230B: Public Economics
Labor Supply Responses to Taxes and Transfers

Emmanuel Saez

Berkeley
MOTIVATION

1) Labor supply responses to taxation are of fundamental importance for income tax policy [efficiency costs and optimal tax formulas]

2) Labor supply responses along many dimensions:

(a) Intensive: hours of work on the job, intensity of work, occupational choice [including education]

(b) Extensive: whether to work or not [e.g., retirement and migration decisions]

3) Reported earnings for tax purposes can also vary due to (a) tax avoidance [legal tax minimization], (b) tax evasion [illegal under-reporting]

4) Different responses in short-run and long-run: long-run response most important for policy but hardest to estimate
STATIC MODEL: SETUP

Baseline model: (a) static, (b) linearized tax system, (c) pure intensive margin choice, (d) single hours choice, (e) no frictions

Let $c$ denote consumption and $l$ hours worked, utility $u(c, l)$ increases in $c$, and decreases in $l$

Individual earns wage $w$ per hour (net of taxes) and has $y$ in non-labor income

Key example: pre-tax wage rate $w^p$ and linear tax system with tax rate $\tau$ and demogrant $G \Rightarrow c = w^p (1 - \tau)l + G$

Individual solves

$$\max_{c,l} u(c, l) \quad \text{subject to} \quad c = wl + y$$
LABOR SUPPLY BEHAVIOR

FOC: \( wu_c + u_l = 0 \) defines uncompensated (Marshallian) labor supply function \( l^u(w, y) \)

Uncompensated elasticity of labor supply: \( \varepsilon^u = \frac{w}{l} \frac{\partial l^u}{\partial w} \) [\% change in hours when net wage \( w \uparrow \) by 1%]

Income effect parameter: \( \eta = w \frac{\partial l}{\partial y} \leq 0 \): $ increase in earnings if person receives $1 extra in non-labor income

Compensated (Hicksian) labor supply function \( l^c(w, u) \) which minimizes cost \( wl - c \) st to constraint \( u(c, l) \geq u \).

Compensated elasticity of labor supply: \( \varepsilon^c = \frac{w}{l} \frac{\partial l^c}{\partial w} > 0 \)

Slutsky equation: \( \frac{\partial l}{\partial w} = \frac{\partial l^c}{\partial w} + l \frac{\partial l}{\partial y} \Rightarrow \varepsilon^u = \varepsilon^c + \eta \)
$l = \text{consumption}$

Marshallian Labor Supply $l(w, R)$

Indifference Curve

Labor Supply Theory
Labor Supply Theory

\[ c = consumption \]

utility \( u \)

slope = \( w \)

Hicksian Labor Supply \( l^c(w, u) \)
Labor Supply Income Effect

\[ l(w, R + \Delta R) \leq l(w, R) \]

\[ \eta = w \frac{\partial l}{\partial R} \leq 0 \]
Labor Supply Substitution Effect

\[ \varepsilon_c = \frac{w \partial l^c}{l \partial w} > 0 \]

where

\[ l^c(w, u) \]

and

\[ l^c(w + \Delta w, u) \]
Uncompensated Labor Supply Effect

\[ \varepsilon^u = \varepsilon^c + \eta \]

slope = \( w + \Delta w \)

income effect
\[ \eta \leq 0 \]

substitution effect: \( \varepsilon^c > 0 \)

\[ l(w, R) \]

\[ l(w + \Delta w, R) \]

labor supply \( l \)
BASIC CROSS SECTION ESTIMATION

Data on hours or work, wage rates, non-labor income started becoming available in the 1960s when first micro surveys and computers appeared:

Simple OLS regression:

\[ l_i = \alpha + \beta w_i + \gamma y_i + X_i \delta + \epsilon_i \]

\( w_i \) is the net-of-tax wage rate

\( y_i \) measures non-labor income [including spousal earnings for couples]

\( X_i \) are demographic controls [age, experience, education, etc.]

\( \beta \) measures uncompensated wage effects, and \( \gamma \) income effects [can be converted to \( \varepsilon^u, \eta \)]
BASIC CROSS SECTION RESULTS


    a) Small effects $\epsilon_u = 0, \eta = -0.1, \epsilon^c = 0.1$ with some variation across estimates (sometimes $\epsilon^c < 0$).

2. Female workers [secondary earners when married] (Killingsworth and Heckman, 1986):

    Much larger elasticities on average, with larger variations across studies. Elasticities go from zero to over one. Average around 0.5. Significant income effects as well

    Female labor supply elasticities have declined overtime as women become more attached to labor market (Blau-Kahn JOLE’07)
KEY ISSUE: $w$ correlated with tastes for work

$$l_i = \alpha + \beta w_i + \gamma y_i + \epsilon_i$$

Identification is based on cross-sectional variation in $w_i$: comparing hours of work of highly skilled individuals (high $w_i$) to hours of work of low skilled individuals (low $w_i$)

If highly skilled workers have more taste for work (independent of the wage effect), then $\epsilon_i$ is positively correlated with $w_i$ leading to an upward bias in OLS

Plausible scenario: hard workers acquire better education and hence have higher wages

Controlling for $X_i$ can help but can never be sure that we have controlled for all the factors correlated with $w_i$ and tastes for work: **Omitted variable bias**

⇒ Tax changes provide more compelling identification
Negative Income Tax (NIT) Experiments

1) Best way to resolve identification problems: exogenously change taxes/transfers with a randomized experiment

