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

Emmanuel Saez

Berkeley: Spring 2010

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

OUTLINE

1) Labor Supply Elasticity Estimation: Methodological Issues

2) Estimates of hours/participation elasticities

3) Responses to low-income transfer programs (EITC)

4) Inter-temporal Labor Supply Models

5) Macro Estimates of Labor Supply

6) Elasticity of Taxable Income

REFERENCES

- Surveys in labor economics:

  a) Pencavel (1986) *Handbook of Labor Economics* vol 1

  b) Heckman and Killingsworth (1986) *Handbook of Labor Econ* vol 1

  c) Blundell and MaCurdy (1999) *Handbook of Labor Economics* vol 3

- Surveys in public economics:

  a) Hausman (1985) *Handbook of Public Economics* vol 1

  b) Moffitt (2003) *Handbook of Public Economics* vol 4

  c) Saez, Slemrod, and Giertz (2009) JEL in prep

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 $u(c,l)$ $\uparrow c$, and $\downarrow l$

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 τ and grant $G \Rightarrow c = w^p(1-\tau)l + G$]

Individual solves

$$\max_{c,l} u(c,l)$$

subject to $c = wl + y$

LABOR SUPPLY BEHAVIOR

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

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

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

Slutsky equation: $\partial l^c/\partial w = \partial l/\partial w - l\partial l/\partial y \Rightarrow \epsilon^c = \epsilon^u - \eta$

IMPORTANT SPECIAL CASE: NO INCOME EFFECTS

Quasi-linear utility function $u(c,l) = c - h(l)$

$\max_l wl + y - h(l) \Rightarrow h'(l) = w$

$\Rightarrow$ Marshallian $l^u(w,y) = l(w)$ labor supply independent of y

$\Rightarrow$ Hicksian $l^c(w,u) = l(w)$ labor supply independent of y [parallel indifference curves]

$\Rightarrow$ Identical uncompensated and compensated labor supply

$\Rightarrow \eta = 0$ and $\epsilon^c = \epsilon^u > 0$

Iso-elastic utility function: $u(c,l) = c - a\frac{l^{1+1/\varepsilon}}{1+1/\varepsilon} \Rightarrow w = C \cdot l^\varepsilon$

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.]

β measures uncompensated wage effects, and γ income effects [can be converted to ε^u, η]

BASIC CROSS SECTION RESULTS

1. Male workers [primary earners when married] (Pencavel, 1986 survey):

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

PROBLEMS WITH OLS ESTIMATION OF LABOR SUPPLY EQUATION

1) Econometric issues

a) Unobserved heterogeneity [tax instruments]

b) Measurement error in wages and division bias [tax instruments]

c) Selection into labor force [selection models]

d) Endogenous tax rates [non-linear budget set methods]

2) Extensive vs. intensive margin responses [participation models]

3) Non-hours responses [taxable income]

ISSUE 1: 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 ϵ_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** $\Rightarrow$ Tax changes provide more compelling identification

ISSUE 2: Measurement error in hours

In general w computed as earnings / hours $\Rightarrow$ Can create division bias

Let l^* denote true hours, l observed hours

Compute $w = e/l$ where e is earnings

$$\Rightarrow \ \log l = \log l^* + \mu$$
$$\Rightarrow \ \log w = \log e - \log l = \log e - \log l^* - \mu = \log w^* - \mu$$

Spurious negative correlation between $\log l$ and $\log w$ [e.g, workers with high misreported hours also have low imputed wages] biasing elasticity estimate downward

Solution: tax instruments again

ISSUE 3: Non-participation

Consider model with fixed costs of working, where some individuals choose not to work

Wages are unobserved for non-labor force participants

Thus, OLS regression on workers only includes observations with $l_i > 0$

This can bias OLS estimates: low wage earners must have very high unobserved propensity to work to find it worthwhile

Requires a selection correction pioneered by Heckman in 1970s (e.g. Heckit, Tobit, or ML estimation): problem is that identification is based on strong functional form assumptions [See Killingsworth and Heckman (1986) for implementation]

Current approach: use panel data to distinguish entry/exit from intensive-margin changes

Extensive vs. Intensive Margin

Related issue: want to understand effect of taxes on labor force participation decision

With fixed costs of work, individuals may jump from non-participation to part time or full time work (non-convex budget set)

This can be handled using a discrete choice model:

$$P = \phi(\alpha + \varepsilon \log(1 - \tau) - \eta y)$$

where $P \in \{0, 1\}$ is an indicator for whether the individual works

Function ϕ typically specified as logit, probit, or linear prob model

Note: here it is critical to have tax variation; regression cannot be run with wage variation

Non-hours responses

Traditional literature focused purely on hours of work and labor force participation

Problem: income taxes distort many margins beyond hours of work

a) Non-hours margins may be more important quantitatively

b) Hours very hard to measure (most ppl report 40 hours per week)

Two solutions in modern literature:

a) Focus on total earnings (wl) [or taxable income] as a broader measure of labor supply

b) Focus on subgroups of workers for whom hours are better measured, e.g. taxi drivers

ISSUE 4: NON-LINEAR BUDGET SETS

Actual tax system is not linear but piece-wise linear with varying marginal tax rate τ due to (a) means-tested transfer programs, (b) progressive individual income tax, (c) ceiling in payroll tax

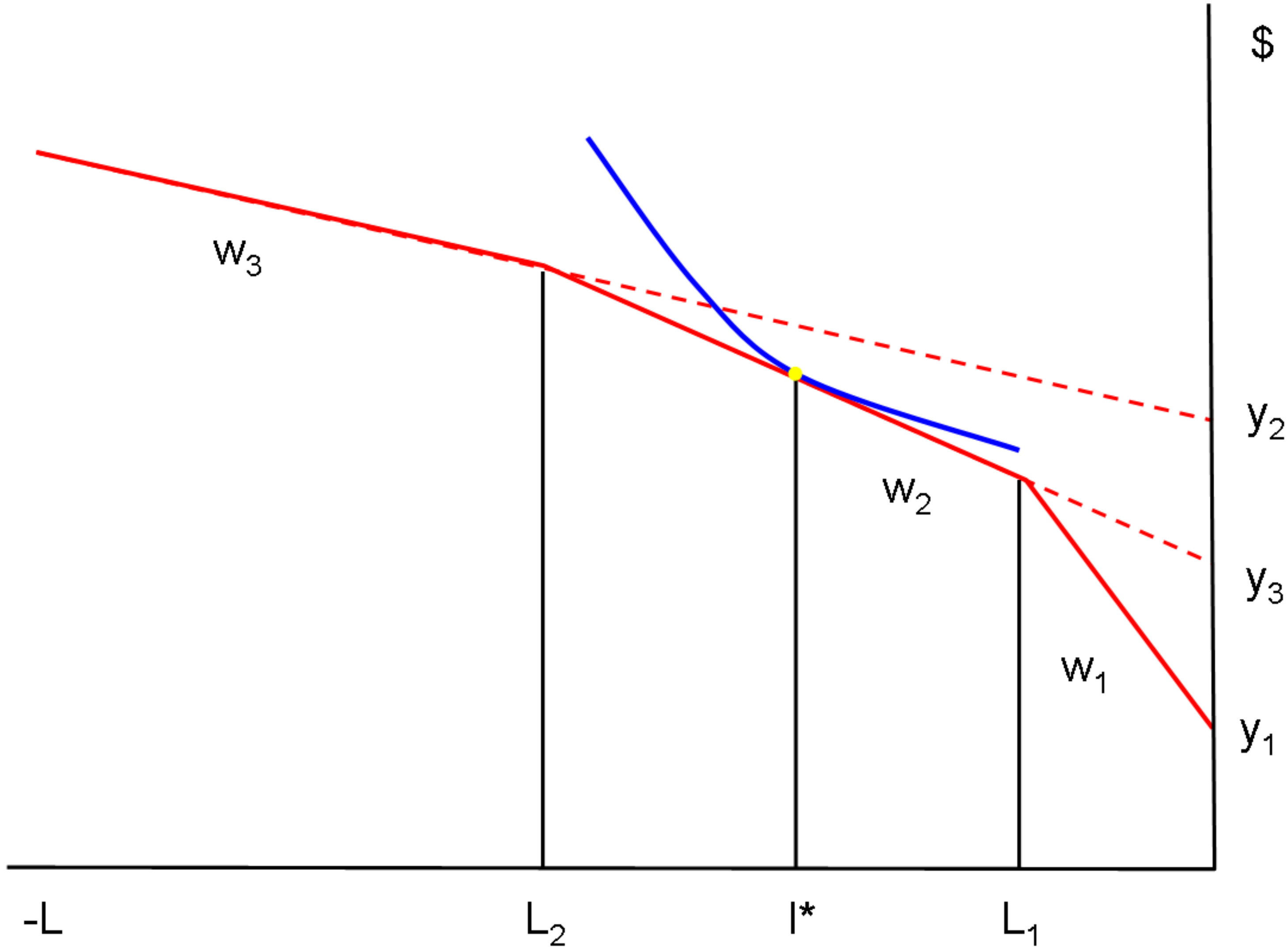
Individual maximization problem:

$\max u(w^p l - T(w^p l), l) \Rightarrow$ FOC $u_c w^p (1 - T') + u_l = 0$

Same theory applies when considering the linearized tax system $c = wl + y$ with $w = w^p(1 - T')$ and y defined as virtual income (intercept of budget with x-axis when setting $l = 0$)

Main complications: (a) w [and y] become endogenous to choice of l, (b) FOC may not hold if individual bunches at a kink, (c) FOC may not characterize the optimum choice

Non-Linear Budget Set Estimation: Virtual Incomes



Source: Hausman (Hbk 1985)

ISSUE 4: NON-LINEAR BUDGET SETS

Non-linear budget set creates two problems:

1) Model mis-specification: OLS regression no longer recovers structural elasticity parameter ε of interest

Two reasons: (a) underestimate response because people pile up at kink and (b) mis-estimate income effects

2) Econometric bias: τ_i depends on income $w_i l_i$ and hence on l_i

Tastes for work are positively correlated with $\tau_i \rightarrow$ downward bias in OLS regression of hours worked on net-of-tax rates

Solution to problem #2: only use reform-based variation in tax rates. But problem #1 requires fundamentally different estimation method

NON-LINEAR BUDGET SET METHOD

Issue addressed by non linear budget set studies pioneered by Hausman in late 1970s (Hausman, 1985 PE handbook chapter)

Method uses a structural model of labor supply

Key point: the method uses the standard cross-sectional variation in pre-tax wages w^p for identification. Taxes are seen as a problem to deal with rather than an opportunity for identification.

New literature identifying labor supply elasticities using tax changes has a totally different perspective: taxes are seen as an **opportunity** to identify labor supply

NON-LINEAR BUDGET SET METHOD

1) Assume an uncompensated labor supply equation:

$$l = \alpha + \beta w(1-\tau) + \gamma y + \epsilon$$

2) Error term ϵ is normally distributed with variance σ^2

3) Observed variables: w_i, τ_i, y_i, and l_i

4) Technique: (a) construct likelihood function given observed labor supply choices on NLBS, (b) find parameters (α, β, γ) that maximize likelihood

5) Important insight: need to use "virtual incomes" in lieu of actual unearned income with NLBS

NLBS Likelihood Function (2 brackets)

Individual can locate on first bracket, on second bracket, or at the kink l_K

Likelihood = probability that we see individual i at labor supply l_i given a parameter vector

Decompose likelihood into three components

Component 1: individual i on first bracket: $0 < l_i < l_K$

$$l_i = \alpha + \beta w_i(1 - \tau^1) + \gamma y^1 + \epsilon_i$$

Error $\epsilon_i = l_i - (\alpha + \beta w_i(1 - \tau^1) + \gamma y^1)$. Likelihood:$L_i = \phi((l_i - (\alpha + \beta w_i(1 - \tau^1) + \gamma y^1)/\sigma)$

Component 2: individual i on second bracket: $l_K < l_i$: $L_i = \phi((l_i - (\alpha + \beta w_i(1 - \tau^2) + \gamma y^2)/\sigma)$

NLBS Likelihood Function

Now consider individual i located at the kink point

1) If tax rate is τ^1 and virtual income y^1 individual wants to work $l > l_K$

2) If tax is τ^2 and virtual income y^2 individual wants to work $l < l_K$

3) These inequalities imply:

$$
\begin{aligned}
\alpha + \beta w_i(1-\tau^1) + \gamma y^1 + \epsilon_i &> l_K > \alpha + \beta w_i(1-\tau^2) + \gamma y^2 + \epsilon_i \\
l_K - (\alpha + \beta w_i(1-\tau^1) + \gamma y^1 &< \epsilon_i < l_K - (\alpha + \beta w_i(1-\tau^2) + \gamma y^2)
\end{aligned}
$$

4) Contribution to likelihood is probability that error lies in this range:

$$
\begin{aligned}
L_i &= \Psi[(l_K - (\alpha + \beta w_i(1-\tau^2) + \gamma y^2))/\sigma] \\
&\quad - \Psi[(l_K - (\alpha + \beta w_i(1-\tau^1) + \gamma y^1))/\sigma]
\end{aligned}
$$

Maximum Likelihood Estimation

1) Log likelihood function is $L = \sum_i \log L_i$

2) Final step is solving

$$\max L(\alpha, \beta, \gamma, \sigma)$$

3) In practice, likelihood function much more complicated because of more kinks, non-convexities, and covariates

4) But basic technique remains the same

Hausman (1981) Application

1) Hausman applies method to 1975 PSID cross-section

a) Finds significant compensated elasticities and large income effects

b) Elasticities larger for women than for men

2) Shortcomings of this implementation

a) Sensitivity to functional form choices, which is a larger issue with structural estimation

b) No tax reforms, so does not solve fundamental econometric problem that tastes for work may be correlated with w

NLBS and Bunching at Kinks

Subsequent studies obtain different estimates (MaCurdy, Green, and Paarsh 1990, Blomquist 1995)

a) Several studies find **negative** compensated wage elasticity estimates

b) Debate: impose requirement that compensated elasticity is positive or conclude that data rejects model?

