<td>0.895***</td>
<td>0.969***</td>
<td>0.955***</td>
</tr>
<tr>
<td></td>
<td>(0.206)</td>
<td>(0.208)</td>
<td>(0.210)</td>
<td>(0.212)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 1-5</td>
<td>0.443***</td>
<td>0.452***</td>
<td>0.522***</td>
<td>0.502***</td>
</tr>
<tr>
<td></td>
<td>(0.133)</td>
<td>(0.134)</td>
<td>(0.133)</td>
<td>(0.134)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 5-10</td>
<td>0.148</td>
<td>0.162</td>
<td>0.233**</td>
<td>0.209*</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.114)</td>
<td>(0.111)</td>
<td>(0.112)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 10-25</td>
<td>-0.123</td>
<td>-0.0987</td>
<td>-0.0222</td>
<td>-0.0460</td>
</tr>
<tr>
<td></td>
<td>(0.0934)</td>
<td>(0.0933)</td>
<td>(0.0890)</td>
<td>(0.0895)</td>
</tr>
<tr>
<td>Log Retention Rate × Below Top 25</td>
<td>-0.405***</td>
<td>-0.349***</td>
<td>-0.267**</td>
<td>-0.278**</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
<td>(0.119)</td>
<td>(0.119)</td>
<td>(0.120)</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>.02</td>
<td>.02</td>
<td>.024</td>
<td>.024</td>
</tr>
<tr>
<td>s.e</td>
<td>(.0047)</td>
<td>(.0047)</td>
<td>(.005)</td>
<td>(.005)</td>
</tr>
<tr>
<td>Foreign elasticity</td>
<td>.75</td>
<td>.754</td>
<td>.811</td>
<td>.811</td>
</tr>
<tr>
<td>s.e</td>
<td>(.174)</td>
<td>(.175)</td>
<td>(.177)</td>
<td>(.177)</td>
</tr>
<tr>
<td>Observations</td>
<td>8644280</td>
<td>8616336</td>
<td>8616336</td>
<td>8616336</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>1.376***</td>
<td>1.508***</td>
<td>1.451***</td>
<td>1.404***</td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.486)</td>
<td>(0.489)</td>
<td>(0.489)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 1-5</td>
<td>0.926**</td>
<td>1.065**</td>
<td>1.004**</td>
<td>0.950**</td>
</tr>
<tr>
<td></td>
<td>(0.449)</td>
<td>(0.455)</td>
<td>(0.458)</td>
<td>(0.457)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 5-10</td>
<td>0.629</td>
<td>0.773*</td>
<td>0.713</td>
<td>0.654</td>
</tr>
<tr>
<td></td>
<td>(0.449)</td>
<td>(0.455)</td>
<td>(0.457)</td>
<td>(0.456)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 10-25</td>
<td>0.357</td>
<td>0.511</td>
<td>0.454</td>
<td>0.396</td>
</tr>
<tr>
<td></td>
<td>(0.441)</td>
<td>(0.447)</td>
<td>(0.448)</td>
<td>(0.447)</td>
</tr>
<tr>
<td>Log Retention Rate × Below Top 25</td>
<td>0.0775</td>
<td>0.263</td>
<td>0.210</td>
<td>0.166</td>
</tr>
<tr>
<td></td>
<td>(0.444)</td>
<td>(0.451)</td>
<td>(0.449)</td>
<td>(0.449)</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 Domestic elasticity</td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.005)</td>
</tr>
<tr>
<td>Foreign elasticity</td>
<td>.63</td>
<td>.62</td>
<td>.62</td>
<td>.63</td>
</tr>
<tr>
<td></td>
<td>(.18)</td>
<td>(.18)</td>
<td>(.19)</td>
<td>(.19)</td>
</tr>
<tr>
<td>Control: Top 10-25 Domestic elasticity</td>
<td>.03</td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.005)</td>
</tr>
<tr>
<td>Foreign elasticity</td>
<td>.85</td>
<td>.84</td>
<td>.83</td>
<td>.84</td>
</tr>
<tr>
<td></td>
<td>(.18)</td>
<td>(.18)</td>
<td>(.18)</td>
<td>(.18)</td>
</tr>
<tr>
<td>Control: Below Top 25 Domestic elasticity</td>
<td>.03</td>
<td>.03</td>
<td>.03</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.006)</td>
<td>(.006)</td>
</tr>
<tr>
<td>Foreign elasticity</td>
<td>1.09</td>
<td>1.05</td>
<td>1.04</td>
<td>1.04</td>
</tr>
<tr>
<td></td>
<td>(.190)</td>
<td>(.196)</td>
<td>(.201)</td>
<td>(.203)</td>
</tr>
<tr>
<td>Observations</td>
<td>8644280</td>
<td>8616336</td>
<td>8616336</td>
<td>8616336</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>1.376***</td>
<td>1.508***</td>
<td>1.451***</td>
<td>1.404***</td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.486)</td>
<td>(0.489)</td>
<td>(0.489)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 1-5</td>
<td>0.926**</td>
<td>1.065**</td>
<td>1.004**</td>
<td>0.950**</td>
</tr>
<tr>
<td></td>
<td>(0.449)</td>
<td>(0.455)</td>
<td>(0.458)</td>
<td>(0.457)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 5-10</td>
<td>0.629</td>
<td>0.773*</td>
<td>0.713</td>
<td>0.654</td>
</tr>
<tr>
<td></td>
<td>(0.449)</td>
<td>(0.455)</td>
<td>(0.457)</td>
<td>(0.456)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 10-25</td>
<td>0.357</td>
<td>0.511</td>
<td>0.454</td>
<td>0.396</td>
</tr>
<tr>
<td></td>
<td>(0.441)</td>
<td>(0.447)</td>
<td>(0.448)</td>
<td>(0.447)</td>
</tr>
<tr>
<td>Log Retention Rate × Below Top 25</td>
<td>0.0775</td>
<td>0.263</td>
<td>0.210</td>
<td>0.166</td>
</tr>
<tr>
<td></td>
<td>(0.444)</td>
<td>(0.451)</td>
<td>(0.449)</td>
<td>(0.449)</td>
</tr>
</tbody>
</table>