2) NIT experiment conducted in 1960s/70s in Denver, Seattle, and other cities

3) First major social experiment in U.S. designed to test proposed transfer policy reform

4) Provided lump-sum welfare grants $G$ combined with a steep phaseout rate $\tau$ (50%-80%) [based on family earnings]


6) Several groups, with randomization within each; approx. $N = 75$ households in each group
### Table 1
**Parameters of the 11 Negative Income Tax Programs**

<table>
<thead>
<tr>
<th>Program Number</th>
<th>G ($)</th>
<th>$\tau$</th>
<th>Declining Tax Rate</th>
<th>Break-even Income ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3,800</td>
<td>.5</td>
<td>No</td>
<td>7,600</td>
</tr>
<tr>
<td>2</td>
<td>3,800</td>
<td>.7</td>
<td>No</td>
<td>5,429</td>
</tr>
<tr>
<td>3</td>
<td>3,800</td>
<td>.7</td>
<td>Yes</td>
<td>7,367</td>
</tr>
<tr>
<td>4</td>
<td>3,800</td>
<td>.8</td>
<td>Yes</td>
<td>5,802</td>
</tr>
<tr>
<td>5</td>
<td>4,800</td>
<td>.5</td>
<td>No</td>
<td>9,600</td>
</tr>
<tr>
<td>6</td>
<td>4,800</td>
<td>.7</td>
<td>No</td>
<td>6,857</td>
</tr>
<tr>
<td>7</td>
<td>4,800</td>
<td>.7</td>
<td>Yes</td>
<td>12,000</td>
</tr>
<tr>
<td>8</td>
<td>4,800</td>
<td>.8</td>
<td>Yes</td>
<td>8,000</td>
</tr>
<tr>
<td>9</td>
<td>5,600</td>
<td>.5</td>
<td>No</td>
<td>11,200</td>
</tr>
<tr>
<td>10</td>
<td>5,600</td>
<td>.7</td>
<td>No</td>
<td>8,000</td>
</tr>
<tr>
<td>11</td>
<td>5,600</td>
<td>.8</td>
<td>Yes</td>
<td>10,360</td>
</tr>
</tbody>
</table>

Source: Ashenfelter and Plant (1990), p. 403
NIT Experiments: Findings

See Ashenfelter and Plant JHR’ 90 for non-parametric evidence. More parametric evidence in earlier work. Key results:

1) Significant labor supply response but small overall

2) Implied earnings elasticity for males around 0.1

3) Implied earnings elasticity for women around 0.5

4) Academic literature not careful to decompose response along intensive and extensive margin

5) Response of women is concentrated along the extensive margin (can only be seen in official govt. report)

6) Earnings of treated women who were working before the experiment did not change much
From true experiment to “natural experiments”

True experiments are costly to implement and hence rare

However, real economic world (nature) provides variation that can be exploited to estimate behavioral responses ⇒ “Natural Experiments”

Natural experiments sometimes come very close to true experiments: Imbens, Rubin, Sacerdote AER ’01 did a survey of lottery winners and non-winners matched to Social Security administrative data to estimate income effects

Lottery generates random assignment conditional on playing

Find significant but relatively small income effects: \( \eta = w \partial l / \partial y \) between -0.05 and -0.10

Identification threat: differential response-rate among groups
Figure 2. Proportion with Positive Earnings for Nonwinners, Winners, and Big Winners

Note: Solid line = nonwinners; dashed line = winners; dotted line = big winners.

On average the individuals in our basic sample won yearly prizes of $26,000 (averaged over the $55,000 for winners and zero for nonwinners). Typically they won 10 years prior to completing our survey in 1996, implying they are on average halfway through their 20 years of lottery payments when they responded in 1996. We asked all individuals how many tickets they bought in a typical week in the year they won the lottery. As expected, the number of tickets bought is considerably higher for winners than for nonwinners. On average, the individuals in our basic sample are 50 years old at the time of winning, which, for the average person was in 1986; 35 percent of the sample was over 55 and 15 percent was over 65 years old at the time of winning; 63 percent of the sample was male. The average number of years of schooling, calculated as years of high school plus years of college plus 8, is equal to 13.7; 64 percent claimed at least one year of college.

We observe, for each individual in the basic sample, Social Security earnings for six years preceding the time of winning the lottery, for the year they won (year zero), and for six years following winning. Average earnings, in terms of 1986 dollars, rise over the pre-winning period from $13,930 to $16,330, and then decline back to $13,290 over the post-winning period. For those with positive Social Security earnings, average earnings rise over the entire 13-year period from $20,180 to $24,300. Participation rates, as measured by positive Social Security earnings, gradually decline over the 13 years, starting at around 70 percent before going down to 56 percent. Figures 1 and 2 present graphs for average earnings and the proportion of individuals with positive earnings for the three groups, nonwinners, winners, and big winners. One can see a modest decline in earnings and proportion of individuals with positive earnings for the full winner sample compared to the nonwinners after winning the lottery, and a sharp and much larger decline for big winners at the time of winning. A simple difference-in-differences type estimate of the marginal propensity to earn out of unearned income (mpe) can be based on the ratio of the difference in the average change in earnings before and after winning the lottery for two groups and the difference in the average prize for the same two groups.

For the winners, the difference in average earnings over the six post-lottery years and the six pre-lottery years is -$1,877 and for the nonwinners the average change is $448. Given a difference in average prize of $55,000 for the winner/nonwinners comparison, the estimated mpe is \((-1,877 - 448)/(55,000 - 0) = -0.042\) (SE 0.016). For the big-winners/small-winners comparison, this estimate is -0.059 (SE 0.018). In Section IV we report estimates for this quantity using more sophisticated analyses.

