Labor Supply Responses to Taxes and Transfers

131 Undergraduate Public Economics
Emmanuel Saez
UC 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 (same as previous lecture):

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

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

Individual solves

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

FOC: $w \partial u / \partial c + \partial u / \partial l = 0$ defines uncompensated (Marshallian) labor supply function $l^u(w, R)$

Uncompensated elasticity of labor supply: $\varepsilon^u = (w/l) \partial l^u / \partial w$ [% change in hours when net wage $w$ increases by 1%]

Income effect parameter: $\eta = w \partial l / \partial R \leq 0$: $\$ 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: $\varepsilon^c = (w/l) \partial l^c / \partial w > 0$

Slutsky equation: $\partial l / \partial w = \partial l^c / \partial w + l \partial l / \partial R \Rightarrow \varepsilon^u = \varepsilon^c + \eta$
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 (Ordinary Least Square) regression:

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

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

\( R_i \) measures non-labor income [including spousal earnings for married person]

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

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


a) Small effects $\varepsilon^u = 0$, $\eta = -0.1$, $\varepsilon^c = 0.1$ with some variation across estimates

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)
ISSUE WITH OLS REGRESSION:

$w_i$ correlated with taste for work $\epsilon_i$

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

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 regression

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** $\Rightarrow$ Tax changes provide more compelling identification
Negative Income Tax (NIT) Experiments

1) Best identification method: exogenously increase the tax rate / non-labor income with a **randomized experiment**

2) NIT experiments 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) Lump-sum transfers $G$ combined with a steep phaseout rate $\tau$ (50%-80%) [based on family earnings] for 3 or 5 years.


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
Negative Income Tax Experiment

$c = z - T(z)$

NIT Treatment: Transfer $G$ phased-out with earnings $z$ at tax rate $\tau$

slope = $1 - \tau$

Control group: slope = 1

45°
Negative Income Tax Experiment

\[ c = z - T(z) \]

NIT Treatment
Negative income and substitution effects on \( z \)

slop=1-\( \tau \)

Control group: slope=1

45°
NIT Experiments: Findings

1) Significant labor supply response but small overall

2) Implied earnings elasticity for males around 0.1

3) Implied earnings elasticity for married women around 0.5

4) Response of married women is concentrated along the extensive margin (dropping out of work)

5) Earnings of treatment group bounce back after experiment ends
From true experiment to “natural experiments”:
Estimating income effects with lottery winnings

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 small income effects: \( \eta = w \partial l / \partial R \) between -0.05 and -0.10: $1 in lottery reduces earnings by 5-10¢.

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

Figure 1. Average Earnings for Nonwinners, Winners, and Big Winners

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

Source: Imbens et al. (2001), p. 783
Labor Supply Substitution Effects: Tax Free Second Jobs in Germany

In 2003, Germany made secondary jobs (paying less than 400 Euros/month) tax free: amounts to a 20-60% subsidy on second job earnings: substitution labor supply effect

Tazhitdinova '22 uses social security admin monthly earnings data

Fraction of population holding second jobs increased sharply (from 2.5% to 6-7%) with bigger response overtime

Finds no offsetting effect on primary earnings ⇒ People did work more

Likely happened because employers willing to create lots of mini-jobs to accommodate supply
Figure 4: Secondary Job Holding Rates by Secondary Earnings Level
Source: Tazhïtïñova (2019)

(a) same axis
Responses to Low-Income Transfer Programs

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

2) Complicated set of transfer programs in US

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

b) Cash: Temporary Aid to Need Families (TANF) and Earned Income Tax Credit for families with kids, Supplemental Security Income (SSI) for aged and disabled

US government (fed+state and local) spent 6% of national income in 2019 on income-tested programs

a) 60% is health care (Medicaid)

b) Cash goes almost only to families with kids
Means-tested Transfers in the US, 1960-2019