Quality × Country FE

Quality × Country FE × Year

Quality × Country FE × Year × Field FE

| Control: Top 5-10 | Domestic elasticity | .02 | .02 | .02 | .02 |
|                  | s.e                | (.005) | (.005) | (.005) | (.005) |
| Foreign elasticity | .63 | .62 | .62 | .63 |
| s.e                | (.18) | (.18) | (.19) | (.19) |

Control: Top 10-25

| Domestic elasticity | .03 | .02 | .02 | .02 |
| s.e                | (.005) | (.005) | (.005) | (.005) |
| Foreign elasticity | .85 | .84 | .83 | .84 |
| s.e                | (.18) | (.18) | (.18) | (.18) |

Control: Below Top 25

| Domestic elasticity | .03 | .03 | .03 | .03 |
| s.e                | (.005) | (.005) | (.006) | (.006) |
| Foreign elasticity | 1.09 | 1.05 | 1.04 | 1.04 |
| s.e                | (.190) | (.196) | (.201) | (.203) |

Observations

8644280  8616336  8616336  8616336
## 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>1.376***</td>
<td>1.508***</td>
<td>1.451***</td>
<td>1.404***</td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.486)</td>
<td>(0.489)</td>
<td>(0.489)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 1-5</td>
<td>0.926**</td>
<td>1.065**</td>
<td>1.004**</td>
<td>0.950**</td>
</tr>
<tr>
<td></td>
<td>(0.449)</td>
<td>(0.455)</td>
<td>(0.458)</td>
<td>(0.457)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 5-10</td>
<td>0.629</td>
<td>0.773*</td>
<td>0.713</td>
<td>0.654</td>
</tr>
<tr>
<td></td>
<td>(0.449)</td>
<td>(0.455)</td>
<td>(0.457)</td>
<td>(0.456)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 10-25</td>
<td>0.357</td>
<td>0.511</td>
<td>0.454</td>
<td>0.396</td>
</tr>
<tr>
<td></td>
<td>(0.441)</td>
<td>(0.447)</td>
<td>(0.448)</td>
<td>(0.447)</td>
</tr>
<tr>
<td>Log Retention Rate × Below Top 25</td>
<td>0.0775</td>
<td>0.263</td>
<td>0.210</td>
<td>0.166</td>
</tr>
<tr>
<td></td>
<td>(0.444)</td>
<td>(0.451)</td>
<td>(0.449)</td>
<td>(0.449)</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>
</tbody>
</table>

### Control: Top 5-10
- Domestic elasticity 0.02 0.02 0.02 .02
- s.e (0.005) (0.005) (0.005) (0.005)
- Foreign elasticity 0.63 0.62 0.62 0.63
- s.e (0.18) (0.18) (0.19) (0.19)

### Control: Top 10-25
- Domestic elasticity 0.03 0.02 0.02 0.02
- s.e (0.005) (0.005) (0.005) (0.005)
- Foreign elasticity 0.85 0.84 0.83 0.84
- s.e (0.18) (0.18) (0.18) (0.18)

### Control: Below Top 25
- Domestic elasticity 0.03 0.03 0.03 0.03
- s.e (0.005) (0.005) (0.006) (0.006)
- Foreign elasticity 1.09 1.05 1.04 1.04
- s.e (0.190) (0.196) (0.201) (0.203)