On average the value of all cars was $18,200. For housing the average value was $166,300, with an average mortgage of $44,200. We aggregated the responses to financial wealth into two categories. The first concerns retirement

Note: Solid line = nonwinners; dashed line = winners; dotted line = big winners.

Source: Imbens et al. (2001), p. 783
Difference-in-Difference (DD) methodology

Two groups: Treatment group (T) which faces a change [lottery winners] and control group (C) which does not [non winners]

Compare the evolution of T group (before and after change) to the evolution of the C group (before and after change)

DD identifies the treatment effect if the parallel trend assumption holds:

Absent the change, T and C would have evolved in parallel

DD most convincing when groups are very similar to start with

Should always test DD using data from more periods and plot the two time series to check parallel trend assumption
Cesarini et al. (2015) use Swedish population wide administrative data with more compelling setting: (1) bank accounts with random prizes (PLS), (2) monthly lottery subscription (Kombi), and (3) TV show participants (Triss)

**Key results:**

1) Effects on both extensive and intensive labor supply margin, time persistent

2) Significant but relatively small income effects: \( \eta = w \partial l / \partial y \) around -0.10

3) Effects on spouse but not as large as on winner

⇒ Rejects the **unitary** model of household labor supply:

\[
\max u(c_1, c_2, l_1, l_2) \quad \text{st} \quad c_1 + c_2 \leq w_1 l_1 + w_2 l_2 + R
\]
Table 1. Distribution of Prizes

<table>
<thead>
<tr>
<th></th>
<th>Pooled Sample</th>
<th>Individual Lottery Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>Share</td>
</tr>
<tr>
<td>0 to 1K SEK</td>
<td>25,172</td>
<td>10.0%</td>
</tr>
<tr>
<td>1K to 10K SEK</td>
<td>204,626</td>
<td>81.3%</td>
</tr>
<tr>
<td>10K to 100K SEK</td>
<td>16,429</td>
<td>6.5%</td>
</tr>
<tr>
<td>100K to 500K SEK</td>
<td>3,685</td>
<td>1.5%</td>
</tr>
<tr>
<td>500K to 1M SEK</td>
<td>355</td>
<td>0.1%</td>
</tr>
<tr>
<td>≥1M SEK</td>
<td>1,481</td>
<td>0.6%</td>
</tr>
<tr>
<td>TOTAL</td>
<td>251,748</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table reports the distribution of lottery prizes for the pooled sample and the four lottery subsamples.

Cesarini, Lindqvist, Notowidigdo, Östling NBER WP 2015
Figure 1: Effect of Wealth on Individual Gross Labor Earnings

Notes: This figure reports estimates obtained from equation (2) estimated in the pooled lottery sample with gross labor earnings as the dependent variable. A coefficient of 1.00 corresponds to an increase in annual labor earnings of 1 SEK for each 100 SEK won. Each year corresponds to a separate regression and the dashed lines show 95% confidence intervals.

Cesarini, Lindqvist, Notowidigdo, Östling NBER WP 2015
Figure 5: Effect of Wealth on Gross Labor Earnings of Winners and Spouses

Notes: This figure reports estimates obtained from equation (2) estimated separately for winners, their spouses, and the household. The dependent variable is gross labor earnings. Each year corresponds to a separate regression.

Cesarini, Lindqvist, Notowidigdo, Östling NBER WP 2015
Married Women Elasticities: Blau and Kahn ’07

1) Identify elasticities from 1980-2000 using grouping instrument

a) Define cells (year × age × education) and compute mean wages

b) Instrument for actual wage with mean wage in cell

2) Identify purely from group-level variation, which is less contaminated by individual endogenous choice

3) Results: (a) total hours elasticity for married women (including intensive + extensive margin) shrank from 0.4 in 1980 to 0.2 in early 2000s, (b) effect of husband earnings ↓ overtime

4) Interpretation: elasticities shrink as women become more attached to the labor force
Summary of Static Labor Supply Literature (SKIP)

1) Small elasticities for prime-age males

Probably institutional restrictions, need for at least one income, etc. prevent a short-run response

2) Larger responses for workers who are less attached to labor force: Married women, low income earners, retirees

3) Responses driven primarily by extensive margin

a) Extensive margin (participation) elasticity around 0.2-0.5

b) Intensive margin (hours) elasticity smaller
Responses to Low-Income Transfer Programs

1) Particular interest in treatment of low incomes in a progressive tax system: are they responsive to incentives?

2) Complicated set of transfer programs in US

a) In-kind: food stamps, Medicaid, public housing, job training, education subsidies

b) Cash: TANF, EITC, SSI

3) See Gruber undergrad textbook for details on institutions
1996 US Welfare Reform

1) Reform modified AFDC cash welfare program to provide more incentives to work (renamed TANF)
   a) Requiring recipients to go to job training or work
   b) Limiting the duration for which families able to receive welfare
   c) Reducing phase-out rate of benefits


4) EITC also expanded during this period: general shift from welfare to “workfare”

Did welfare reform and EITC increase labor supply?
Figure 1. Number of Families Receiving AFDC/TANF Cash Assistance, 1959-2013

Source: Falk (2016)

(Families in millions)

Source: Congressional Research Service (CRS), based on data from the U.S. Department of Health and Human Services (HHS).