1996 US Welfare Reform (PRWORA)

1) Reform modified AFDC (Aid for Families with Dependent Children) cash welfare program to provide more incentives to work (renamed Temporary Aid to Needy Families, TANF)

   a) Requiring recipients to go to job training or work

   b) Limiting the duration of benefits (5 year max lifetime)

   c) Reducing phase-out rate of benefits

2) States got welfare waivers from Federal government to experiment during 1992-1996 before Federal welfare reform

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

Did welfare reform and EITC increase labor supply?
**Figure 1: Long-Run Evolution of EITC and Cash Welfare**

![Graph showing the evolution of EITC and cash welfare recipients over time.](image)

- **EITC Recipients** (Red line)
- **AFDC/TANF Recipients** (Blue line)

Source: Internal Revenue Service (EITC) and Department of Health and Human Services (AFDC/TANF).

Notes: The red series show the annual number of federal EITC recipients between 1966-2016. The blue series show the average monthly number of Aid to Families with Dependent Children (AFDC) recipients between 1966-1996, and the average monthly number of Temporary Assistance for Needy Families (TANF) recipients between 1997-2016.
SNAP Tracks Changes in Share of Population Near or Below the Poverty Line

- Share of individuals with income at or below 130% of poverty line
- Share of individuals participating in SNAP

Note: Poverty estimates are annual estimates. SNAP shares of resident population are calendar year averages.
Sources: U.S. Census Bureau, U.S. Department of Agriculture
Randomized welfare experiment: 
Canadian Self Sufficiency Project

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 ’05 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).18 Third, there appear to be relatively large recall errors and seam biases in the earnings and wage data.19

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, while the program group shows a slight increase followed by a decrease, and then a gradual increase. The difference between the two groups is also shown, with a peak in the first 12 months and then a gradual decrease.

---

**Figure 3.**—Monthly employment rates.

Source: Card and Hyslop, 2005, p. 1734

18The distribution of response patterns to the 18-, 36-, and 54-month surveys is fairly similar for the program and control groups (chi-squared statistic = 11.4 with 7 degrees of freedom, p-value = 0.12). However, a slightly larger fraction of the program group have complete labor market data for 52 months—85.4% versus 84.0% for the controls. Moreover, the difference in mean IA participation between the treatment and control groups in month 52 is a little different in the overall sample (2.5%) than in the subset with complete labor market histories (3.3%).

19Each of the three post-random-assignment surveys asked people about their labor market outcomes in the 18 months since the previous survey. Many people report constant earnings over the recall period, leading to a pattern of measured pay increases that are concentrated at the seams, rather than occurring more smoothly over the recall period.
Earned Income Tax Credit (EITC) program

1) EITC started small in the 1970s but was expanded in 1986-88, 1994-96, 2008-09: today, largest means-tested cash transfer program [$75bn in 2019, 30m families recipients]

2) Eligibility: families with kids and low earnings.

3) Refundable Tax credit: administered through income tax 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) Theoretically, EITC should encourage labor force participation (extensive labor supply margin)

Kleven (2019) who looks at participation of single women (aged 20-50) with kids (treatment) vs without kids (control)
EITC Schedule in 2017
Labor Force Participation of Single Women
With and Without Children

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

50 years of relative stability, apart from these 5 years

Unemployment Rate

50 60 70 80 90 100

Labor Force Participation (%)

With Children Without Children

Annual Employment Low Education

14 / 167

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

50 years of relative stability, apart from these 5 years

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)
Welfare Reform and EITC Expansion: Labor supply

Kleven (2019) looks at participation of single women (aged 20-50) with kids (treatment) vs without kids (control)

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

Unlikely that the EITC can explain it fully 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

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:** EITC makes work more attractive (vs. non-work) ⇒ positive effect on Labor Force Participation

**Intensive margin:** earnings conditional on working;

1) Phase in: (a) Substitution effect: work more due to 40% increase in net wage, (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
EITC and intensive labor supply

\[ c = z - T(z) \]