Observations

- 8644280 8616336 8616336 8616336
# 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>1.376***</td>
<td>1.508***</td>
<td>1.451***</td>
<td>1.404***</td>
</tr>
<tr>
<td></td>
<td>(0.478)</td>
<td>(0.486)</td>
<td>(0.489)</td>
<td>(0.489)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 1-5</td>
<td>0.926**</td>
<td>1.065**</td>
<td>1.004**</td>
<td>0.950**</td>
</tr>
<tr>
<td></td>
<td>(0.449)</td>
<td>(0.455)</td>
<td>(0.458)</td>
<td>(0.457)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 5-10</td>
<td>0.629</td>
<td>0.773*</td>
<td>0.713</td>
<td>0.654</td>
</tr>
<tr>
<td></td>
<td>(0.449)</td>
<td>(0.455)</td>
<td>(0.457)</td>
<td>(0.456)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 10-25</td>
<td>0.357</td>
<td>0.511</td>
<td>0.454</td>
<td>0.396</td>
</tr>
<tr>
<td></td>
<td>(0.441)</td>
<td>(0.447)</td>
<td>(0.448)</td>
<td>(0.447)</td>
</tr>
<tr>
<td>Log Retention Rate × Below Top 25</td>
<td>0.0775</td>
<td>0.263</td>
<td>0.210</td>
<td>0.166</td>
</tr>
<tr>
<td></td>
<td>(0.444)</td>
<td>(0.451)</td>
<td>(0.449)</td>
<td>(0.449)</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>.02</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>s.e</td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.005)</td>
</tr>
<tr>
<td></td>
<td>Foreign elasticity</td>
<td>.63</td>
<td>.62</td>
<td>.62</td>
</tr>
<tr>
<td></td>
<td>s.e</td>
<td>(.18)</td>
<td>(.18)</td>
<td>(.19)</td>
</tr>
<tr>
<td>Control: Top 10-25</td>
<td>Domestic elasticity</td>
<td>.03</td>
<td>.02</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>s.e</td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.005)</td>
</tr>
<tr>
<td></td>
<td>Foreign elasticity</td>
<td>.85</td>
<td>.84</td>
<td>.83</td>
</tr>
<tr>
<td></td>
<td>s.e</td>
<td>(.18)</td>
<td>(.18)</td>
<td>(.18)</td>
</tr>
<tr>
<td>Control: Below Top 25</td>
<td>Domestic elasticity</td>
<td>.03</td>
<td>.03</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>s.e</td>
<td>(.005)</td>
<td>(.005)</td>
<td>(.006)</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>(.190)</td>
<td>(.196)</td>
<td>(.201)</td>
</tr>
<tr>
<td>Observations</td>
<td>8644280</td>
<td>8616336</td>
<td>8616336</td>
<td>8616336</td>
</tr>
</tbody>
</table>
## The Role of Companies

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Retention Rate × Top 1</td>
<td>1.400***</td>
<td>0.980*</td>
</tr>
<tr>
<td></td>
<td>(0.500)</td>
<td>(0.537)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 1-5</td>
<td>0.868*</td>
<td>0.548</td>
</tr>
<tr>
<td></td>
<td>(0.473)</td>
<td>(0.493)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 5-10</td>
<td>0.514</td>
<td>0.199</td>
</tr>
<tr>
<td></td>
<td>(0.473)</td>
<td>(0.491)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 10-25</td>
<td>0.181</td>
<td>-0.0974</td>
</tr>
<tr>
<td></td>
<td>(0.468)</td>
<td>(0.481)</td>
</tr>
<tr>
<td>Log Retention Rate × Below Top 25</td>
<td>-0.254</td>
<td>-0.560</td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.485)</td>
</tr>
<tr>
<td>Log Retention Rate × Not Multinational</td>
<td><strong>-0.216</strong>*</td>
<td><em>(0.129)</em></td>
</tr>
<tr>
<td>Log Retention Rate × Activity abroad</td>
<td>-1.470***</td>
<td><em>(0.137)</em></td>
</tr>
</tbody>
</table>

|                         |                         |                         |
| 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: .018, s.e (.0045) vs .011, s.e (.0047)
- Foreign elasticity: .809, s.e (.201) vs .420, s.e (.154)

### Control: Top 10-25
- Domestic elasticity: .024, s.e (.0045) vs .016, s.e (.0046)
- Foreign elasticity: 1.113, s.e (.197) vs .579, s.e (.151)

### Control: Below Top 25
- Domestic elasticity: .034, s.e (.0047) vs .027, s.e (.0049)
- Foreign elasticity: 1.511, s.e (.211) vs .828, s.e (.159)