Fundamental source of problem: labor supply model predicts that individuals should bunch at the kink points of the tax schedule

a) But we observe very little bunching at kinks (Heckman vs. Hausman), so model is rejected by the data

b) Interest in NLBS models diminished despite their conceptual advantages over OLS methods

Bunching at Kinks (Saez AEJ-EP'10)

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

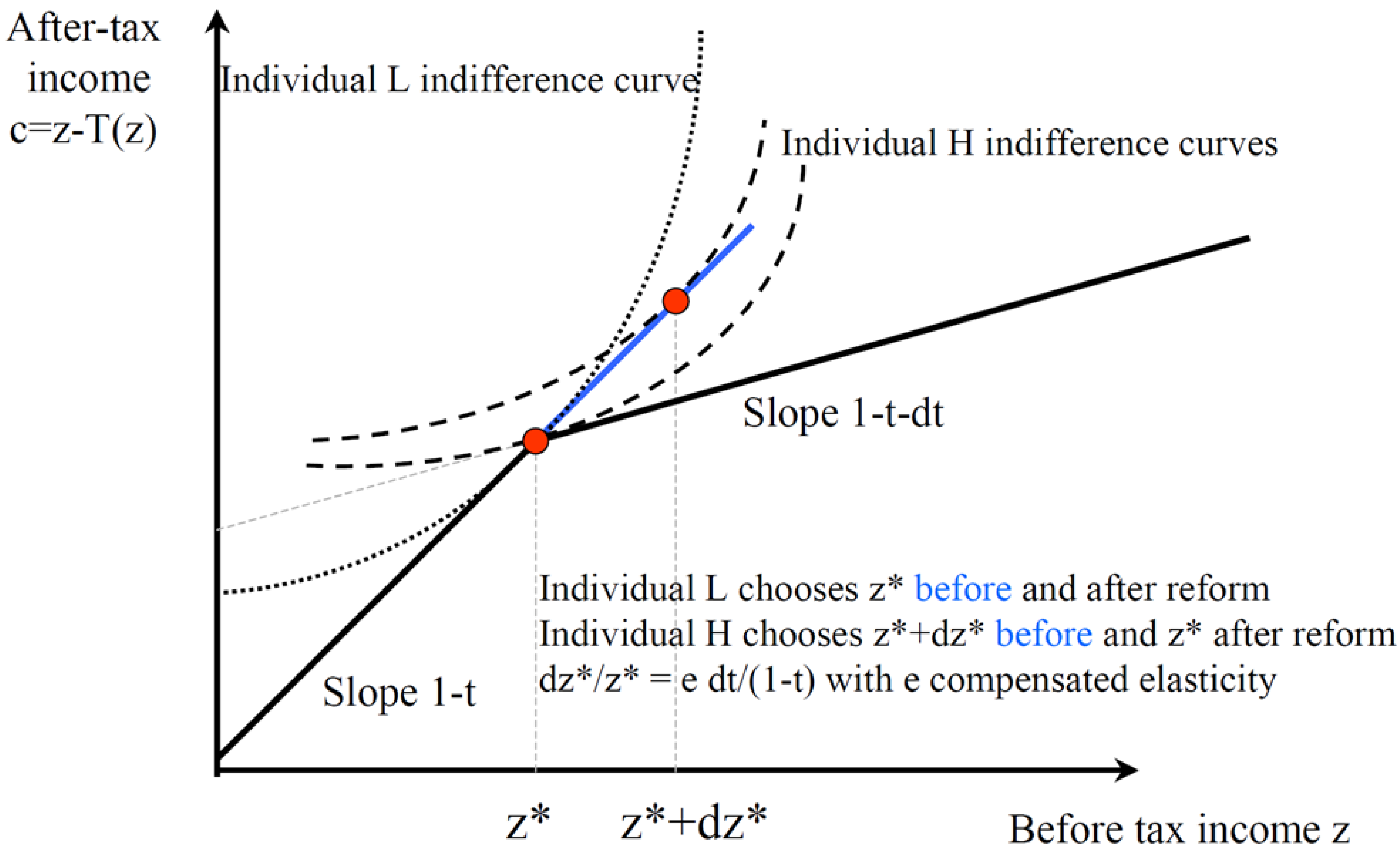
2) All other tax variation is contaminated by heterogeneity in tastes

3) Develops method of using bunching at kinks to estimate the compensated income elasticity

4) Idea: if this simple, non-parametric method does not recover positive compensated elasticities, then little value in additional structure of NLBS models

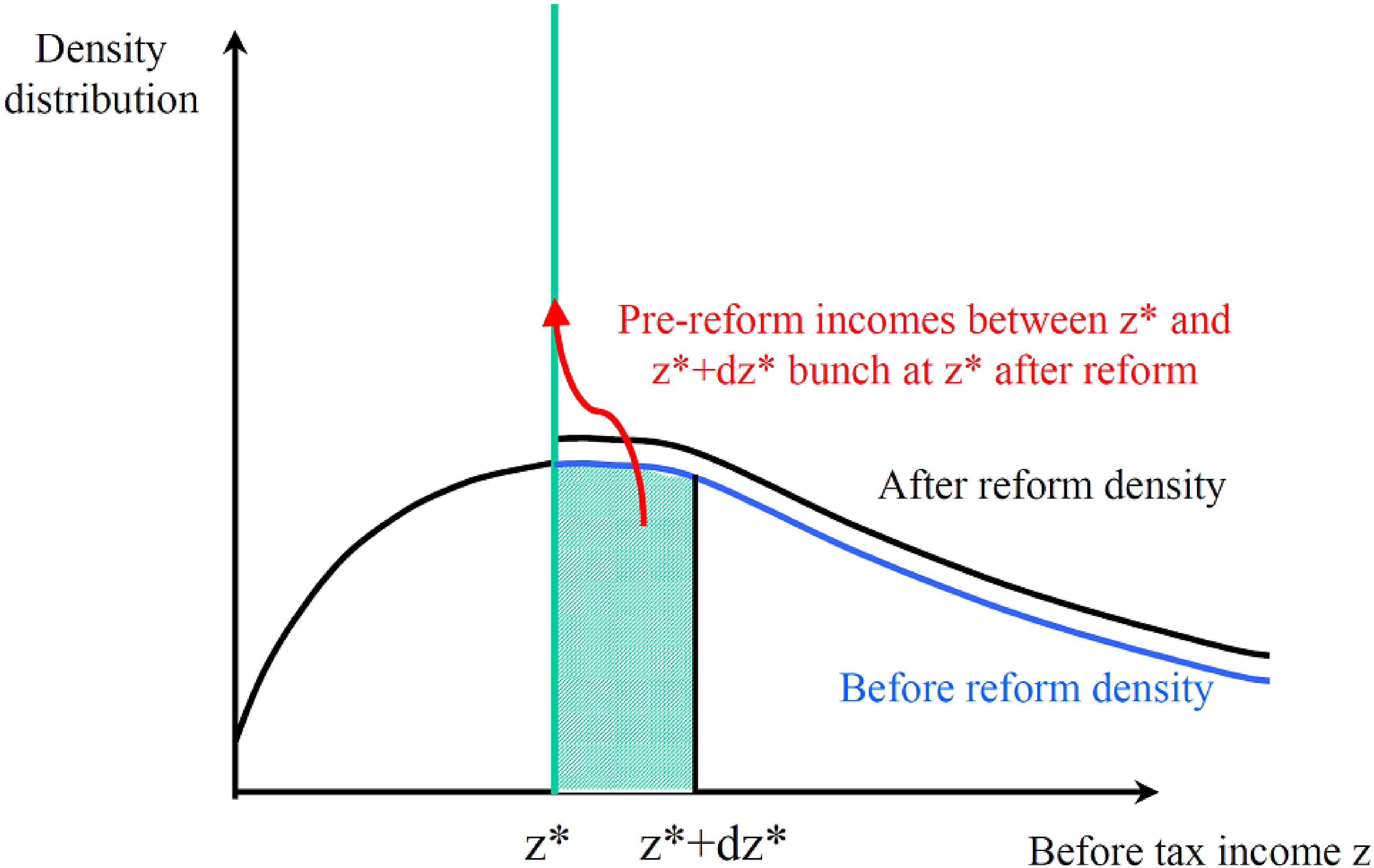
Formula for elasticity: $\varepsilon^c = \frac{dz/z^*}{dt/(1-t)}$ = excess mass at kink / change in NTR

A. Indifference curves and bunching



Source: Saez (2009)

B. Density Distributions and Bunching



Source: Saez (2009)

Bunching at Kinks (Saez, 2009)

1) Saez implements this method using individual tax return micro data (IRS public use files) from 1960 to 2004

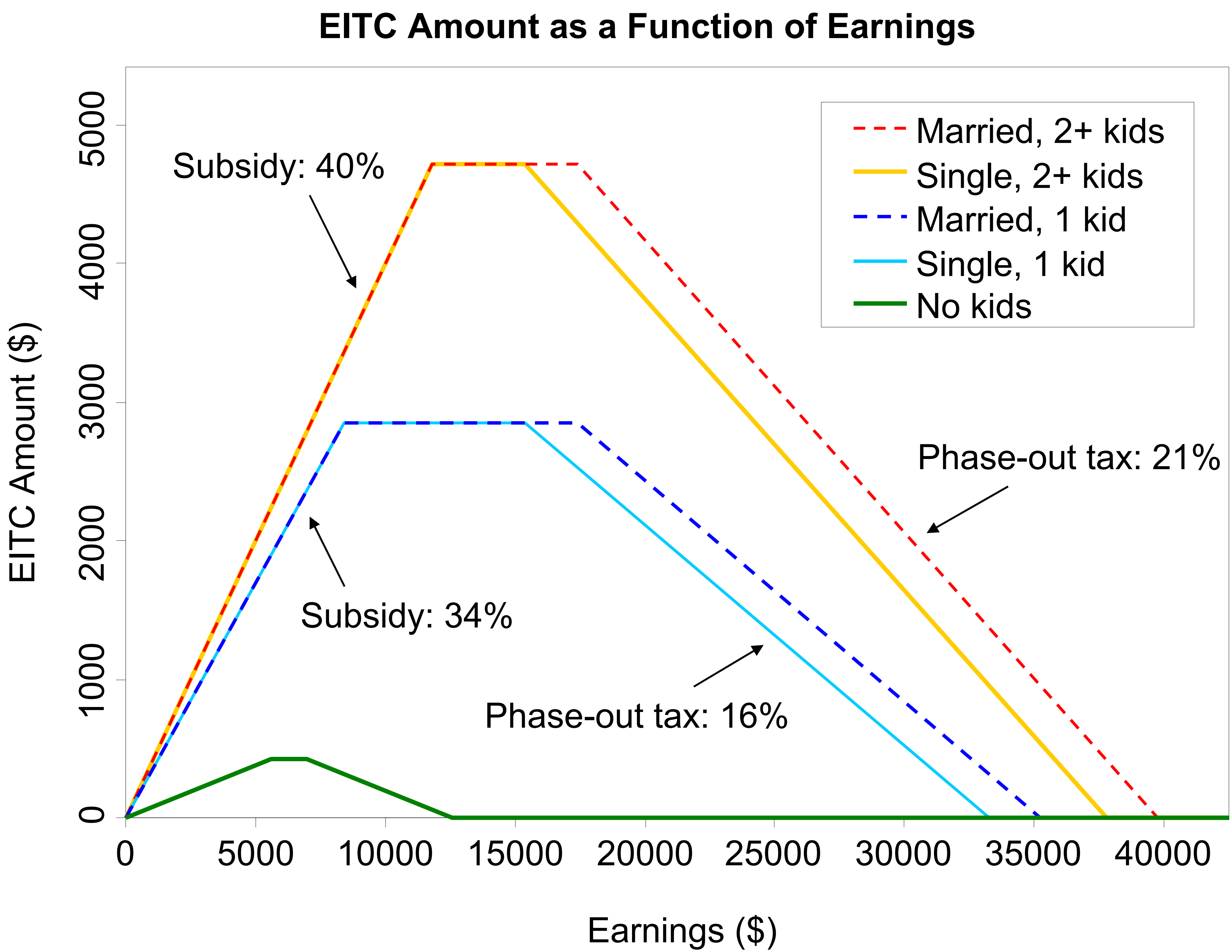
2) Advantage of dataset over PSID: very little measurement error