Notes: Shaded areas represent recessionary periods. Families receiving TANF cash assistance since October 1, 1999, include families receiving cash assistance from separate state programs (SSPs) with expenditures countable toward the TANF maintenance of effort requirement (MOE).
The landscape providing assistance to poor families with children has changed substantially.

![Graph showing changes in assistance programs over time.](image-url)
Randomized welfare experiment: 
SSP Welfare Demonstration in Canada

Canadian Self Sufficiency Project (SSP): randomized experiment that gave welfare recipients an earnings subsidy for 36 months in 1990s (but need to start working by month 12 to get it)

3 year temporary participation tax rate cut from average rate of 74.3% to 16.7% [get to keep 83 cents for each $ earned instead of 26 cents]

Card and Hyslop (EMA 2005) provide classic analysis. Two results:

1) Strong effect on employment rate during experiment (peaks at 14 points)

2) Effect quickly vanishes when the subsidy stops after 36 months (entirely gone by month 52)
and control groups. Unfortunately, these data have some critical limitations relative to the administratively based Income Assistance data. Most importantly, they are only available for 52 months after random assignment. Since some program group members were still receiving subsidy payments as late as month 52, this time window is too short to assess the long-run effects of the program. Indeed, looking at Figure 1a, there is still an impact on IA participation in month 52 that does not fully dissipate until month 69. Second, because of nonresponses and refusals, labor market information is only available for 85% of the experimental sample (4,757 people). Third, there appear to be relatively large recall errors and seam biases in the earnings and wage data.

Nevertheless, the labor market outcomes provide a valuable complement to the administratively based welfare participation data. Figure 3 shows the average monthly employment rates of the program and control groups, along with the associated experimental impacts. After random assignment the employment rate of the control group shows a steady increase.
Earned Income Tax Credit (EITC) program

See Kleven (2018) provides comprehensive ex-post re-analysis using women aged 20-50 and CPS data

1) EITC started small in the 1970s but was expanded in 1986-88, 1994-96, 2008-09: today, largest means-tested cash transfer program [$70bn in 2016, 30m families recipients]

2) Eligibility: families with kids and low earnings.

3) Refundable Tax credit: administered as annual tax refund received in Feb-April, year $t+1$ (for earnings in year $t$)

4) EITC has flat pyramid structure with phase-in (negative MTR), plateau, (0 MTR), and phase-out (positive MTR)

5) States have added EITC components to their income taxes [in general a percentage of the Fed EITC, great source of natural experiments, understudied bc CPS too small]
EITC Schedule in 2017

- **0 children**
  - Annual Credit (USD): 0
  - Earnings (USD): 0-6000

- **1 child**
  - Annual Credit (USD): 1000-6000
  - Earnings (USD): 0-6000

- **2 children**
  - Annual Credit (USD): 2000-6000
  - Earnings (USD): 0-6000

- **3+ children**
  - Annual Credit (USD): 4000-6000
  - Earnings (USD): 0-6000
EITC Maximum Credit Over Time

Source: Kleven (2018)
Labor Force Participation of Single Women
With and Without Children

Source: Kleven (2018)
50 years of relative stability, apart from these 5 years

Source: Kleven (2018)
Labor Force Participation of Single Women
With and Without Children

50 years of relative stability, apart from these 5 years

14.5pp 14pp

Unemployment Rate

With Children Without Children

Annual Employment Low Education

Source: Kleven (2018)
Labor Force Participation of Single Women
With and Without Children

Source: Kleven (2018)
Labor Force Participation of Single Women
With and Without Children

Source: Kleven (2018)
Labor Force Participation of Single Women
With and Without Children

Source: Kleven (2018)
Labor Force Participation of Single Women
With and Without Children

Source: Kleven (2018)
Labor Force Participation of Single Women
By Number of Children

Source: Kleven (2018)
Labor Force Participation of Single Women
By Number of Children

Tax Reduction Act of 1975 TRA86 OBRA90 OBRA93 ARRA

Much larger increase by those with 3+ kids

Source: Kleven (2018)
Labor Force Participation of Single Women
By Number of Children

Source: Kleven (2018)
Welfare Reform and EITC Expansion: Labor supply

Incredible increase in labor force participation of single mothers during the 1990s when welfare reform and EITC expansion happened

Unlikely that the EITC can explain it because other EITC changes haven’t generated such large effects

Sociological evidence shows that welfare reform “scared” single mothers into working

Single moms in the US were suddenly expected to work

Kleven (2018): Maybe a unique combination of EITC reform, welfare reform, economic upturn, and changing social norms lead to this shift
**Theoretical Behavioral Responses to the EITC**

**Extensive margin:** positive effect on Labor Force Participation as EITC makes work more attractive

**Intensive margin:** earnings conditional on working, mixed effects

1) Phase in: (a) Substitution effect: work more due to wage subsidy, (b) Income effect: work less ⇒ Net effect: ambiguous; probably work more

2) Plateau: Pure income effect (no change in net wage) ⇒ Net effect: work less

3) Phase out: (a) Substitution effect: work less, (b) Income effect: also work less ⇒ Net effect: work less

Should expect bunching at the EITC kink points
Bunching at Kinks (Saez AEJ-EP’10)

Key prediction of standard labor supply model: individuals should bunch at (convex) kink points of the budget set

1) The only non-parametric source of identification for intensive elasticity in a single cross-section of earnings is amount of bunching at kinks creating by tax/transfer system

2) Saez ’10 develops method of using bunching at kinks to estimate the compensated income elasticity

Formula for elasticity: \[ \varepsilon^c = \frac{dz/z^*}{dt/(1-t)} = \text{excess mass at kink / change in NTR} \]

⇒ Amount of bunching proportional to compensated elasticity

Blomquist-Newey 2017: Bunching method requires making assumptions on counterfactual density (but testable using tax changes see Londono-Avila ’18 below)
elasticity would no longer be a pure compensated elasticity, but a mix of the compensated elasticity and the uncompensated elasticity. Four points should be noted.