Budget with EITC
EITC and intensive labor supply

\[ c = z - T(z) \]

Negative income and substitution effects on \( z \)

Negative income effects on \( z \)

Positive substitution effects on \( z \)
EITC and Intensive Labor Supply Response: Bunching at Kinks

1) Basic labor supply theory predicts that we should observe bunching of individuals at the EITC kink points:

Some individuals find it worthwhile to work more when subsidy rate is 40% (2 kids) but not when subsidy rate falls to 0% ⇒ Utility maximizing labor supply is to be exactly at the kink

2) Amount of bunching is proportional to compensated elasticity: if labor supply is inelastic, then kinks in the budget set are irrelevant and do not create bunching

Saez AEJ’10 finds bunching around 1st kink point of EITC but only for the self-employed ⇒ likely due to cheating to maximize tax refund (and not labor supply)
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 would be no bunching.

Panel A. Indifference curves and bunching

Panel B. Density distributions and bunching

Source: Saez (2010), p. 184
B. Density Distributions and Bunching

Before tax income $z$

Pre-reform incomes between $z^*$ and $z^*+dz^*$ bunch at $z^*$ after reform

Before reform density

After reform density

Before tax income $z$
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...
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

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 (and hence whose...). 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
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 (and hence whose earnings density

Panel A. One child

Panel B. Two or more children

Figure 4. Earnings Density and the EITC: Wage Earners versus Self-Employed

Notes:
The figure displays the kernel density of earnings for wage earners (those with no self-employment earnings) and for the self-employed (those with nonzero self employment earnings). 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 per-cent 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
Some 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 increases or decreases with earnings

Such confusion might be good for the government as the EITC induces work along participation margin without discouraging work along intensive margin
Chetty, Friedman, Saez AER’13 EITC information

Use US population wide tax return data since 1996

1) Substantial heterogeneity fraction of EITC recipients bunching (using self-employment) across geographical areas

⇒ Information about EITC varies across areas

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

⇒ Evidence of wage earnings response to EITC along intensive margin

3) Omitted variable test: use birth of first child to test causal effect of EITC on wage earnings
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
Long-term effects of Redistribution: Evidence from the Israeli Kibbutz

Abramitzky (2018) book based on series of academic papers

Kibbutz are egalitarian and socialist voluntary communities in Israel, thrived for almost a century within a 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
Long-term effects of Redistribution: Evidence from the Israeli Kibbutz

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

They use a DD strategy: pre-post reform and comparing reform Kibbutz to non-reform Kibbutz. They find that

1) Students are 3 percentage points more likely to graduate

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

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’14 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 severe, others 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)
Long-term benefits for children of support programs

Traditional economic view: equity vs. efficiency tradeoff

But support programs can also have positive efficiency long-term impacts on children later on (birth weight, health, education, earnings) as shown in series of papers by H. Hoynes

Example: Hoynes, Schanzenbach, Almond AER’16 studies effect of Food Stamp Program rollout across US counties in the 1960s-1970s on health (based on age of children at rollout)

Find large positive impact of food stamps on health for being exposed to food stamps in early childhood (-.4 std dev. of having metabolic syndrome precursor to diabetes)

⇒ Social state support should also be seen as investment
but the Act also stated that no county could run both the FSP and the CDP. Thus, for counties which previously ran a CDP, adoption of the FSP implies termination of the CDP. The political accounts of the time suggest that debates about adopting the FSP pitted powerful agricultural interests (who favored the CDP) against advocates for the poor (who favored the FSP: see MacDonald 1977; Berry 1984). In particular, counties with strong support for farming interests (e.g., Southern or rural counties) may be late adopters of the FSP. On the other hand, counties with strong support for the low-income population (e.g., Northern, urban counties with large poor populations) may adopt FSP earlier in the period. This systematic variation in food stamp adoption could lead to spurious estimates of the program impact if those same county characteristics are associated with differential trends in the outcome variables.