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
Married, 2+ kids
Single, 2+ kids
Married, 1 kid
Single, 1 kid
No kids
Subsidy: 40%
Subsidy: 34%
Phase-out tax: 21%
Phase-out tax: 16%
EITC Amount ($)
Earnings ($)
0
1000
2000
3000
4000
5000
0
5000
10000
15000
20000
25000
30000
35000
40000

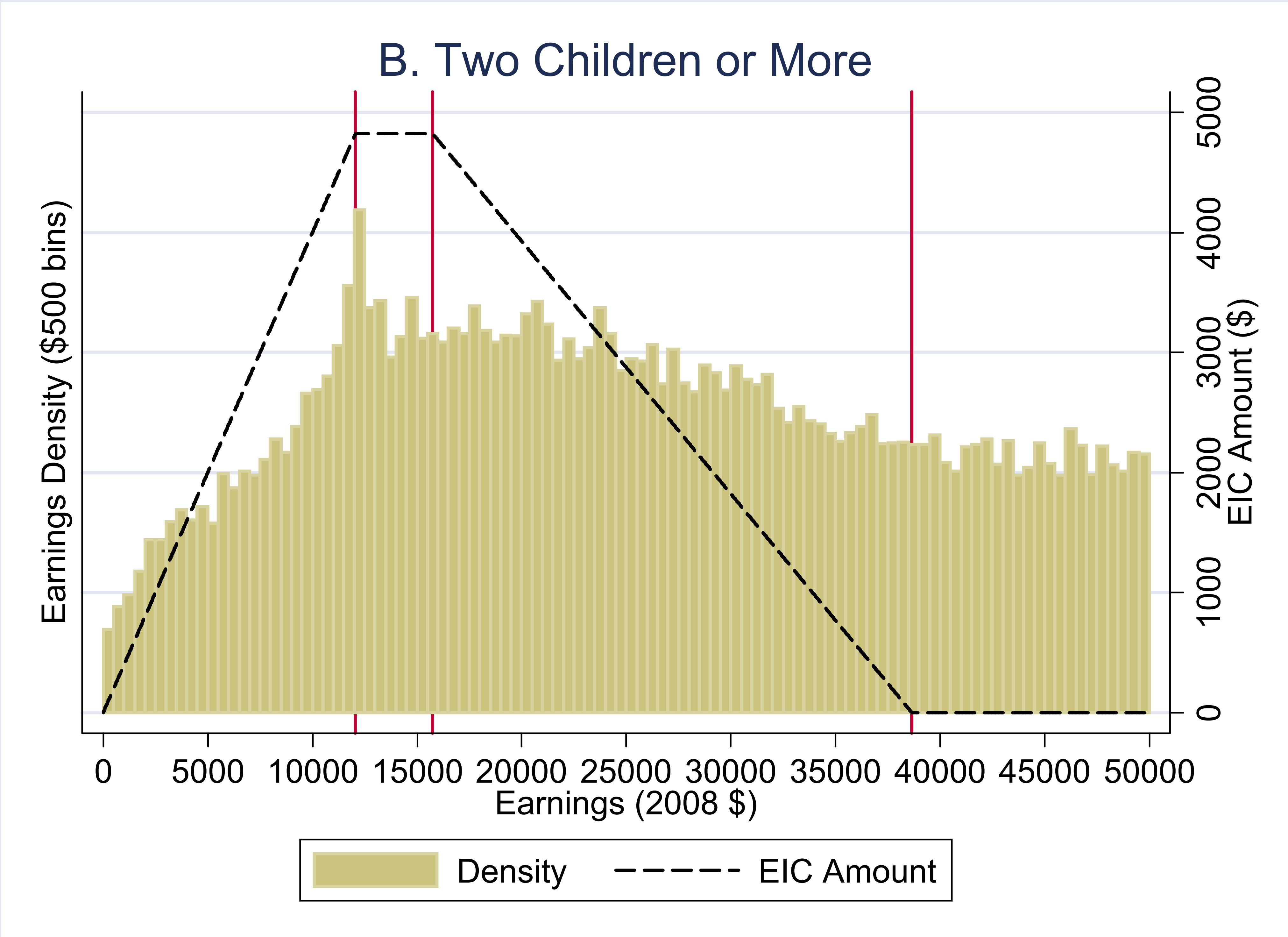
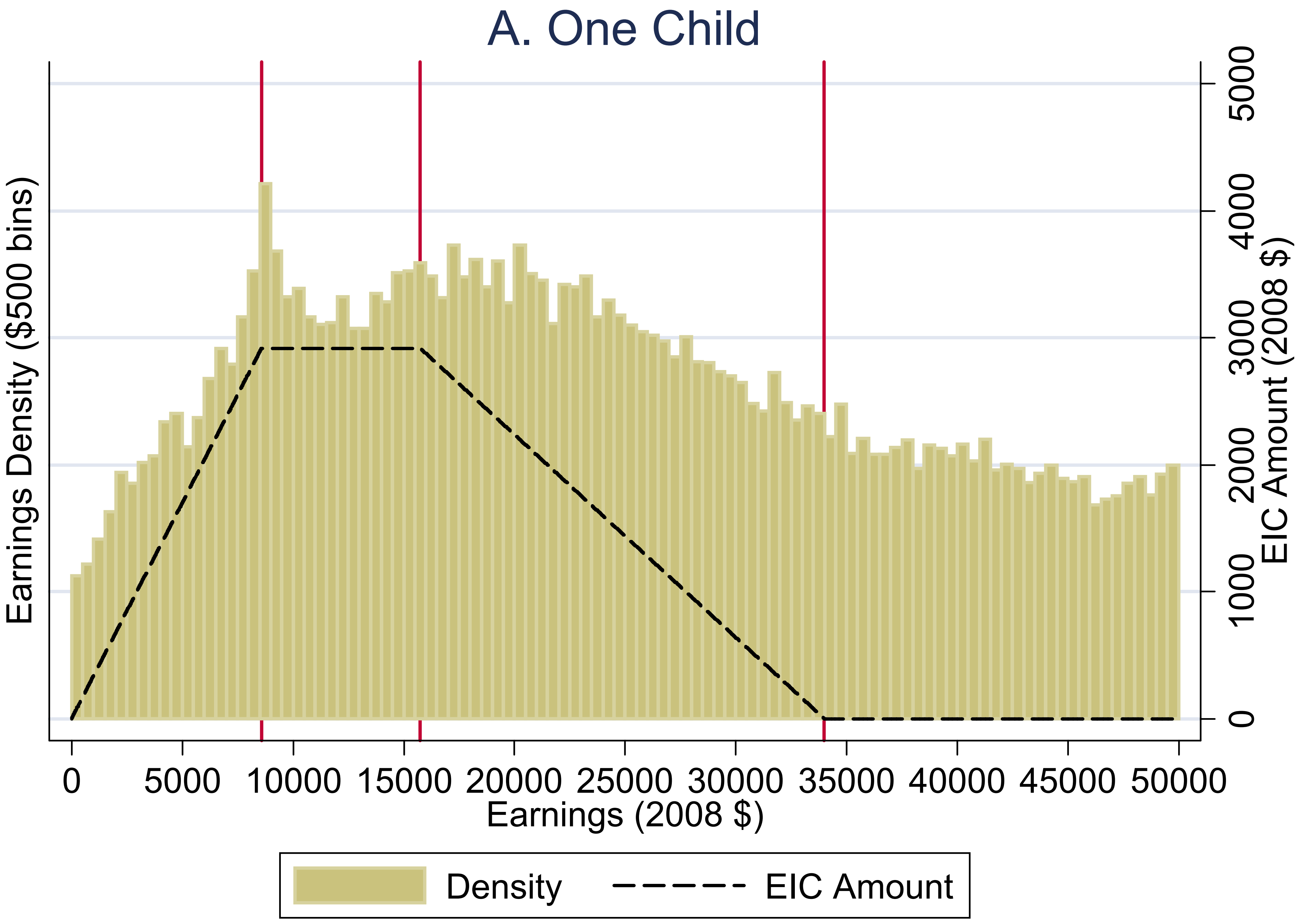
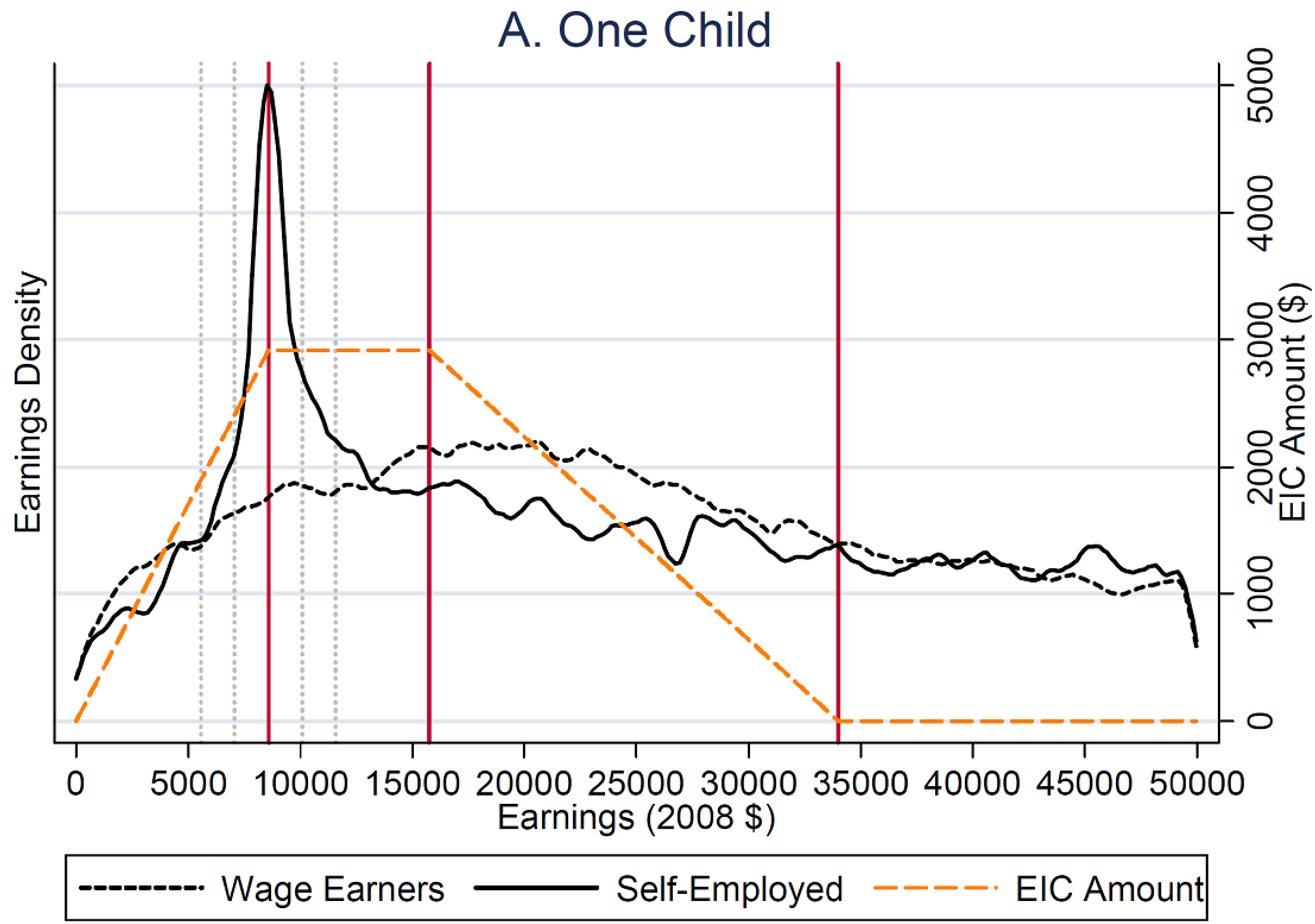