First, the larger the behavioral elasticity, the more bunching we should expect.

Unsurprisingly, if there are no behavioral responses to marginal tax rates, there

Panel A. Indifference curves and bunching

Panel B. Density distributions and bunching

Before tax income \( z \)

Before tax income \( z \)

After-tax income \( c = z - T(z) \)

Individual L indifference curve

Individual H indifference curves

Individual L chooses \( z^* \) before and after reform

Individual H chooses \( z^* + dz^* \) before and \( z^* \) after reform

\[
\frac{dz^*}{z^*} = e \frac{dt}{1 - t}
\]

with \( e \) compensated elasticity

Source: Saez (2010), p. 184
Elasticity would no longer be a pure compensated elasticity, but a mix of the compensated elasticity and the uncompensated elasticity. Four points should be noted.

First, the larger the behavioral elasticity, the more bunching we should expect. Unsurprisingly, if there are no behavioral responses to marginal tax rates, there

Panel A: Indifference curves and bunching

Before tax income \( z \)

\[ \text{slope} = 1 - t \]

\[ z^* \]

\[ z^* + dz^* \]

Individual L chooses \( z^* \) before and after reform

Individual H chooses \( z^* + dz^* \) before and after reform

\( dz^*/z^* = e dt/(1-t) \) with compensated elasticity

Panel B: Density distributions and bunching

Pre-reform incomes between \( z^* \) and \( z^* + dz^* \) bunch at \( z^* \) after reform

After reform density

Before reform density

Before tax income \( z \)

Source: Saez (2010), p. 184
Before reform: linear tax rate $t_0$, density $h_0(z)$

After reform: tax rate $t_0$ below $z^*$
Tax rate $t_1$ above $z^*$ ($t_1 > t_0$), density $h(z)$

$B = H^*- (H_-^* + H_+^*) = \text{excess bunching}$

Figure 2. Estimating Excess Bunching Using Empirical Densities
Bunching at Kinks (Saez AEJ-EP’10)

1) Uses individual tax return micro data (IRS public use files) from 1960 to 2004

2) Advantage of dataset over survey data: very little measurement error

3) Finds bunching around:
   a) First kink point of the Earned Income Tax Credit (EITC), especially for self-employed
   b) At threshold of the first tax bracket where tax liability starts, especially in the 1960s when this point was very stable

4) However, no bunching observed around all other kink points
EITC Amount as a Function of Earnings

Source: Federal Govt
Two elements are worth noting in Figure 3. First, there is a clear clustering of tax filers around the first kink point of the EITC. In both panels, the density is maximum exactly at the first kink point. The fact that the location of the first kink point differs between EITC recipients with one child, versus those with two or more children, constitutes strong evidence that the clustering is driven by behavioral responses to the EITC as predicted by the standard model. Second, however, we cannot discern any

Figure 3. Earnings Density Distributions and the EITC

Notes:
The figure displays the histogram of earnings (by $500 bins) for tax filers with one dependent child (panel A) and tax filers with two or more dependent children (panel B). The histogram includes all years 1995–2004 and inflates earnings to 2008 dollars using the IRS inflation parameters (so that the EITC kinks are aligned for all years).

Earnings are defined as wages and salaries plus self-employment income (net of one-half of the self-employed payroll tax). The EITC schedule is depicted in dashed line and the three kinks are depicted with vertical lines. Panel A is based on 57,692 observations (representing 116 million tax returns), and panel B on 67,038 observations (representing 115 million returns).

Source: Saez (2010), p. 191
indexes earnings to 2008 using the IRS inflation parameters, so that the EITC kinks are perfectly aligned for all years.

Two elements are worth noting in Figure 3. First, there is a clear clustering of tax filers around the first kink point of the EITC. In both panels, the density is maximum exactly at the first kink point. The fact that the location of the first kink point differs between EITC recipients with one child, versus those with two or more children, constitutes strong evidence that the clustering is driven by behavioral responses to the EITC as predicted by the standard model. Second, however, we cannot discern any

Panel A. One child

Panel B. Two children or more

Figure 3. Earnings Density Distributions and the EITC

Notes: The figure displays the histogram of earnings (by $500 bins) for tax filers with one dependent child (panel A) and tax filers with two or more dependent children (panel B). The histogram includes all years 1995–2004 and inflates earnings to 2008 dollars using the IRS inflation parameters (so that the EITC kinks are aligned for all years). Earnings are defined as wages and salaries plus self-employment income (net of one-half of the self-employed payroll tax). The EITC schedule is depicted in dashed line and the three kinks are depicted with vertical lines. Panel A is based on 57,692 observations (representing 116 million tax returns), and panel B on 67,038 observations (representing 115 million returns).

Source: Saez (2010), p. 191
systematic clustering around the second kink point of the EITC. Similarly, we cannot discern any gap in the distribution of earnings around the concave kink point where the EITC is completely phased-out. This differential response to the first kink point, versus the other kink points, is surprising in light of the standard model predicting that any convex (concave) kink should produce bunching (gap) in the distribution of earnings.