### Observations
- 7059856
- 6168504
### The Role of Companies

<table>
<thead>
<tr>
<th>Interaction</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Retention Rate × Top 1</td>
<td>1.400***</td>
<td>0.980*</td>
</tr>
<tr>
<td></td>
<td>(0.500)</td>
<td>(0.537)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 1-5</td>
<td>0.868*</td>
<td>0.548</td>
</tr>
<tr>
<td></td>
<td>(0.473)</td>
<td>(0.493)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 5-10</td>
<td>0.514</td>
<td>0.199</td>
</tr>
<tr>
<td></td>
<td>(0.473)</td>
<td>(0.491)</td>
</tr>
<tr>
<td>Log Retention Rate × Top 10-25</td>
<td>0.181</td>
<td>-0.0974</td>
</tr>
<tr>
<td></td>
<td>(0.468)</td>
<td>(0.481)</td>
</tr>
<tr>
<td>Log Retention Rate × Below Top 25</td>
<td>-0.254</td>
<td>-0.560</td>
</tr>
<tr>
<td></td>
<td>(0.472)</td>
<td>(0.485)</td>
</tr>
<tr>
<td>Log Retention Rate × Not Multinational</td>
<td>-0.216*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.129)</td>
<td></td>
</tr>
<tr>
<td>Log Retention Rate × Activity abroad</td>
<td>-1.470***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interaction</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality × Country FE</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Quality × Country FE × Year</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Quality × Country FE × Year × Field FE</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

| Control: Top 5-10                                | Domestic elasticity | .018            | .011            |
|                                                  | s.e               | (.0045)         | (.0047)         |
|                                                  | Foreign elasticity | .809            | .420            |
|                                                  | s.e               | (.201)          | (.154)          |

| Control: Top 10-25                               | Domestic elasticity | .024            | .016            |
|                                                  | s.e               | (.0045)         | (.0046)         |
|                                                  | Foreign elasticity | 1.113           | .579            |
|                                                  | s.e               | (.197)          | (.151)          |

| Control: Below Top 25                            | Domestic elasticity | .034            | .027            |
|                                                  | s.e               | (.0047)         | (.0049)         |
|                                                  | Foreign elasticity | 1.511           | .828            |
|                                                  | s.e               | (.211)          | (.159)          |

| Observations                                     | 7059856          | 6168504         |
OPTIMAL TRANSFERS: MIRRLEES MODEL

Mirrlees model predicts that optimal transfer at bottom takes the form of a “Negative Income Tax”:

1) Lumpsum grant \(-T(0)\) for those with no earnings

2) High MTRs \(T'(z)\) at the bottom to phase-out the lumpsum grant quickly

Intuition: high MTRs at bottom are efficient because:

(a) they target transfers to the most needy

(b) earnings at the bottom are low to start with so intensive response does not generate large output losses

Diamond-Saez JEP’11: \(T'(0) = (g_0 - 1)/(g_0 - 1 + e_0)\) with \(e_0\) elasticity of the fraction non-working wrt to \(1 - T'(0)\) and \(g_0\) social marginal welfare weight on non workers

\[\Rightarrow T'(0)\] large: e.g. \(g_0 = 3\) and \(e_0 = 0.5 \Rightarrow T'(0) = 80\%\]
Reform: Increase $\tau_1$ by $d\tau_1$ and $c_0$ by $dc_0 = z_1 d\tau_1$

1) Mechanical fiscal cost: $dM = -H_0 dc_0 = -H_0 z_1 d\tau_1$
2) Welfare effect: $dW = g_0 H_0 dc_0 = g_0 H_0 z_1 d\tau_1$
3) Fiscal cost due to behavioral responses:
$$dB = -dH_0 \tau_1 z_1 = d\tau_1 e_0 H_0 \tau_1/(1-\tau_1) z_1$$

Optimal phase-out rate $\tau_1$:
$$dM + dW + dB = 0 \implies \frac{\tau_1}{1-\tau_1} = \frac{g_0-1}{e_0}$$
Optimal Transfers: Participation Responses

Empirical literature shows that participation labor supply responses [due to fixed costs of working] are large at the bottom [much larger and clearer than intensive responses]

Diamond JpubE’80, Saez QJE’02, Laroque EMA’05 incorporate such extensive labor supply responses in the optimal income tax model

Participation depends on participation tax rate: \( \tau_p = \frac{T(z) - T(0)}{z} \): individual keeps fraction \( 1 - \tau_p \) of earnings when moving from zero earnings to earnings \( z \):

\[
z - T(z) = -T(0) + z - [T(z) - T(0)] = -T(0) + z \cdot (1 - \tau_p)
\]

Key result: in-work subsidies with \( T'(z) < 0 \) (such as EITC) become optimal when labor supply responses are concentrated along extensive margin and social marginal welfare weight on low skilled workers \( > 1 \).
Model with discrete earnings outcomes: \( w_0 = 0 < w_1 < \ldots < w_I \)

Tax/transfer \( T_i \) when earning \( w_i \), \( c_i = w_i - T_i \)

Participation labor supply: Skill \( i \) individual compares \( c_i \) and \( c_0 \) when deciding to work \( \Rightarrow \) Participation tax rate \( \tau_i \) such that \( c_i - c_0 = w_i \cdot (1 - \tau_i) \)

Person works iff \( c_i - \theta \geq c_0 \) where \( \theta \) is fixed cost of working