Figure 3. Earnings Density Distributions and the EITC

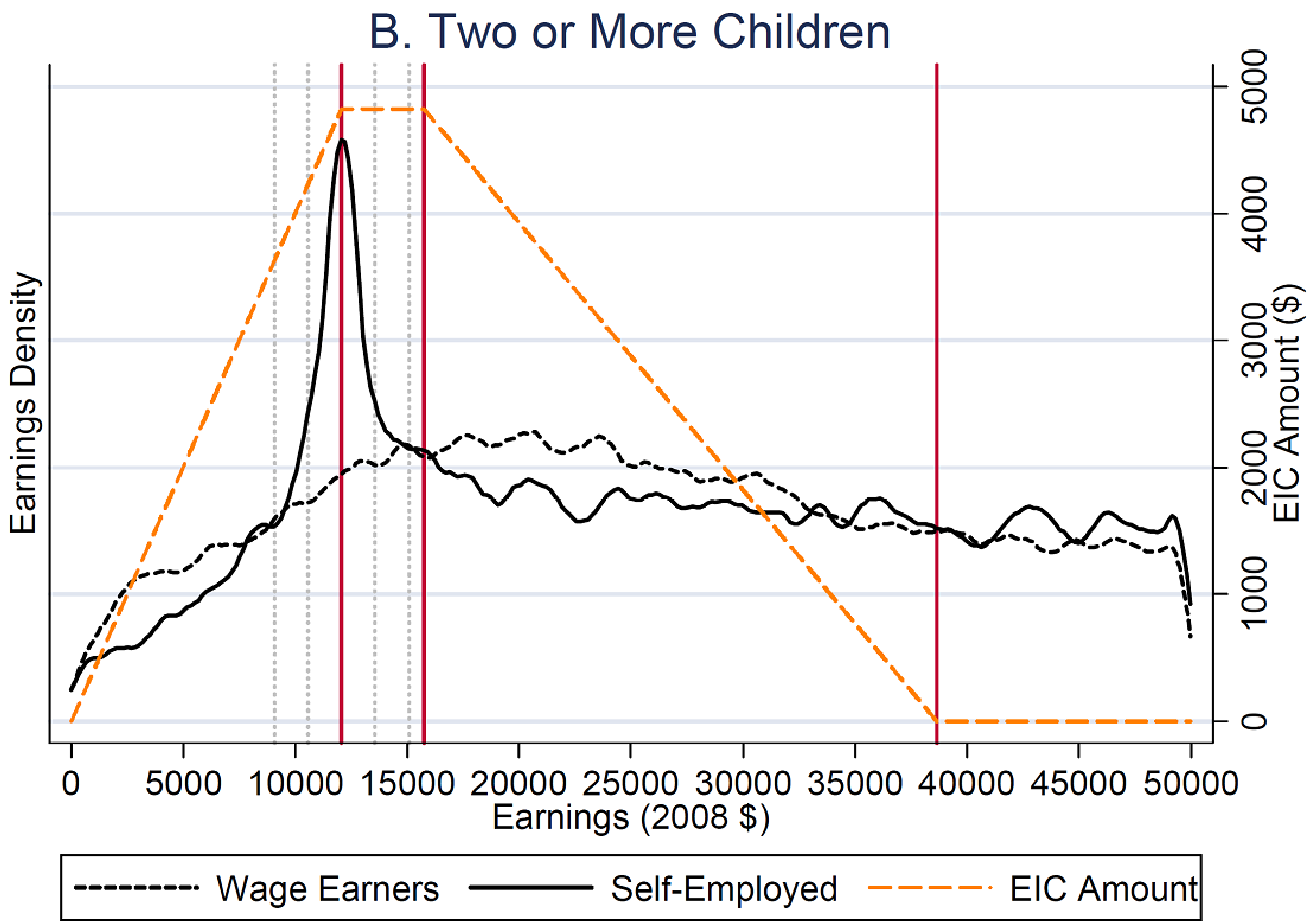

A. One Child
Earnings Density ($500 bins)
EIC Amount (2008 $)
0
1000
2000
3000
4000
5000
0
5000
10000
15000
20000
25000
30000
35000
40000
45000
50000
Earnings (2008 $)
Density
EIC Amount

Earnings Density and the EITC: Wage Earners vs. Self-Employed



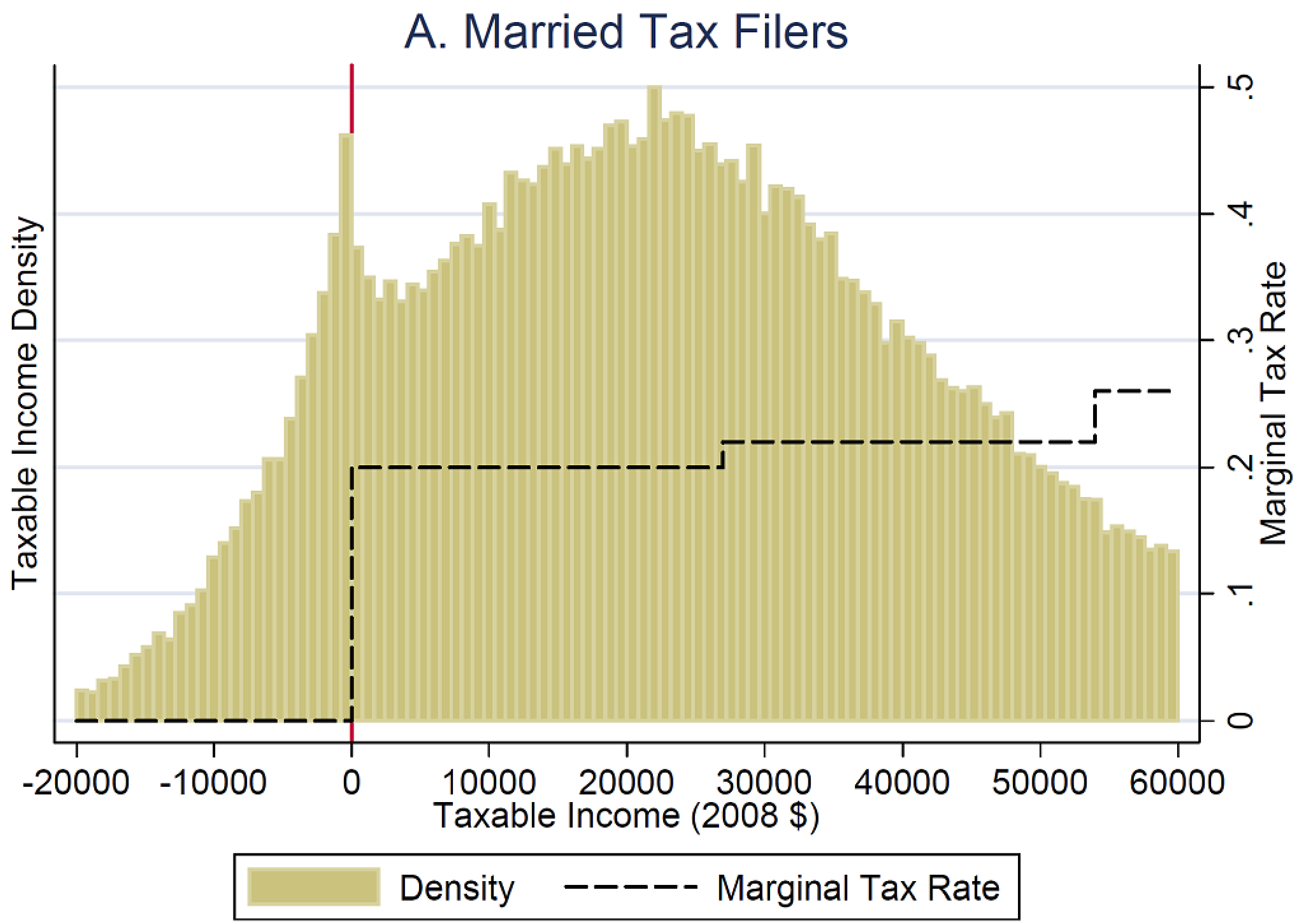
Source: Saez (2009)

Earnings Density and the EITC: Wage Earners vs. Self-Employed



Source: Saez (2009)

Taxable Income Density, 1960-1969: Bunching around First Kink



Source: Saez (2009)

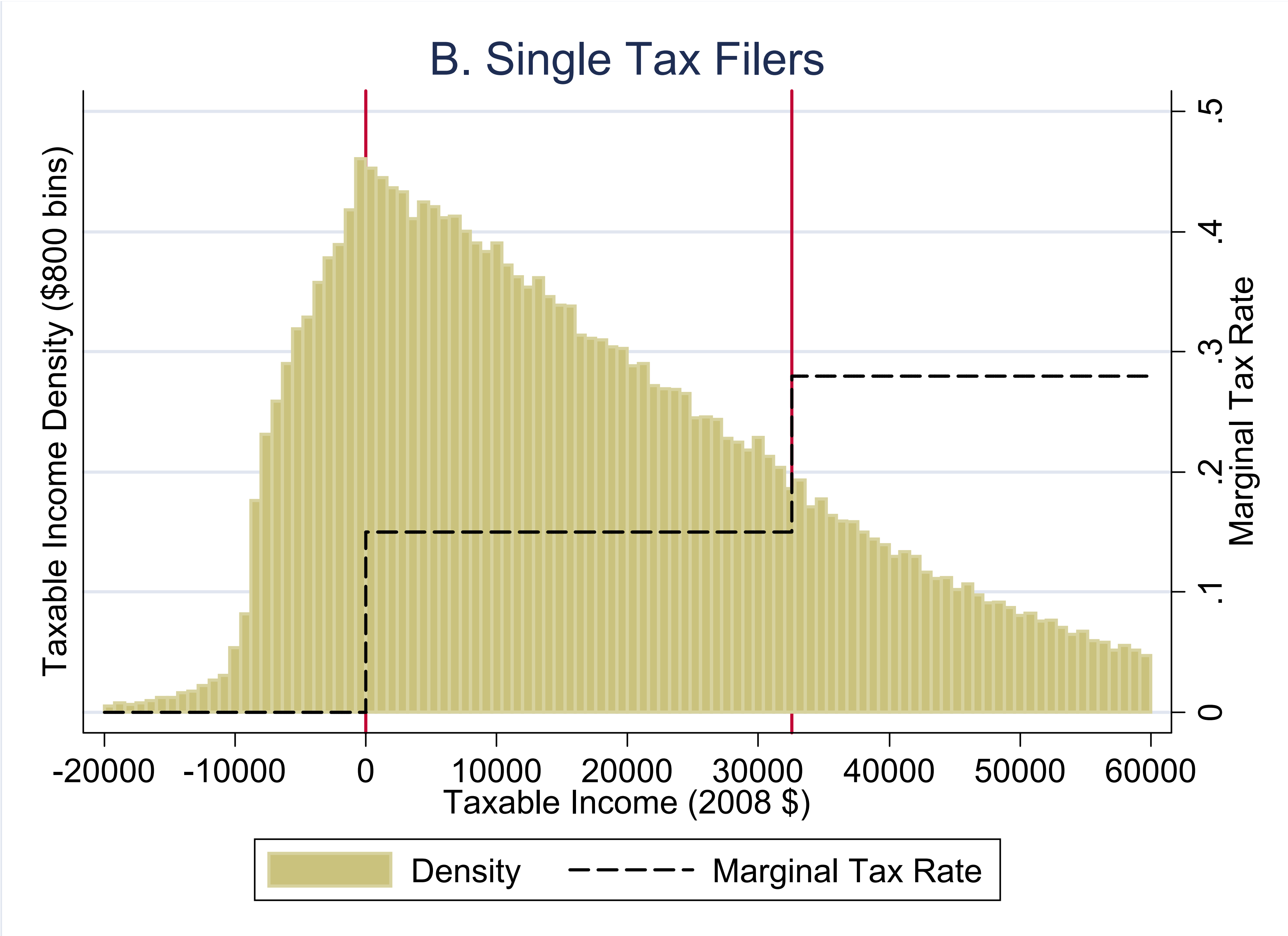


Figure 9. Taxable Income Density, 1988-2002

Friedberg 2000: Social Security Earnings Test

1) Uses CPS data on labor supply of retirees receiving Social Security benefits

2) Studies bunching based on responses to Social Security earnings test

3) Earnings test: phaseout of SS benefits with earnings above an exempt amount around \$14K/year

4) Today: Phaseout rate varies by age group: 50% (below 66), 33% (age 66), 0 (above 66)

5) Friedberg exploits 1983 reform (CPS age = age + 1):