In Figure 4, we break down the sample of earners into those with nonzero self-employment income versus those zero self-employment income. Panel A reports the density for tax filers with one dependent child and panel B for tax filers with two or more dependent children. The charts include all years 1995–2004. The bandwidth is $400 in all kernel density estimations. The fraction self-employed in 16.1 percent and 20.5 percent in the population depicted on panels A and B (in the data sample, the unweighted fraction self-employed is 32 percent and 40 percent). We display in dotted vertical lines around the first kink point the three bands used for the elasticity estimation with \( \delta = 1,500 \).
systematic clustering around the second kink point of the EITC. Similarly, we cannot discern any gap in the distribution of earnings around the concave kink point where the EITC is completely phased-out. This differential response to the first kink point, versus the other kink points, is surprising in light of the standard model predicting that any convex (concave) kink should produce bunching (gap) in the distribution of earnings.

In Figure 4, we break down the sample of earners into those with nonzero self-employment income versus those zero self-employment income. Panel A reports the density for tax filers with one dependent child and panel B for tax filers with two or more dependent children. The charts include all years 1995–2004. The bandwidth is $400 in all kernel density estimations. The fraction self-employed in 16.1 percent and 20.5 percent in the population depicted on panels A and B (in the data sample, the unweighted fraction self-employed is 32 percent and 40 percent). We display in dotted vertical lines around the first kink point the three bands used for the elasticity estimation with $\delta = \$1,500.$

Source: Saez (2010), p. 192
Why not more bunching at kinks?

1) True intensive elasticity of response may be small

2) Randomness in income generation process: Saez (1999) shows that year-to-year income variation too small to erase bunching if elasticity is large

3) Frictions: Adjustment costs and institutional constraints (Chetty, Friedman, Olsen, and Pistaferri QJE’11)

4) Information and salience
EITC Behavioral Studies

Evidence of response along extensive margin but 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 (Liebman-Zeckhauser ’04, Rees-Jones and Taubinsky ’16)
Chetty, Friedman, Saez AER’13 EITC heterogeneity

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 eff‘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

Source: Chetty, Friedman, and Saez NBER'12
Earnings Distribution in the Year of 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
Value of Administrative data

Key advantages of admin data (in most advanced countries such as Scandinavia):

1) Size (often full population available)

2) Longitudinal structure (can follow individual across years)

3) Ability to match wide variety of data (tax records, earnings records, family records, health records, education records)

US is lagging behind in terms of admin data access [hard to match across agencies]

Private sector also generates valuable **big data** (Google, Credit Bureaus, personnel/health data from large companies)
Bunching at Notches: Kleven and Waseem ’13

Taxes and transfers sometimes also generate notches (=discontinuities) in the budget set.

Such discontinuities should create bunching (and gaps) in the resulting distributions.

Kleven and Waseem QJE’13 pioneered tax notch analysis in the case of the Pakistani income tax where average tax rate jumps.

⇒ Bunching below the notch and gap in density just above the notch.

Recently Londono-Velez and Avila (2018) use notch analysis to study wealth tax in Columbia.

They show clean prior-year counterfactual overcoming the Blomquist-Newey ’17 critique.
 FIGURE 1  
Effect of Notch on Taxpayer Behavior

Panel A: Bunching at the Notch
FIGURE 2
Effect of Notch on Density Distribution

Panel A: Theoretical Density Distributions

Source: Kleven and Waseem '11
Figure 1: The Personal Wealth Tax Schedule in Colombia

(a) Wealth Tax Liability as a Function of Reported Net Wealth (FY 2010)

Wealth tax $T(W_r)$ (million COP)

<table>
<thead>
<tr>
<th>Reported wealth $W_r$ (billion COP)</th>
<th>Wealth percentile</th>
<th>Wealth tax 3% (million COP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0.5] P99.88</td>
<td>0%</td>
<td>10</td>
</tr>
<tr>
<td>[1.0] P99.96</td>
<td>1%</td>
<td>28</td>
</tr>
<tr>
<td>[1.6] P99.98</td>
<td>1.4%</td>
<td>300</td>
</tr>
<tr>
<td>[2.6] P99.99</td>
<td>6%</td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- These figures depict the personal wealth tax schedule for Colombia.
- Panel (a) plots wealth tax liability by reported wealth $W_r$ in FY 2010. Each bracket of $W_r$ is associated with a fixed average tax rate on taxable net wealth. As a result, wealth tax liability $T(W_r)$ jumps discretely at the notch points. That year, the wealth tax brackets affected the top 0.12%, top 0.04%, top 0.02%, and top 0.01%, respectively.

Source: Table A.1
Source Londono 2018
(b) Evolution of Statutory Annual Wealth Tax Rates by Bracket Cutoff

Tax rate $\tau$

Bracket cutoff:
- **1 billion pesos**
- **2 billion pesos**
- **3 billion pesos**
- **5 billion pesos**

Source: Table A.1
Figure 2: Distribution of Reported Net Worth in 2009 (Before Reform) and 2010 (After Reform)

Notes: This figure overlays the distribution of tax filers by reported net wealth before and after a reform introduced two wealth tax notches at 1 and 2 billion pesos (red vertical lines), as depicted in Figure 1. These notches imply that wealth tax liability jumps discontinuously, as illustrated in Figure 1. The figure shows that the distribution of individuals is smooth in the absence of wealth tax notches (2009). The two notches result in the immediate emergence of excess mass below the notch points, and corresponding missing mass just above them (2010). This
Bunching at notches: elasticity estimation

With optimization frictions (lack of information, costs of adjustment), a fraction of individuals fail to respond to notch