\( \Rightarrow \) In aggregate, fraction \( h_i (c_i - c_0) \) of population earns \( w_i \)

Participation elasticity \( e_i = (c_i - c_0) / h_i \cdot \partial h_i / \partial (c_i - c_0) \)

Social Welfare function is summarized by social marginal welfare weights at each earnings level \( g_i \downarrow i \), and average to one \( \sum_i g_i h_i = 1 \) (if no income effects)
Starting from a Means-Tested Program

Consumption $c$

Earnings $w$

$45^\circ$ $G$

$0$ $w^*$

Source: revised version of Saez (2002), p. 1050
Starting from a Means-Tested Program

Introducing a small EITC is desirable for redistribution

Source: revised version of Saez (2002), p. 1050
Introducing a small EITC is desirable for redistribution. Participation response saves government revenue.

Source: revised version of Saez (2002), p. 1050
Starting from a positive phasing-out rate \( \tau_1 > 0 \):

1) Increasing transfers by \( dc_1 \) at \( z_1 \) is desirable for redistribution: net effect \( (g_1-1)h_1 dc_1 > 0 \) if \( g_1 > 1 \)

2) Participation response saves government revenue

\[ \tau_1 z_1 dh_1 = e_1 \frac{\tau_1}{(1-\tau_1)} h_1 dc_1 > 0 \]

\( \Rightarrow \) Win-win reform …if intensive response is small

Optimal phase-out rate \( \tau_1 \):

\[ (g_1-1)h_1 dc_1 + e_1 \frac{\tau_1}{(1-\tau_1)} h_1 dc_1 = 0 \]

\( \Rightarrow \frac{\tau_1}{(1-\tau_1)} = \frac{(1-g_1)}{e_1} < 0 \) if \( g_1 > 1 \)
Figure 3a: Optimal Tax/Transfer Derivation

Source: revised version of Saez (2002), p. 1052
Figure 3a: Optimal Tax/Transfer Derivation (assuming $g_1 > 1$)

Consumption $c$

Wage $w$

Fiscal Effect: $-h_1 dc_1 < 0$

Welfare Effect: $h_1 g_1 dc_1 > 0$

Source: revised version of Saez (2002), p. 1052
Figure 3a: Optimal Tax/Transfer Derivation (assuming $g_1 > 1$)

Net Welfare effect: $h_1 d c_1 (g_1 - 1) > 0$

Labor Supply: $d h_1 w_1 \tau_1 < 0$

Source: revised version of Saez (2002), p. 1052
Figure 3a: Optimal Tax/Transfer Derivation (assuming $g_1 > 1$)

At the optimum:
\[ dh_1 w_1 \tau_1 + h_1 dc_1 (g_1 - 1) = 0 \]
implies
\[ \tau_1 / (1 - \tau_1) = (1 - g_1) / e_1 < 0 \]

Source: revised version of Saez (2002), p. 1052
Small reform $dc_i = -dT_i > 0$. Three effects:

1) Mechanical Change in tax revenue $dM = h_i dT_i$

2) Behavioral Effect: $dh_i = -e_i h_i dT_i / (c_i - c_0) \Rightarrow$ Tax loss: $dB = -(T_i - T_0) dh_i = -e_i h_i dT_i (T_i - T_0) / (c_i - c_0)$

3) Welfare Effect: each worker in job $i$ looses $dT_i$ so welfare loss $dW = -g_i h_i dT_i$ [No first order welfare loss for switchers]

FOC: $dM + dB + dW = 0 \Rightarrow$

$$\frac{\tau_i}{1 - \tau_i} = \frac{T_i - T_0}{c_i - c_0} = \frac{1}{e_i} (1 - g_i)$$

$g_1 > 1 \Rightarrow T_1 - T_0 < 0 \Rightarrow$ in-work subsidy
ACTUAL TAX/TRANSFER SYSTEMS

1) Transfer programs used to be of the traditional form with high phasing-out rates (sometimes above 100%) ⇒ No incentives to work (even with modest elasticities)

Initially designed for groups not expected to work [widows in the US] but later attracting groups who could potentially work [single mothers]

2) In-work benefits have been introduced and expanded in OECD countries since 1980s (US EITC, UK Family Credit, etc.) and have been politically successful ⇒ (a) Redistribute to low income workers, (b) improve incentives to work
We have assumed that \( T(z) \) depends only on earnings \( z \).

In reality, govt can observe many other characteristics \( X \) also correlated with ability [gender, race, age, disability, family structure, height,...] and set \( T(z, X) \). Two theory results:

1) If characteristic \( X \) is **immutable** then redistribution across the \( X \) groups will be complete [until average social marginal welfare weights are equated across \( X \) groups]

2) If characteristic \( X \) can be **manipulated** [behavioral response or cheating] but \( X \) correlated with ability then taxes will still depend on both \( X \) and \( z \).