In earlier work (Hoynes and Schanzenbach 2009), we documented that larger counties with a greater fraction of the population that was urban, black, or low income indeed implemented the FSP earlier (i.e., consistent with the historical accounts). Nevertheless, we found that the county characteristics explain very little of the variation in adoption dates. This is consistent with the characterization of funding limits controlling the movement of counties off the waiting list to start up

11 This transition in nutritional assistance would tend to bias downward FSP impact estimates, but we do not think this bias is substantial because of the limited scope of the CDP. The CDP was not available in all counties and recipients often had to travel long distances to pick up the items. Further, the commodities were distributed infrequently and inconsistently, and provided a very narrow set of commodities—the most frequently available were flour, cornmeal, rice, dried milk, peanut butter, and rolled wheat (Citizens' Board of Inquiry 1968). In contrast, food stamp benefits can be used to purchase a wide range of grocery food items.

12 For more detail, see Table 1 in Hoynes and Schanzenbach (2009).

Figure 2. Food Stamp Program Start Date, by County, 1961–1974

Notes: Authors’ tabulations of food stamp administrative data (US Department of Agriculture, various years). The shading corresponds to the county FSP start date, where darker shading indicates later county implementation.
Amendments to the Food Stamp Act, which mandated that all counties offer FSP by 1975.

Figure 1 plots the percent of counties with a FSP from 1960 to 1975. During the pilot phase (1961–1964), FSP coverage increased slowly. Beginning in 1964, program growth accelerated; coverage expanded at a steady pace until all counties were covered in 1974. Furthermore, there was substantial heterogeneity in timing of adoption of the FSP, both within and across states. The map in Figure 2 shades counties according to date of FSP adoption (darker shading denotes a later start-up date). Our basic identification strategy considers the month of FSP adoption for each county to trigger the beginning of the FSP “treatment.”

For our identification strategy to yield causal estimates of the program, it is important to establish that the timing of FSP adoption appears to be exogenous and here we summarize what we have examined in our earlier work (Hoynes and Schanzenbach 2009). Prior to the FSP, some counties provided food aid through the commodity distribution program (CDP) which took surplus food purchased by the Federal government as part of an agricultural price support policy and distributed those goods to the poor. The 1964 Food Stamp Act allowed for counties to voluntarily set up a FSP.

Counties are weighted by their 1970 population. Note this is not the food stamp caseload, but represents the percent of the US population that lived in a county with a FSP. Online Appendix Figure 1 reproduces this figure and adds the county-level coverage rate using the PSID data. The data available in the PSID line up well with the national rollout trends.

Figure 1. Weighted Percent of Counties with Food Stamp Program, 1960–1975

Source: Authors’ tabulations of food stamp administrative data (US Department of Agriculture, various years). Counties are weighted by their 1960 population.

Source: Hoynes, Schanzenbach, and Almond AER'16
given that we find zero impact on economic self-sufficiency for men (Table 6). Note that these are the reverse of a typical event study graph, in that negative “event time” is the case where a person was fully treated (food stamps was in place in their county prior to birth). Further, treatment (exposure to the program) increases as we move from the right (treated in later life) to the left (treated in early life). Finally, as we have said before, once the treatment turns on it does not turn off. While we do not have a strong prediction about the precise shape of the treatment effects, our hypothesis is that the impact of the FSP treatment should decline as age at initial exposure increases. Or to state the reverse, the younger the initial age of exposure the larger the (cumulative) effect of the FSP. If exposure in later childhood does not matter, then the event study coefficients should be flat on the right end of the graph (suggesting no “pretrend”). Eventually, once we hit the point in early childhood when exposure matters, a movement left (toward earlier initial exposure) should reduce the metabolic syndrome index (or increase economic self-sufficiency).