Kleven-Waseem use empirical density in the theoretical gap area to measure the fraction of unresponsive individuals

This allows them to back up the frictionless elasticity (i.e. the elasticity among responsive individuals)

The frictionless elasticity is much higher than the reduced form elasticity but remains still relatively modest
Many Recent Bunching Studies

Bunching method applied to many settings with nonlinear budgets with convex kink points or notches (Kleven ’16 survey):

- Individual tax (Bastani-Selin ’14 Sweden, Mortenson-Whitten ’16 US)
- Payroll tax (Tazhidinova ’15 on UK)
- Corporate tax (Devereux-Liu-Loretz ’14, Bachas-Soto ’17)
- Wealth tax (Seim ’17, Jakobsen et al. ’17, Londono-Velez and Avila ’18)
- Health spending (Einav-Finkelstein-Schrimpf ’13 on Medicare Part D)
- Retirement savings (401(k) matches)
- Retirement age (Brown ’13 on California Teachers)
- Housing transactions (Best and Kleven, 2017)

General findings:

(1) clear bunching when information is salient and outcome easily manipulable

(2) bunching is almost always small relative to conventional elasticity estimates
Responses to Corporate Tax Notches: Bachas-Soto ’18

Bachas and Soto ’18 exploit the notched Costa Rica corporate tax system to estimate compellingly the effects of the corporate tax rate on reported profits.

Corporate tax applies to profits = revenue minus costs

But tax rate depends on size of revenue with 3 rates: 10%, 20%, 30%

1) Firms bunch at the notches to benefit from the lower rates

2) Most importantly: clear evidence that profit rates (profits/revenue) is strongly affected by the corporate tax rate
Figure 1 shows the design of the corporate income tax in Costa Rica, as discussed in section 2.1. Firms face increasing average tax rates on their profits (revenue minus cost) as a function of their revenue. When revenue exceeds the first threshold, the average tax rate jumps from 10% to 20% and from 20% to 30% past the second threshold. Thresholds are adjusted yearly for inflation.
Figure 3: Firm Density and Average Profit Margin

Panel A: Firm Density

Panel B: Profit Margin

Figure 3 presents the key patterns of the corporate tax data, discussed in Section 3.1. The figure Pulls together data from years 2008 to 2014. Panel A shows the density of firms by revenue. Panel B displays the average profit margin by revenue. Profit margin is defined as profits over revenue. The size of the revenue bins is 575,000 CRC.
Macro Long-Run Evidence

1) Macroeconomists also estimate elasticities by examining long-term trends/cross-country comparisons

2) Identification more questionable but estimates perhaps more relevant to long-run policy questions of interest

3) Use aggregate hours data and aggregate measures of taxes (average tax rates)

4) Highly influential in calibration of macroeconomic models
Trend-based Estimates and Macro Evidence

**Long-Run:** US real wage rates multiplied by about 5 from 1900 to present due to economic growth.

Aged 25-54 male hours have fallen 25% and then stabilized (Ramey and Francis AEJ-macro ’09)

⇒ Uncompensated hours of work elasticity is small ($<.1$)

However, taxes are rebated as transfers so can still have labor supply effects if large compensated elasticity/income effects

Alternative plausible story: utility depends on relative consumption ⇒ Earnings $10,000$ is low today but would have been very good in 1900 (reference point labor supply theory)
Figure 2. Average Weekly Hours Worked per Person, by Age Group

Source: Authors' estimates, based on information from Kendrick (1961, 1973), the census, and the CPS.

Ramey and Francis AEJ'09

C. Females
Long-run cross-country panel: Prescott 2004

Uses data on hours worked by country in 1970 and 1995 for 7 OECD countries [total hours/people age 15-64]

Technique to identify elasticity: calibration of GE model

Rough intuition: posit a labor supply model, e.g.

\[ u(c, l) = c - \frac{l^{1+1/\varepsilon}}{1 + 1/\varepsilon} \]

Finds that elasticity of \( \varepsilon = 1.2 \) best matches time series and cross-sectional patterns

Note that this is analogous to a regression without controls for other variables

Results verified in subsequent calibrations by Ohanina-Raffo-Rogerson JME’08 and others using more data
<table>
<thead>
<tr>
<th>Period</th>
<th>Country</th>
<th>Labor Supply*</th>
<th>Differences (Predicted Less Actual)</th>
<th>Prediction Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Actual</td>
<td>Predicted</td>
<td></td>
</tr>
<tr>
<td>1993–96</td>
<td>Germany</td>
<td>19.3</td>
<td>19.5</td>
<td>.2</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>17.5</td>
<td>19.5</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>16.5</td>
<td>18.8</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>Canada</td>
<td>22.9</td>
<td>21.3</td>
<td>−1.6</td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
<td>22.8</td>
<td>22.8</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Japan</td>
<td>27.0</td>
<td>29.0</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>United States</td>
<td>25.9</td>
<td>24.6</td>
<td>−1.3</td>
</tr>
<tr>
<td>1970–74</td>
<td>Germany</td>
<td>24.6</td>
<td>24.6</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>France</td>
<td>24.4</td>
<td>25.4</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Italy</td>
<td>19.2</td>
<td>28.3</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td>Canada</td>
<td>22.2</td>
<td>25.6</td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
<td>25.9</td>
<td>24.0</td>
<td>−1.9</td>
</tr>
<tr>
<td></td>
<td>Japan</td>
<td>29.8</td>
<td>35.8</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>United States</td>
<td>23.5</td>
<td>26.4</td>
<td>2.9</td>
</tr>
</tbody>
</table>

*Labor supply is measured in hours worked per person aged 15–64 per week.
Sources: See Appendix.