(a) Before: test up to age 71, no test at age 72+

(b) After: test up to age 69, no test at age 70+

FIGURE 3-A.—EARNINGS DISTRIBUTION, 1980–81

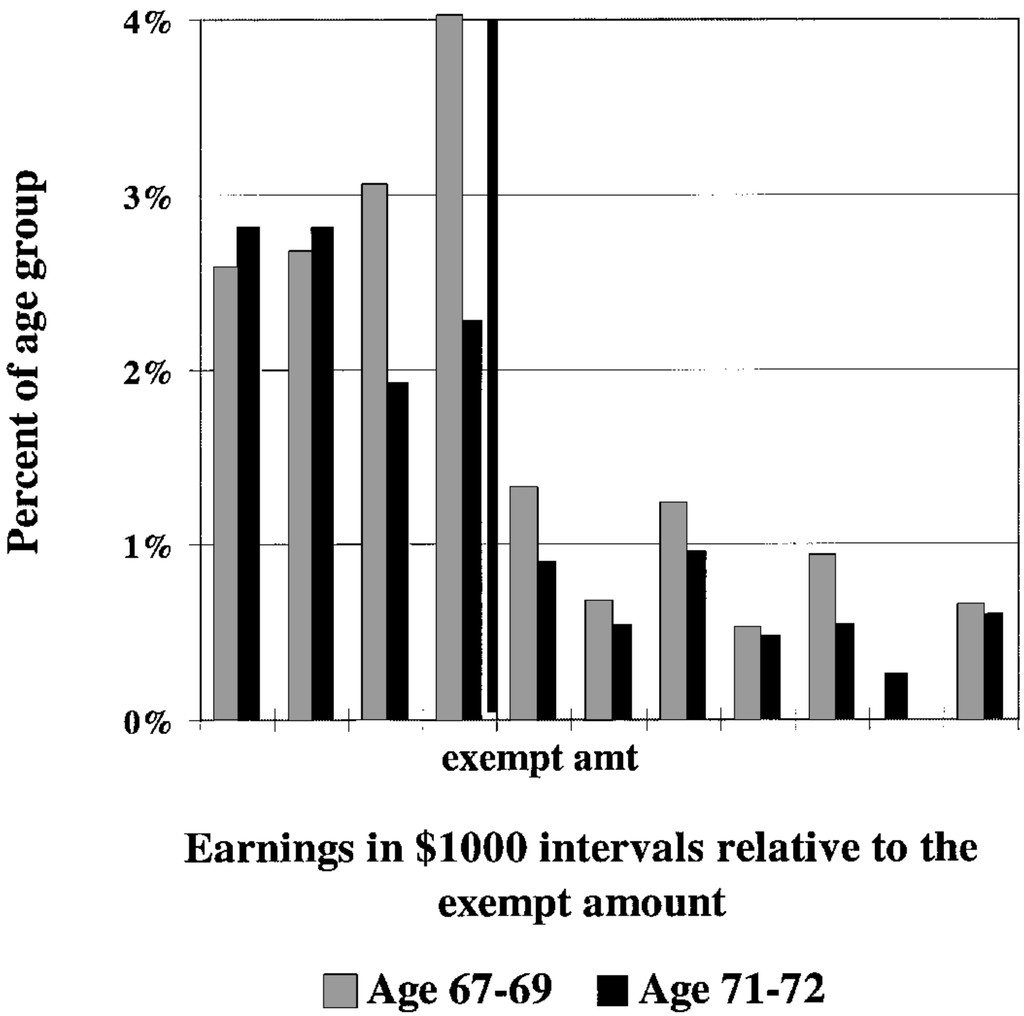


FIGURE 3-B.—EARNINGS DISTRIBUTION, 1980–81

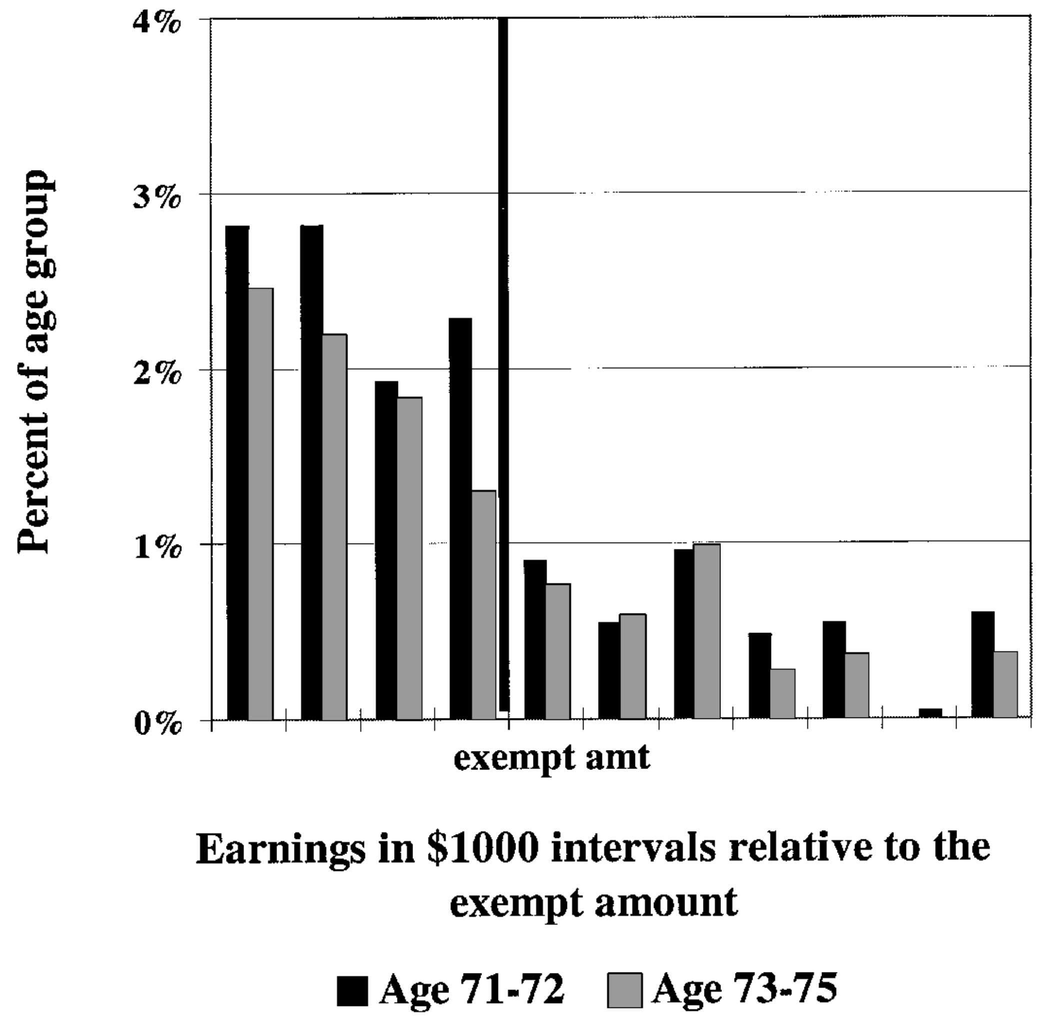


FIGURE 3-C.—EARNINGS DISTRIBUTION, 1984–86

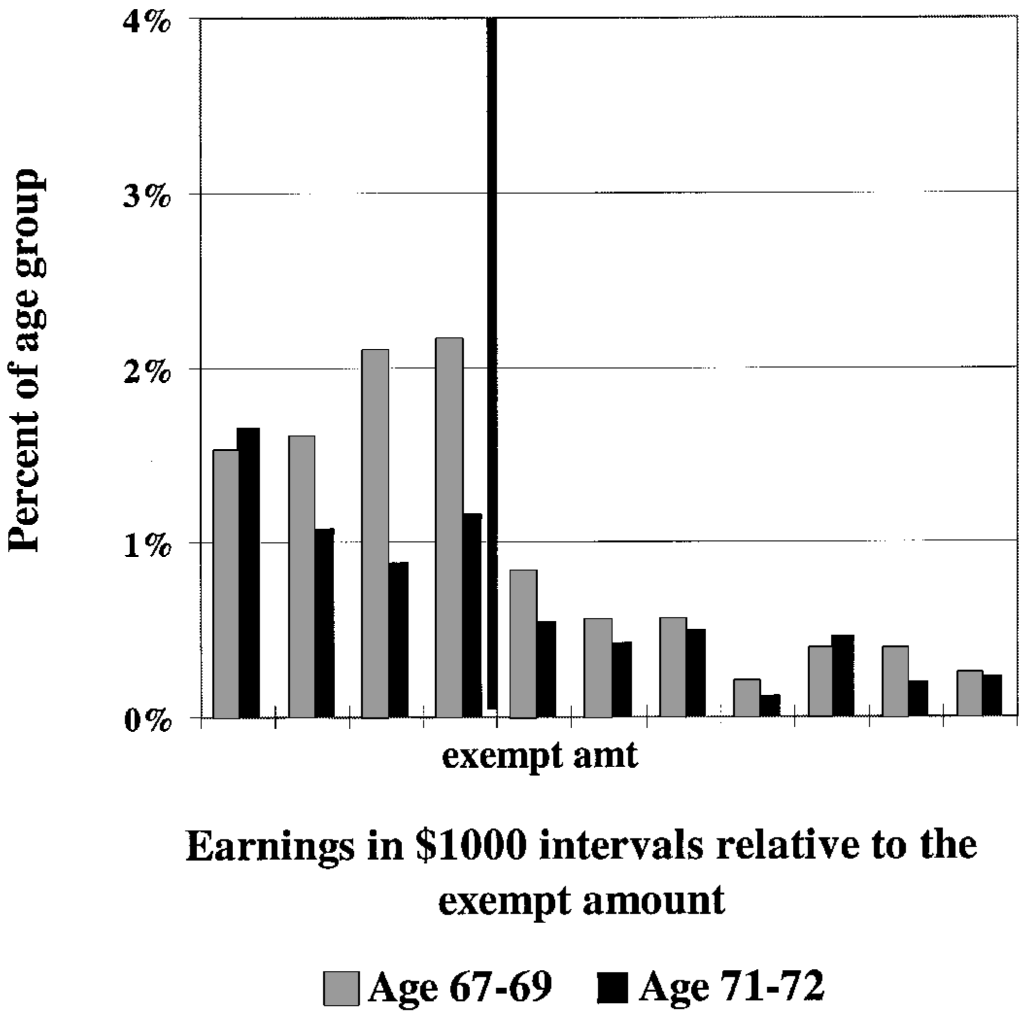
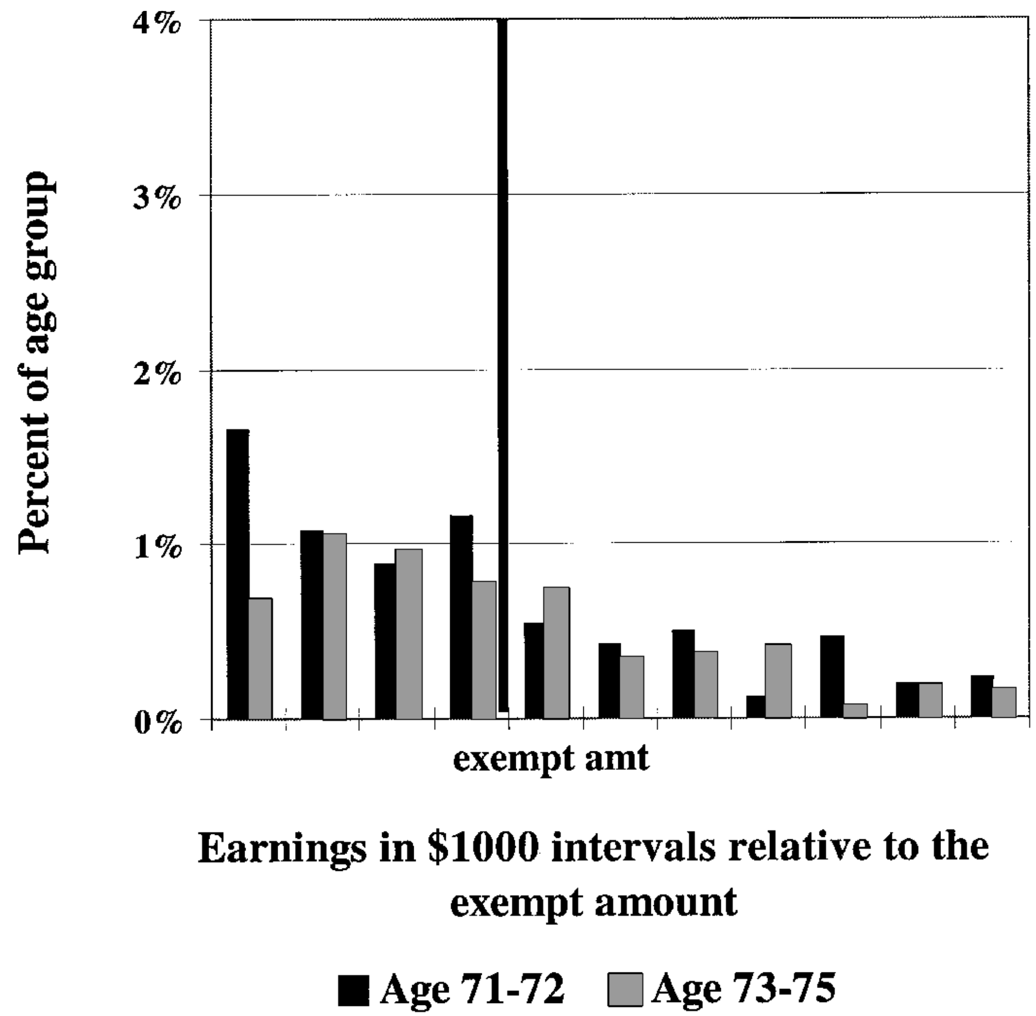


FIGURE 3-D.—EARNINGS DISTRIBUTION, 1984–86



Note: In 1983 the earnings test was eliminated for 70–71 year olds (71–72 year olds in the following March CPS) but was not changed for 62–69 year olds. See Figure 2 note.

Friedberg 2000: Estimates

1) Estimates elasticities using Hausman method, finds relatively large compensated and uncompensated elasticities

2) Ironically, lost social security benefits are considered delayed retirement with an actuarial adjustment of future benefits $\Rightarrow$ (a) No kink if person has average life expectancy, (b) kink if person has less than average life expectancy

3) So the one kink where we do find real bunching is actually not real! (people may not understand rules, or have myopia)

Why not more bunching at kinks?

1) True intensive elasticity of response may be small

2) Randomness in income generation process: Saez, 2002 shows that year-to-year income variation too small to erase bunching if elasticity is large

3) Information and salience

a) Liebman and Zeckhauser: "Schmeduling" (behavioral model where individuals confuse MTR with average tax rate)

b) Chetty and Saez (2009): information significantly affects bunching in EITC field experiment

4) Adjustment costs and institutional constraints (Chetty et al 2009)

Chetty, Friedman, Olsen, and Pistaferri (2009)

1) If workers face adjustment costs, may not reoptimize in response to tax changes of small size and scope in short run

a) Search costs, costs of acquiring information about taxes

b) Institutional constraints imposed by firms (e.g. 40 hour week)

2) Could explain why macro studies find larger elasticities

3) Question: How much are elasticity estimates affected by frictions?