The results in Figure 3 are highly consistent with these predictions and quite encouraging for our research design. They show that the largest effects of the food stamp treatment (in this case a reduction in metabolic syndrome is good and so a

Figure 3. Event Study Estimates of the Impact of FSP Exposure on Metabolic Syndrome Index (High Participation Sample)

Notes: The figure plots coefficients from an event-study analysis. Event time is defined as age when FSP is implemented in the birth county. The models are estimated for the sample of individuals born into families where the head has less than a high school education. Age 10–11 is the omitted year so estimates are relative to that point. See the text for a description of the model.
Crime Reduction Benefits of Welfare Benefits

US has very high incarceration rate .50% of population down from .75% in 2008 peak (but still 5 times more than Europe)

Costs $50K/year per inmate ⇒ Expensive and punitive

Deshpande and Mueller QJE’22 use welfare reform which made it harder for disabled children to keep SSI (supplemental security income for low income aged+disabled) past age 18 ⇒ Can use Regression Discontinuity Design (RDD)

Finds increase in offenses by 20% and incarceration by 60% over next 2 decades [Deshpande AER’16 showed positive but small effect on regular work]

Incarceration effect persists even after 1st stage effect on SSI recipiency has vanished

Extra incarceration costs almost as big as SSI benefits saved
First Stage: Likelihood of Age 18 Medical Review across Cutoff

Figure plots the likelihood of receiving an age 18 medical review and the likelihood of receiving an unfavorable age 18 review (i.e., being removed from SSI at age 18). The sample is SSI children with an 18th birthday within 18 months of the August 22, 1996, cutoff who reside in a county with CJARS coverage. Table I reports point estimates and standard errors.

Source: Deshpande and Mueller-Smith QJE 2023
FIGURE III
Reduced Form: Criminal Justice Outcomes across Cutoff

Importantly, these results mean that the two main CJARS data sources—court charges and corrections—independently show a large and precise increase in criminal justice involvement across the cutoff.11

11. What explains the upward trend in number of charges by birthdate? The answer appears to be a secular increase in criminal charges for younger birth cohorts. Online Appendix Figure B5 shows an upward trend in criminal charges and incarceration for the general population with the same birthdate range.
Social Determinants of Labor Supply (Saez ’21)

Concern that taxes funding social state could discourage work

**Standard econ view:** labor supply $l(w, R)$ coming out of
max $u(c, l)$ st $c = wl + R$ is highly incomplete

**Social determinants of labor supply:**

a) Youth labor is regulated by labor laws/education

b) Old age labor regulated by retirement programs

c) Female market labor driven by norms + child care policy

d) Hours of work regulated by overtime + vacation mandates

Social labor supply with disutility for youth, old, overtime labor
Employment Rates of Men by Age, 2019

Source: Saez AEA-PP’21

Source: OECD database online. Employment to population ratios.

US

France

0%
20%
40%
60%
80%
100%
15-19
20-24
24-29
30-34
35-39
40-44
45-49
50-54
55-59
60-64
65-69
70-74
75-79
80+
Employment Rates of Women by Age, 2019

Source: OECD database online. Employment to population ratios.

Source: Saez AEA-PP’21
Employment Rates of Men and Women, aged 25-54

Source: Saez AEA-PP'21

Source: OECD database online.
Employment Rates of Men and Women, aged 25-54

Source: OECD database online.

Source: Saez AEA-PP'21
US female labor force participation, age 16-64

25% increase in 1943-1945 during WW2 planned economy


Source: Saez AEA-PP’21
CHILD PENALTIES ACROSS COUNTRIES: EVIDENCE AND EXPLANATIONS

The first term on the right-hand side includes extensive margin impacts of children.

In the online Appendix, we present separate results on the ability to keep the zeros in the data.

Large, immediate and persistent drop in earnings sharply after parenthood. Women experience a similarly before parenthood—after adjusting for life cycle and time trends—but diverge differently after. The underlying this approach, compare its results to alternative approaches in the literature, and provide evidence of its ability to identify the causal effect of parenthood.