Reconciling Micro and Macro Estimates

Recent interest in reconciling micro and macro elasticity estimates (see Chetty-Guren-Manoli-Weber '13)

Three potential explanations

a) Statistical Bias: culture differs in countries with higher tax rates [Alesina, Glaeser, Sacerdote 2005, Steinhauer 2018 for Swiss communities by language]

b) Macro-elasticity captures long-term response which could be larger than short-term response (frictions, etc. Chetty '12).

c) Other programs: retirement, education affect labor supply at beginning and end of working life (Blundell-Bozio-Laroque '13) and child care affecting mothers (Kleven JEP’14)
Strong evidence that variation in aggregate hours of work across countries happens among the young and the old: (a) schooling-work margin (b) presence of young children (for women), (c) early retirement

Serious cross-country time series analysis would require to put together a better tax wedge by age groups which includes all those additional govt programs [welfare, retirement, child care]

This has been done quite successfully in the case of retirement by series of books by Gruber and Wise, *Retirement around the world*

⇒ Need to develop a more comprehensive international / time series database of tax wedges by age and family types
There are certain key margins where tax rates impinge on earnings decisions. For many male workers this is at the beginning and at the end of their working lives. These are the schooling-work margins and the early retirement margins. Indeed much of the difference in male employment across OECD countries occurs at these points in the life-cycle.

Male employment by age – US, FR and UK 2005

Source: Blundell (2009), Mirrlees Review
Male Hours by age – US, FR and UK 2005

Source Blundell (2009), Mirrlees Review
Male employment by age – US, FR and UK 1975

Source: Blundell (2009), Mirrlees Review
Female Employment by age – US, FR and UK 2005

Source: Blundell (2009), Mirrlees Review
Female Hours by age – US, FR and UK 2005

Source Blundell (2009), Mirrlees Review
For women earnings are influenced by taxes and benefits not only at these margins but also when there are young children in the family.

For women with younger children it is not usually just an employment decision that is important it is also whether to work part-time or full-time.

Often the employment margin is referred to as the extensive margin of work and the part-time or hours of work decisions more generally as the intensive margin.

---

Source Blundell (2009), Mirrlees Review
Long-term effects: Evidence from the Israeli Kibbutz

Abramitzky ’18 book based on series of academic papers

Kibbutz are egalitarian and socialist communities in Israel, thrived for almost a century within a more capitalist society

1) Social sanctions on shirkers effective in small communities with limited privacy

2) Deal with brain drain exit using communal property as a bond

3) Deal with adverse selection in entry with screening and trial period

4) Perfect sharing in Kibbutz has negative effects on high school students performance but effect is small in magnitude (concentrated among kids with low education parents)
Long-term effects: Evidence from the Israeli Kibbutz

Abramitzky-Lavy ECMA’14 show that high school students study harder once their kibbutz shifts away from equal sharing.

Uses a DD strategy: pre-post reform and comparing reform Kibbutz to non-reform Kibbutz. Finds that

1) Students are 3% points more likely to graduate

2) Students are 6% points more likely to achieve a matriculation certificate that meets university entrance requirements

3) Students get an average of 3.6 more points in their exams

Effect is driven by students whose parents have low schooling; larger for males; stronger in kibbutz that reformed to greater degree.
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]
We note the judge leniency measure is calculated from all cases the judge has ever handled, not just the cases in our estimation sample. On average, each judge has handled a total of 380 cases. The mean of the leniency variable is 0.15 with a standard deviation of 0.06. The histogram reveals a wide spread in judge leniency, with approximately 22% of cases allowed by a judge at the 90th percentile compared to approximately 9% at the 10th percentile.

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


Abramitzky, Ran and Victor Lavy, 2014 “How Responsive is Investment in Schooling to Changes in Redistributive Policies and in Returns?”, Econometrica, 82(4), 1241-1272 (web)


Card, David, Raj Chetty, Martin Feldstein, and Emmanuel Saez “Expanding Access to Administrative Data for Research in the United States,”


eds., Labour Supply and Incentives to Work in Europe, 2005, 44-104. (web)


Heckman, J. and M. Killingsworth “Female Labor Supply: A Survey” Hand- 

Imbens, G.W., D.B. Rubin and B.I. Sacerdote “Estimating the Ef- 
fect of Unearned Income on Labor Earnings, Savings, and Con- 
sumption: Evidence from a Survey of Lottery”, American Economic 

from Denmark NBER working paper No. 24371. (web)

Jones, Damon “Information, Inertia and Public Benefit Participation: Ex- 
perimental Evidence from the Advance EITC and 401(k) Savings“, AEJ: 

Keane, Michael “Labor Supply and Taxes: A Survey?”, Journal of Eco- 
nomic Literature, Vol. 49(4), 2011, 961-1075. (web)

Kleven, Henrik “How Can Scandinavians Tax So Much?”, Journal of Eco- 
nomic Perspectives 28(4), 77-98, 2014 (web)

Kleven, Henrik “Taxation and Labor Force Participation: The EITC 
Reconsidered”, October 2018 slides (web)

Kleven, Henrik “Taxation and Labor Force Participation: The EITC Re- 
considered”, October 2018 slides (web)


Steinhauer, Andreas. 2018 “Working Moms, Childlessness, and Female Identity”, CEPR Discussion Paper No. 12929.[web]