Chetty et al. 2009: Model

1) Firms post jobs with different hours offers

2) Workers draw from this distribution and must pay search cost to reoptimize

3) Therefore not all workers locate at optimal choice

4) Bunching at kink and observed responses to tax reforms attenuated

Chetty et al. 2009: Testable Predictions

Model generates three predictions:

1) **[Size]** Larger tax changes generate larger observed elasticities

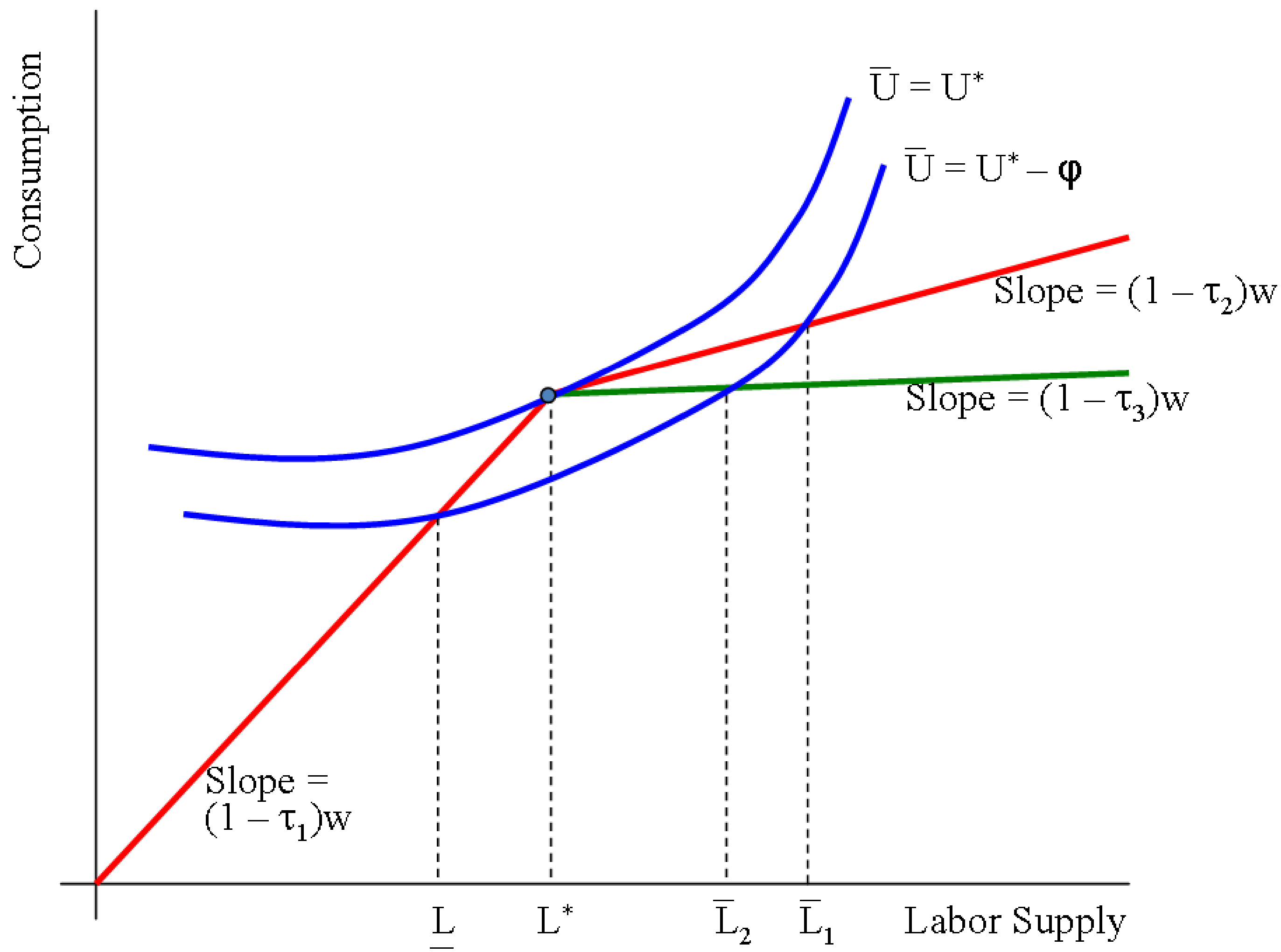
Large tax changes are more likely to induce workers to search for a different job

2) **[Scope]** Tax changes that apply to a larger group of workers generate larger observed elasticities

Firms tailor jobs to preferences of common workers

3) **[Search Costs]** Workers with lower search costs exhibit larger elasticities from individual bunching

Cost of Bunching at Bracket Cutoff Points in Tax Schedule



Source: Chetty et al. (2009)

Chetty et al. 2009: Data

Matched employer-employee panel data with admin tax records for full population of Denmark

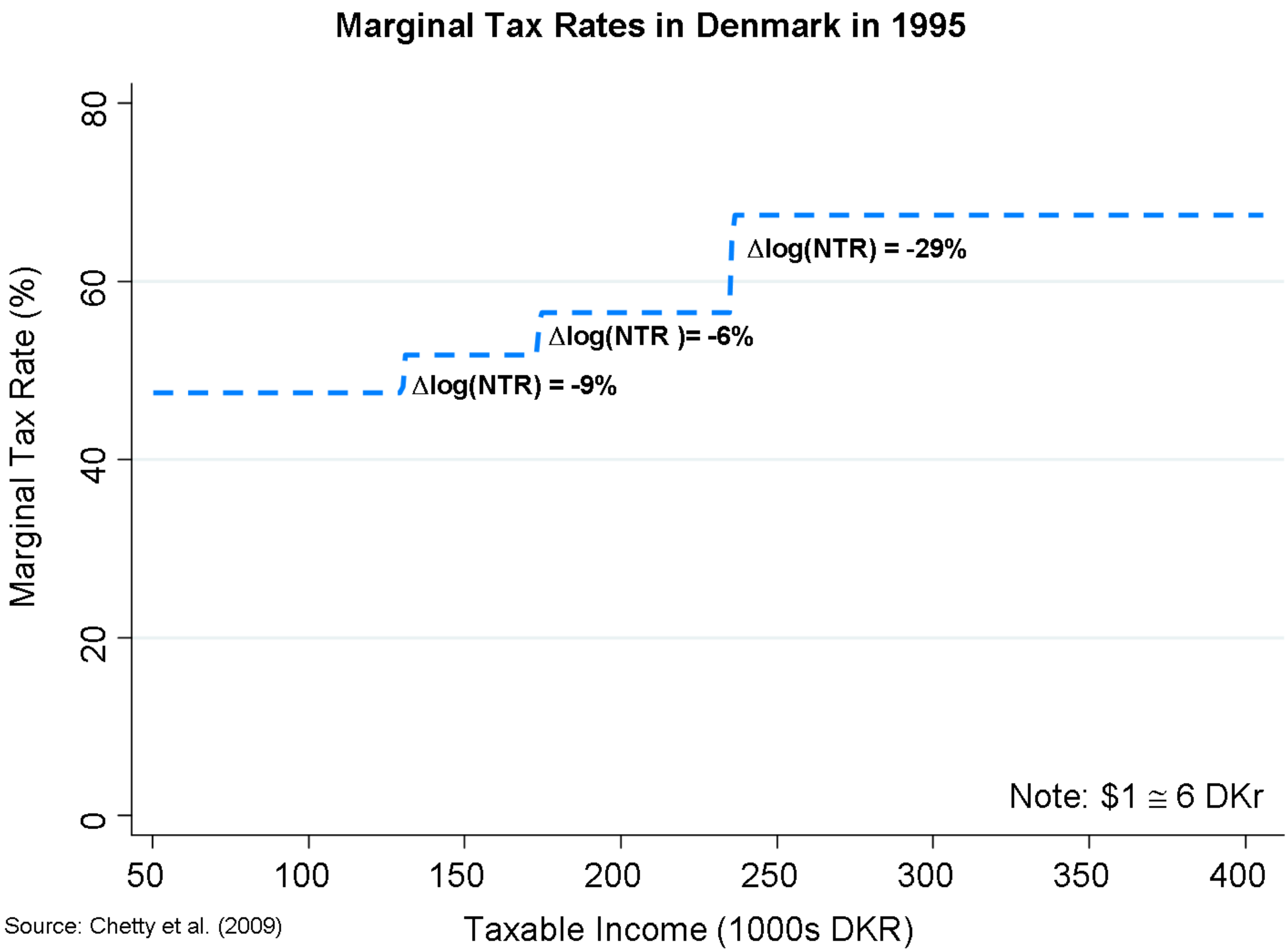
1) Income vars: wage earnings, capital and stock income, pension contributions

2) Employer vars: tenure, occupation, employer ID

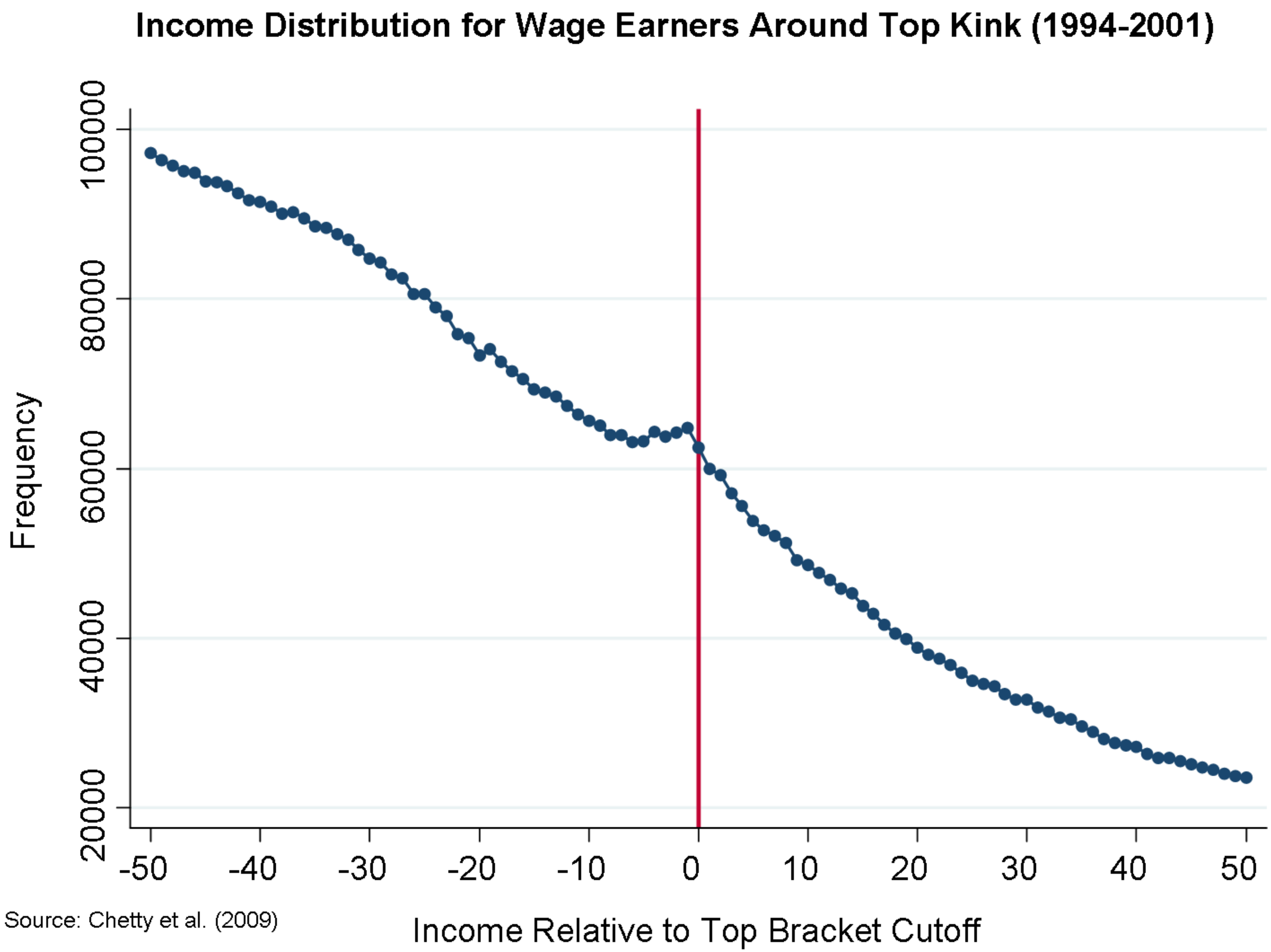
3) Demographics: education, spouse ID, kids, municipality