References: Akerlof AER’78 (welfare), Nichols-Zeckhauser AER’82 (welfare), Weinzierl ’11 (age), Mankiw-Weinzierl ’10 (height), Kaplow ’08 (chapter 7)
Consider $X$ binary immutable (Talls vs. Shorts)

With $T(z)$ independent of $X$, Talls have higher ability on average $\Rightarrow$ Average social marginal welfare weights $\bar{\bar{g}}^T < \bar{\bar{g}}^S$ $\Rightarrow$ Transfer from Talls to Shorts is desirable (surtax on Talls which finances an allowance on Shorts)

Optimal height transfers should be up to point where $\bar{\bar{g}}^T = \bar{\bar{g}}^S$

Mankiw-Weinzierl ’09 compute the optimal $T^{Tall}(z)$ and $T^{Short}(z)$ based on calibrated mode: optimal transfer $T^{Tall}(z) - T^{Short}(z)$ not trivial ($\simeq 10\%$ of income)

They also show that you can get a (very modest) Pareto improvement using taxes on height and income instead of only income
PROBLEM WITH TAGGING

In practice public would oppose height based redistribution because height does not cause high earnings ⇒

1) **Horizontal Equity** concerns [people with same “ability-to-pay” should pay the same tax] impose constraints on feasible policies [not captured by utilitarian framework]

2) Constrained optimization analysis [$T(z)$ instead of $T(z, X)$] remains valid even with heterogeneity in preferences

3) In practice $T(z, X)$ depends on $X$ only when $X$ is directly related to welfare [family structure, # kids, medical expenses] or ability to earn [disability status] (“ability-to-pay” intuition)
IN-KIND REDISTRIBUTION

Majority of actual transfers are in-kind (health care, child care, education, public housing, nutrition subsidies)

1) **Rational Individual** perspective:

(a) In-kind transfer is **tradeable** at market price ⇒ in-kind equivalent to cash

(b) In-kind transfer **non-tradeable** ⇒ in-kind inferior to cash.
IN-KIND REDISTRIBUTION

2) Social perspective: 4 justifications:

a) Commodity Egalitarianism: some goods (education, health, shelter, food) seen as **rights** and ought to be provided to all

b) Paternalism: society imposes its preferences on recipients [recipients prefer cash]

c) Behavioral: Recipients do not make choices in their best interests (self-control, myopia) [recipients understand that in-kind is better for them]

d) Under standard welfarist objective: Efficiency considerations in a 2nd best context
EFFICIENCY OF IN-KIND REDISTRIBUTION

Depends on what income tax tools are available:

1) No income tax: Income $z$ not observable (devo countries) $\Rightarrow$ In-kind provision or subsidies for necessities desirable

2) Linear tax model (Ramsey): Guesnerie–Roberts EMA'84 $\Rightarrow$ rationing goods encouraged by the tax system is desirable [and forcing consumption of goods discouraged by tax]

3) Nonlinear income tax: Under Atkinson–Stiglitz assumption [weak-separability and homogeneity $U^h(v(c_1, \ldots, c_K), z))$] $\Rightarrow$ Any distortion (quota, rationing, subsidy) involving $c$ choices not desirable provided $T(z)$ optimal

If good $c_k$ related to leisure/ability [soup kitchen with queuing requirement] then A-S fails and in-kind redistribution possibly desirable even with optimal $T(z)$
IMPOSING ORDEALS ON TRANSFER RECIPIENTS

Many actual transfer programs impose requirements on beneficiaries (complex application, job search, training, or work requirements) and hence have low take-up (often $< 50\%$)

1) If social objective is welfarist and income $z$ observable: ordeals unlikely to be desirable:

Compare ordeal to benefit cut: (a) only benefit cut saves money mechanically, (b) both reduce welfare of recipients, (c) both reduce take-up [good fiscally]

Need implausible sorting effects for ordeal to be desirable [e.g., ordeal does not hurt much deserving beneficiaries and discourages undeserving take-up, conditional on $z$]

2) If $z$ is not observable then ordeal could be desirable (kitchen soup line)

3) With non-welfarist objective [such as poverty alleviation], ordeal can be desirable [Besley-Coate AER’92]
WORK RESTRICTIONS AND MINIMUM WAGE

Minimum wage creates rationing of low skilled work. Could minimum wage be desirable on top of nonlinear tax/transfer?