Our main outcome variable is gross labor earnings, excluding taxes or transfers, specified in levels. We convert the estimated coefficients to define the child penalty at event time.

To be precise, we define the child penalty

\[ P_t^g = \sum_{i=t}^{t+10} Y_{ist} - \sum_{i=t}^{t+10} \tilde{Y}_{ist} \]

\[ \equiv \sum_{i=t}^{t+10} Y_{ist} - \sum_{i=t}^{t+10} \tilde{Y}_{ist} \]

This measures the impact of children relative to the year just before the first childbirth. We control for time trends through event-time dummies because, conditional on age and year, there is variation in event time driven by variations in ages at which individuals have their first child. Kleven, Landais, and Søgaard (2018) lay out the identification assumptions for each country, the earnings of men and women evolve similarly before parenthood—after adjusting for life cycle and time trends—but diverge differently after. The results confirm that the existence of large child penalties is a pervasive phenomenon. In each country, the earnings of men and women evolve similarly before parenthood—after adjusting for life cycle and time trends—but diverge differently after.

Figures 1–3 show the effects of parenthood on earnings across the different countries. The notes reveal some striking differences. First, the series have plateaued. Ten years after childbirth, men due to children.

Long-run penalty:

- Denmark: 21%
- Sweden: 26%

Note: See the notes to Figure 1.

Source: Kleven et al. AEA-PP 2019

**Figure 1. Child Penalties in Earnings in Scandinavian Countries**
Figure 2. Child Penalties in Earnings in English-Speaking Countries

Source: Kleven et al. AEA-PP 2019

**Figure 2. Child Penalties in Earnings in English-Speaking Countries**

Earnings relative to event time -1

Event time (years)

Long-run penalty:
United Kingdom: 44%
United States: 31%

First childbirth

- Men - United Kingdom
- Men - United States
- Women - United Kingdom
- Women - United States
In the long run, the short-run child penalty is about twice as large in Sweden as it is in Denmark. Swedish mothers catch up with Danish mothers over time such that their child penalty is only slightly larger after 10 years. Sweden is also the only country where childbirth is associated with penalties as high as 51–61 percent. While parenthood reduces female employment across countries is that it allows for direct comparisons. The partial correlation of child penalties focus on government policies such as parental leave and childcare provision that directly affect mothers' incentives. These include taxes, transfers, and family policies such as parental leave and childcare reforms and childcare expansions allow for compelling quasi-experimental analyses of the impacts of parental leave and public childcare on the impact of such policies on female labor supply and gender gaps. There is a voluminous literature on the extensive margin and wage-rate effects. In the United States and the United Kingdom, the employment effects of subsequent children and therefore depend on total fertility. However, differential fertility is unlikely to drive the variation in child penalties across countries. For example, while the English-speaking countries exhibit the largest penalties, Germany is Austria, a country where the combination of high fertility. While parenthood reduces female employment, the extensive margin is a key driver of penalties in those countries. Of particular interest to work. There is a voluminous literature on the impact of such policies on female labor supply and gender gaps. Tax and transfer systems play a role in shaping the dynamics of child penalties. Their setting considers the particularities of their context—perinatal events, the age of children, and educational attainment of the parents. One set of explanations for the differences in child penalties over time. We see that these countries feature less dramatic short-run effects, but that the effects are growing over time.

Figure 3. Child Penalties in Earnings in German-Speaking Countries

Source: Kleven et al. AEA-PP 2019

- First childbirth
- Long-run penalty: Austria: 51%
  Germany: 61%
- Austria: Men - Austria, Women - Austria
- Germany: Men - Germany, Women - Germany
Average Annual Hours of Work of Employees

US has 40 hour/week and no mandatory paid vacation

1968: 4th week of paid vacation

1982: 5th week + 39 hours/week

2000-2: 35 hours/week

Source: Saez AEA-PP’21

Source: OECD database online. Includes all ages, genders, and part-time, full-time, overtime.
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)