Sample restriction: Wage-earners aged 15-70, 1994-2001

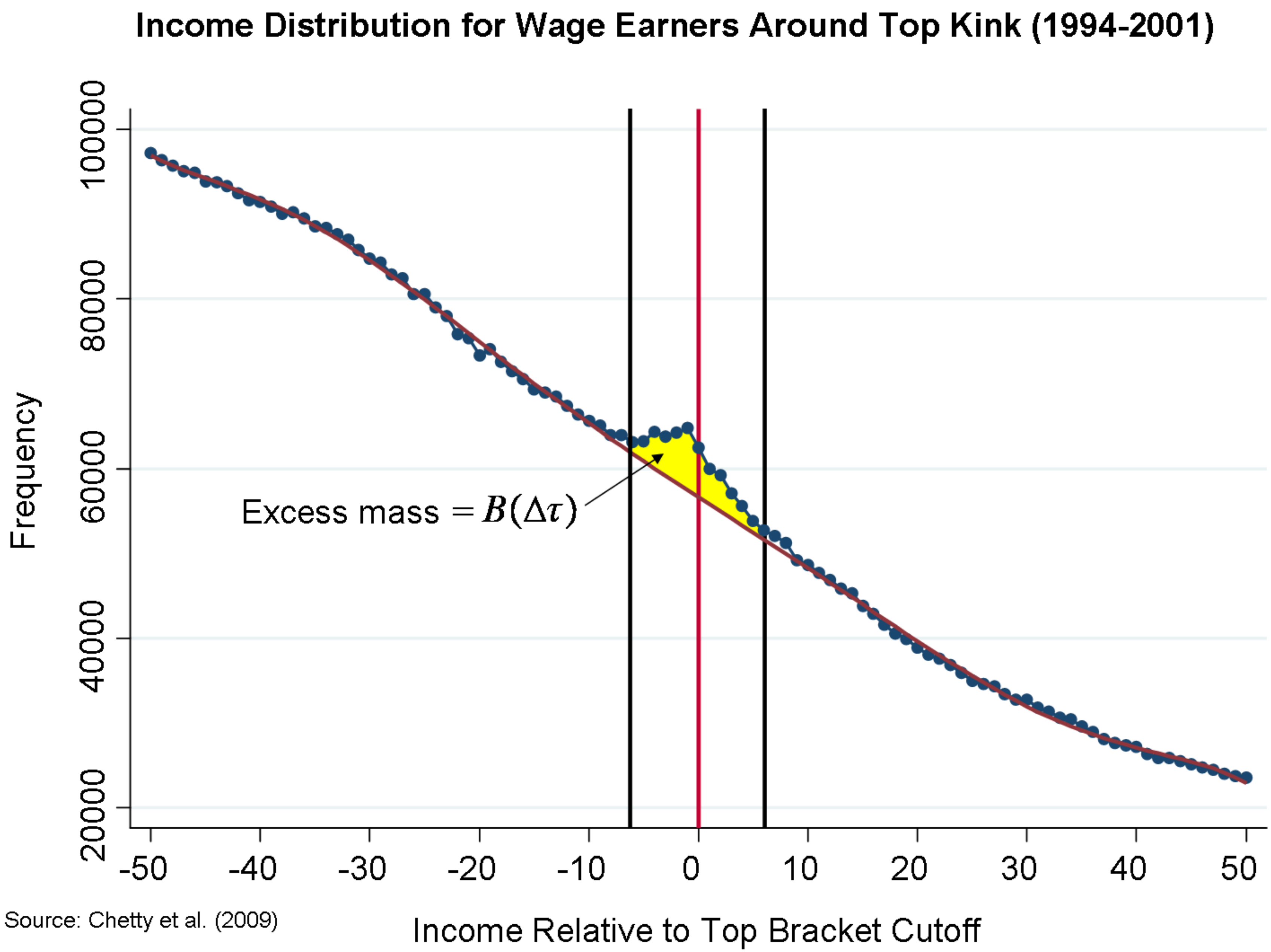
Approximately 2.42 million people per year

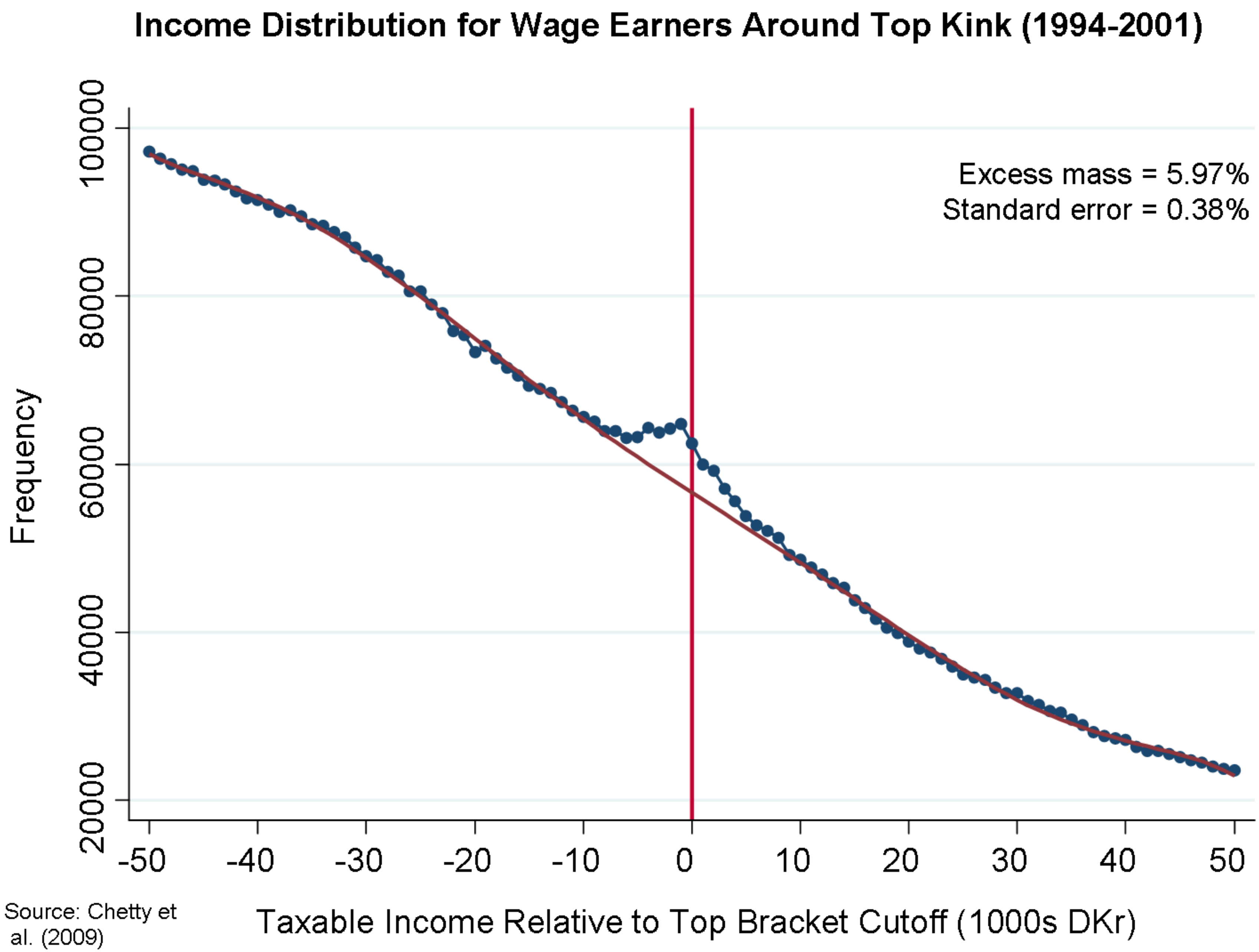


Source: Chetty et al. (2009)

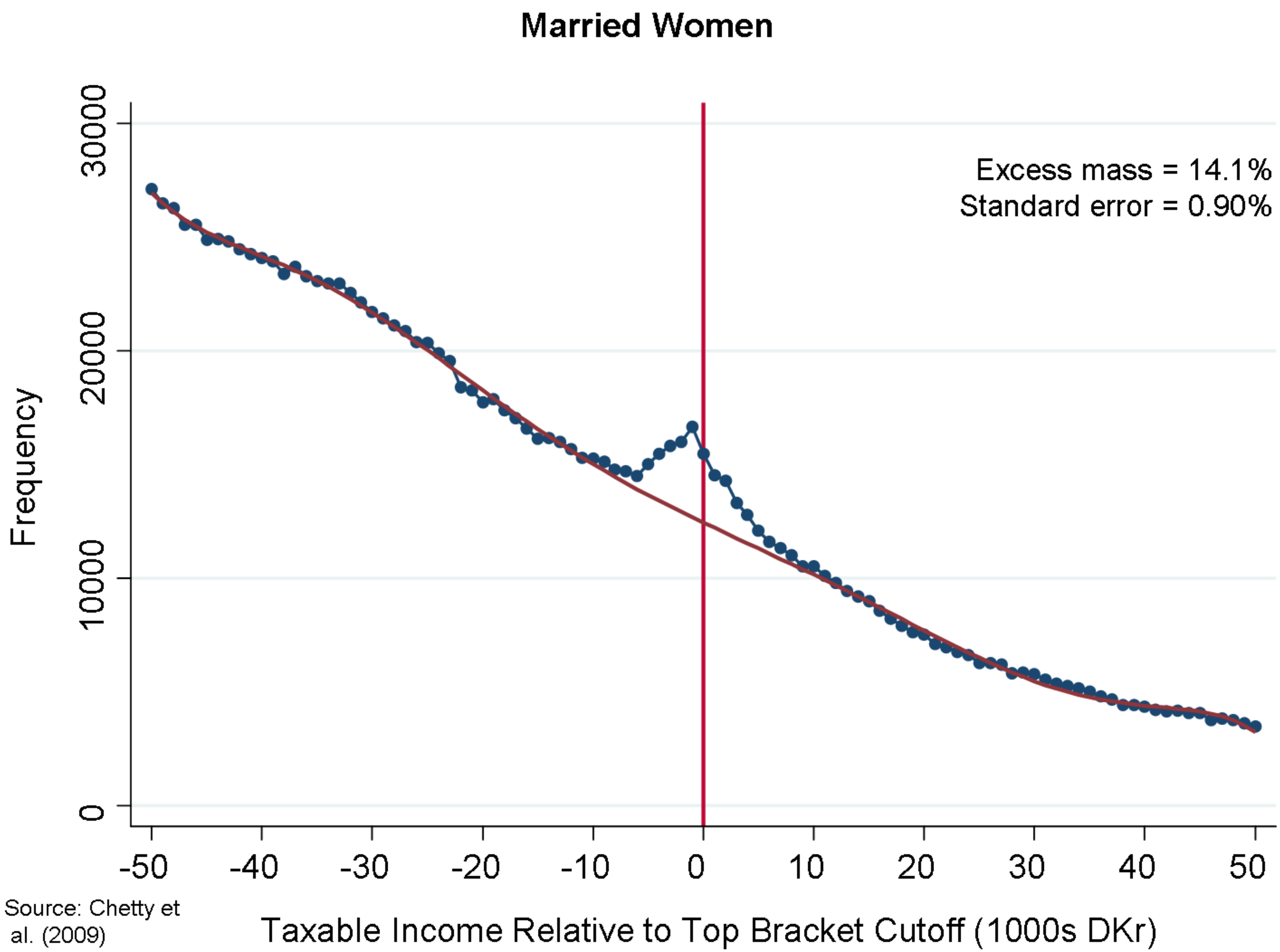


Source: Chetty et al. (2009)

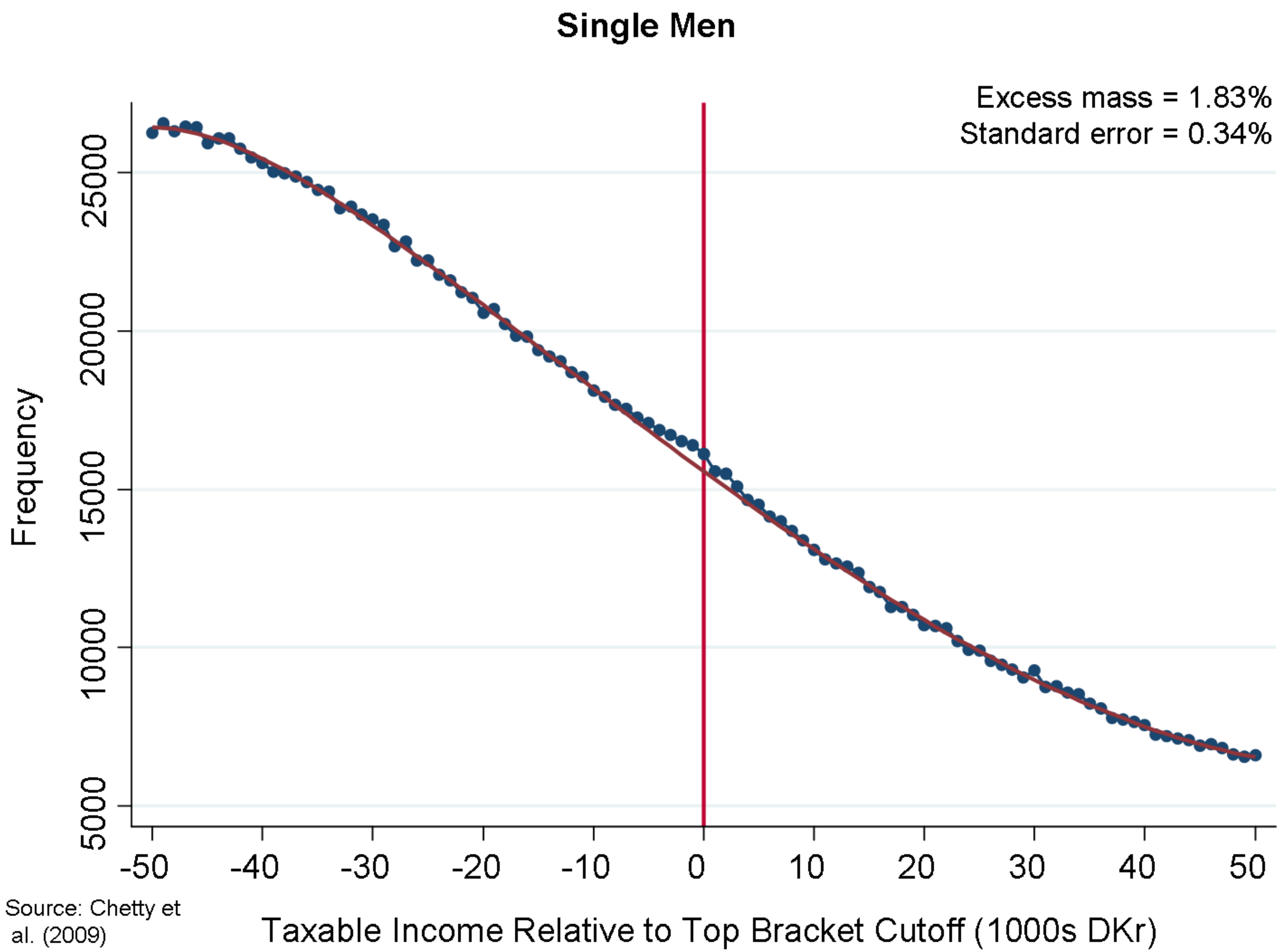

Income Distribution for Wage Earners Around Top Kink (1994-2001)
Frequency
100000
80000
60000
40000
20000
Excess mass = $B(\Delta\tau)$
-50
-40
-30
-20
-10
0
10
20
30
40
50
Income Relative to Top Bracket Cutoff
Source: Chetty et al. (2009)