Lee and Saez JpubE’12 use a job choice model [Saez QJE ’02 with endogenous wages]. Two results:

1) Minimum wage desirable if (a) govt wants to redistribute to low skilled workers \((g_1 > 1)\) and (b) rationing created by min wage is efficient

2) If labor supply responses along extensive margin only then minimum wage with positive tax rate on low skilled work \(\tau_1 > 0\) is 2nd best Pareto inefficient

\(\Rightarrow\) EITC and min wage are complementary
2. Optimal Tax/Transfer System (no min wage)

Consumption \( c \)

Wage \( w \)

\( c_0 \)

\( c_1 \)

\( c_2 \)

\( 45^\circ \)

\( w_1 \)

\( w_2 \)

Source: Lee and Saez (2008)
2. Set Min wage $\bar{w}=w_1$ and increase $c_1$ by $dc_1$

**Welfare Effect > Direct Fiscal Effect** if govt values redistribution to low skill workers

Source: Lee and Saez (2008)
2. Desirability of Min Wage with Optimal Taxes

Consumption $c$

Wage $w$

$w = w_1$

$c_0$

$c_1$

$c_2$

Welfare Effect $>$ Direct Fiscal Effect if govt values redistribution to low skill workers

dc$_1$ > 0 makes low skilled job $w_1$ more attractive $\Rightarrow$ would reduce $w_1$ through demand effects

Source: Lee and Saez (2008)
2. Desirability of Min Wage with Optimal Taxes

With min wage set at $w_1$, $dc_1 > 0$ does not affect labor supply because $w_1$ cannot go down.

$\Rightarrow$ No indirect fiscal effect

$\Rightarrow$ Welfare increases

Welfare Effect > Direct Fiscal Effect if govt values redistribution to low skill workers

Source: Lee and Saez (2008)
3. Pareto Improving Policy when $\tau_1 > 0$ and min wage binds

Source: Lee and Saez (2008)
3. Pareto Improving Policy when $\tau_1 > 0$ and min wage binds

Reduce $\bar{w}$ while keeping $c_1, c_2$ constant:

No direct fiscal effect of $d\bar{w}, dw_2$ as $h_1 d\bar{w} + h_2 dw_2 = 0$ (no profits)

and tax = $(\bar{w} - c_1) h_1 + (w_2 - c_2) h_2$

Source: Lee and Saez (2008)
3. Pareto Improving Policy when $\tau_1 > 0$ and min wage binds

Consumption $c$

Wage $w$

Unemployment decreases $\Rightarrow$ New Workers better off and pay more taxes $\Rightarrow$ Pareto Improvement

Reduce $\bar{w}$ while keeping $c_1$, $c_2$ constant:

No direct fiscal effect of $d\bar{w}$, $dw_2$ as $h_1d\bar{w} + h_2dw_2 = 0$ (no profits)

and tax = ($\bar{w}$-$c_1$) $h_1 + (w_2-c_2)$ $h_2$

Source: Lee and Saez (2008)
FAMILY TAXATION: MARRIAGE AND CHILDREN

Two important issues in policy debate:

1) Marriage: What is the optimal taxation of couples vs. singles? Should secondary earnings be treated differently?

2) Children: What should be the net transfer (transfer or tax reduction) for family with children (as a function of family income and structure)?

Theoretical literature is not great in part because utilitarian framework is not satisfactory.
TAXATION OF COUPLES

1) Economies of scale and sharing in consumption within families $\Rightarrow$ Welfare best measured by family income relative to size [$\equiv$ normalized income]

$\Rightarrow$ Taxes/Transfers should be based on normalized family income which can create a marriage penalty / subsidy

Note: Impossible to have a tax/transfer system that

(1) is family income based $T(z^h + z^w)$

(2) has marriage neutrality $T(z^h, z^w) = T(z^h) + T(z^w)$

(3) is progressive (i.e., not strictly linear)

Proof: (1)+(2) $\Rightarrow T(z^h + z^w) = T(z^h) + T(z^w) \Rightarrow T(z) = \tau \cdot z$
2) If marriage responds to tax/transfer differential ⇒ better to reduce marriage penalty and move toward individualized system

Particularly important cohabitation is close substitute to marriage (Scandinavian countries)

3) Labor supply of secondary earners more elastic than labor supply of primary earner ⇒ Secondary earnings should be taxed less (standard Ramsey intuition, Boskin-Sheshinski JpubE’83)

But labor supply elasticity differential is decreasing as earnings gender gap decreases [Blau and Kahn JOLE’07]

In OECD countries: income tax systems have become individual based but means tested transfers have remained family based
TRANSFERS OR TAX CREDITS FOR CHILDREN

1) Children reduce **normalized income** ⇒ Transfer for children $T_{kid}$ should be positive

In practice, transfers for children are always positive

2) Should $T_{kid}(z)$ increase with income $z$?

Pro: they reduce normalized income most for upper earners [e.g., France computes taxes as $N \cdot T(z/N)$ where $N$ is # family members, kids count as .5 ⇒ $T_{kid}(z)$ increases with $z$].

Cons: lower earners need child transfers most [most OECD countries have means-tested transfers conditional on number of kids ⇒ $T_{kid}(z)$ decreases with $z$, US has $T_{kid}(z)$ inverted U-shape due to EITC and Child Tax Credit]
TRANSFERS OR TAX CREDITS FOR CHILDREN

3) Family does not make decisions as a single unit (Chiappori JPE’92): transfers to mothers has bigger effects on children’s consumption than transfers to fathers [Lundberg et al. ’97, Duflo ’03]

4) Children create externalities [positive: pay-as-you-go retirement programs, negative: global warming]. If fertility responds to transfers, case for subsidizing/taxing children

5) Child care costs are positively related to work ⇒ Such costs should be subsidized by Atkinson-Stiglitz [often they are in practice]:

Public pre-kindergarten in Europe is a huge in-work subsidy for mothers ⇒ Large effect on mothers’ labor force participation (bigger effect than US EITC)
CHILDREN AND LIMITS OF UTILITARIAN MODEL

If fertility decisions unrelated to children tax/transfers ⇒ Social marginal utility should be equated across families with 0 children, families with 1 child, etc.

If ability uncorrelated with children ⇒ Families with kids will get fully compensating transfers

If ability positively correlated with children ⇒ Families with kids might be taxed more heavily [as in the height tax case]

Seems an absurd model to think about transfers for children ⇒ Need to come up with more realistic alternative
EITC Behavioral Studies

Strong evidence of response along extensive margin, little evidence of response along intensive margin (except for self-employed) ⇒ Possibly due to lack of understanding of the program

Qualitative surveys show that:

Low income families know about EITC and understand that they get a tax refund if they work

However very few families know whether tax refund ↑ or ↓ with earnings

Such confusion might be good for the government as the EITC induces work along participation margin without discouraging work along intensive margin
Use US population wide tax return data since 1996 (through IRS special contract)

1) Substantial heterogeneity in fraction of EITC recipients bunching (using self-employment) across geographical areas
   ⇒ Information on EITC varies across areas and grows overtime

2) Places with high self-employment EITC bunching display wage earnings distribution more concentrated around plateau

3) Omitted variable test: use birth of first child to test causal effect of EITC on wage earnings
   ⇒ Evidence of wage earnings response to EITC along intensive margin
Earnings Distributions in Lowest and Highest Bunching Deciles

Source: Chetty, Friedman, and Saez NBER'12
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 1996

Source: Chetty, Friedman, and Saez NBER'12
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 1999

Source: Chetty, Friedman, and Saez NBER'12
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2002

Source: Chetty, Friedman, and Saez NBER'12
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2005

Source: Chetty, Friedman, and Saez NBER’12
Fraction of Tax Filers Who Report SE Income that Maximizes EITC Refund in 2008

Source: Chetty, Friedman, and Saez NBER'12
Income Distribution For Single Wage Earners with One Child

Is the EITC having an effect on this distribution?

Source: Chetty, Friedman, and Saez NBER'12
Income Distribution For Single Wage Earners with One Child
High vs. Low Bunching Areas

Source: Chetty, Friedman, and Saez NBER'12
Earnings Distribution in the Year Before First Child Birth for Wage Earners

Percent of Individuals

- 2%
- 4%
- 0%
- 6%

Wage Earnings
Lowest Sharp Bunching Decile
Middle Sharp Bunching Decile
Highest Sharp Bunching Decile

Source: Chetty, Friedman, and Saez NBER'12
IMPLICATIONS OF ROLE OF INFORMATION

Empirical work:

Information should be a key explanatory variable in estimation of behavioral responses to govt programs

When doing empirical project, always ask the question: did people affected understand incentives?

Cannot identify structural parameters of preferences without modeling information and salience

Normative analysis:

Information is a powerful and inexpensive policy tool to affect behavior

Should be incorporated into optimal policy design problems
Culture of Welfare across Generations

Conservative concern that welfare promotes a culture of dependency: kids growing up in welfare supported families are more likely to use welfare

Correlation in welfare use across generations is obviously not necessarily causal

Dahl, Kostol, Mogstad QJE’2014 analyze causal effect of parental use of Disability Insurance (DI) on children use (as adults) of DI in Norway

Identification uses random assignment of judges to denied DI applicants who appeal [some judges are severe, some lenient]

Find evidence of causality: parents on DI increases odds of kids on DI over next 5 years by 6 percentage points

Mechanism seems to be learning about DI availability rather than reduced stigma from using DI [because no effect on other welfare programs use]
Figure 3: Effect of Judge Leniency on Parents (First Stage) and Children (Reduced Form).

Notes: Baseline sample, consisting of parents who appeal an initially denied DI claim during the period 1989-2005 (see Section 3 for further details). There are 14,893 individual observations and 79 different judges. Panel (A): Solid line is a local linear regression of parental DI allowance on judge leniency. Panel (B): Solid line is a local linear regression of child DI receipt on their parent’s judge leniency measure. All regressions include fully interacted year and department dummies. The histogram of judge leniency is shown in the background of both figures (top and bottom 0.5% excluded from the graph).

Source: Dahl, Kostol, Mogstad (2013)
REFERENCES

All papers referenced in this lecture were cited in the Reading Lists of Lectures 1-5.