Source: Chetty et al. (2009)

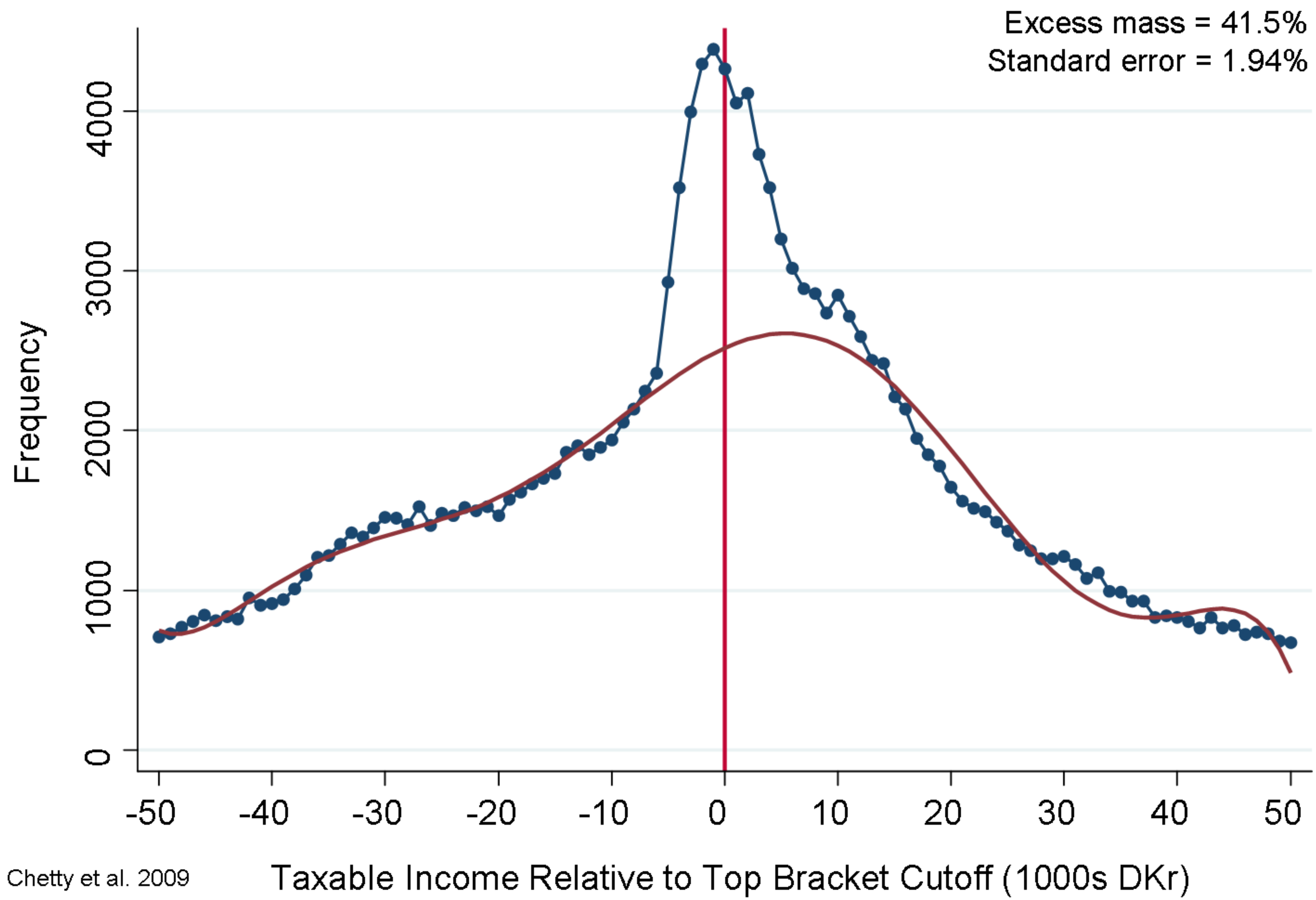


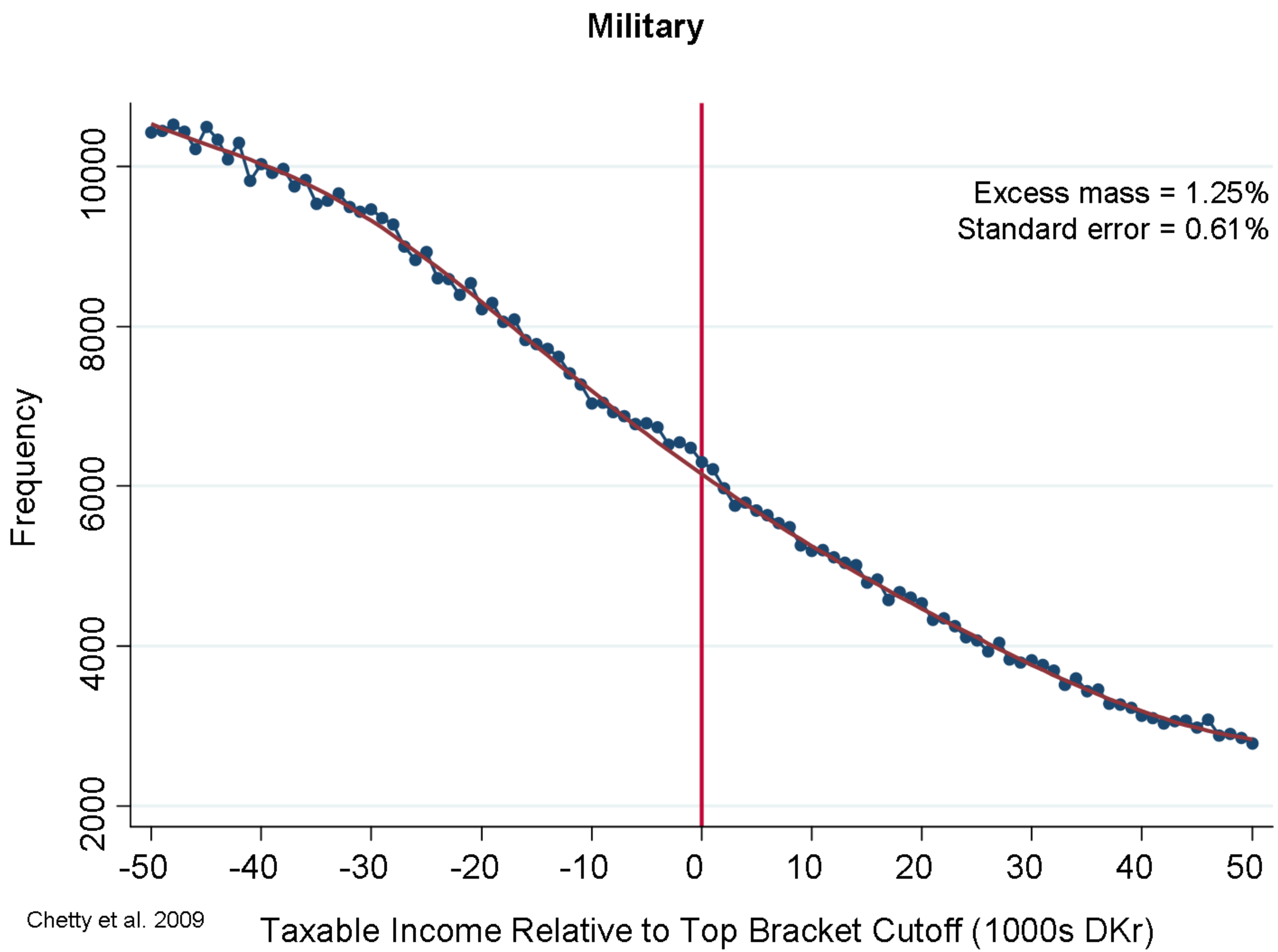
Source: Chetty et al. (2009)



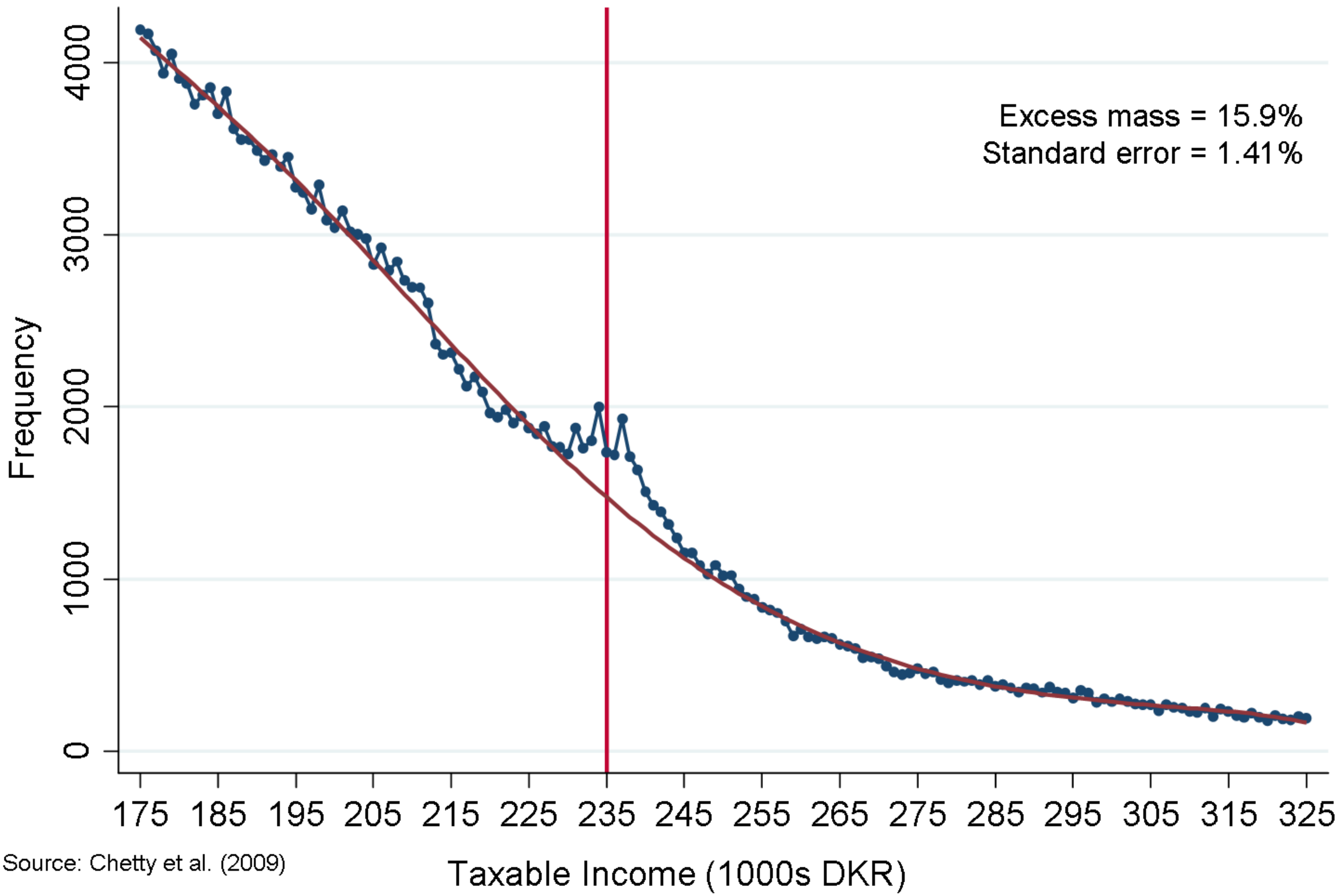
Source: Chetty et al. (2009)

Married Female Professionals with Above Median Experience

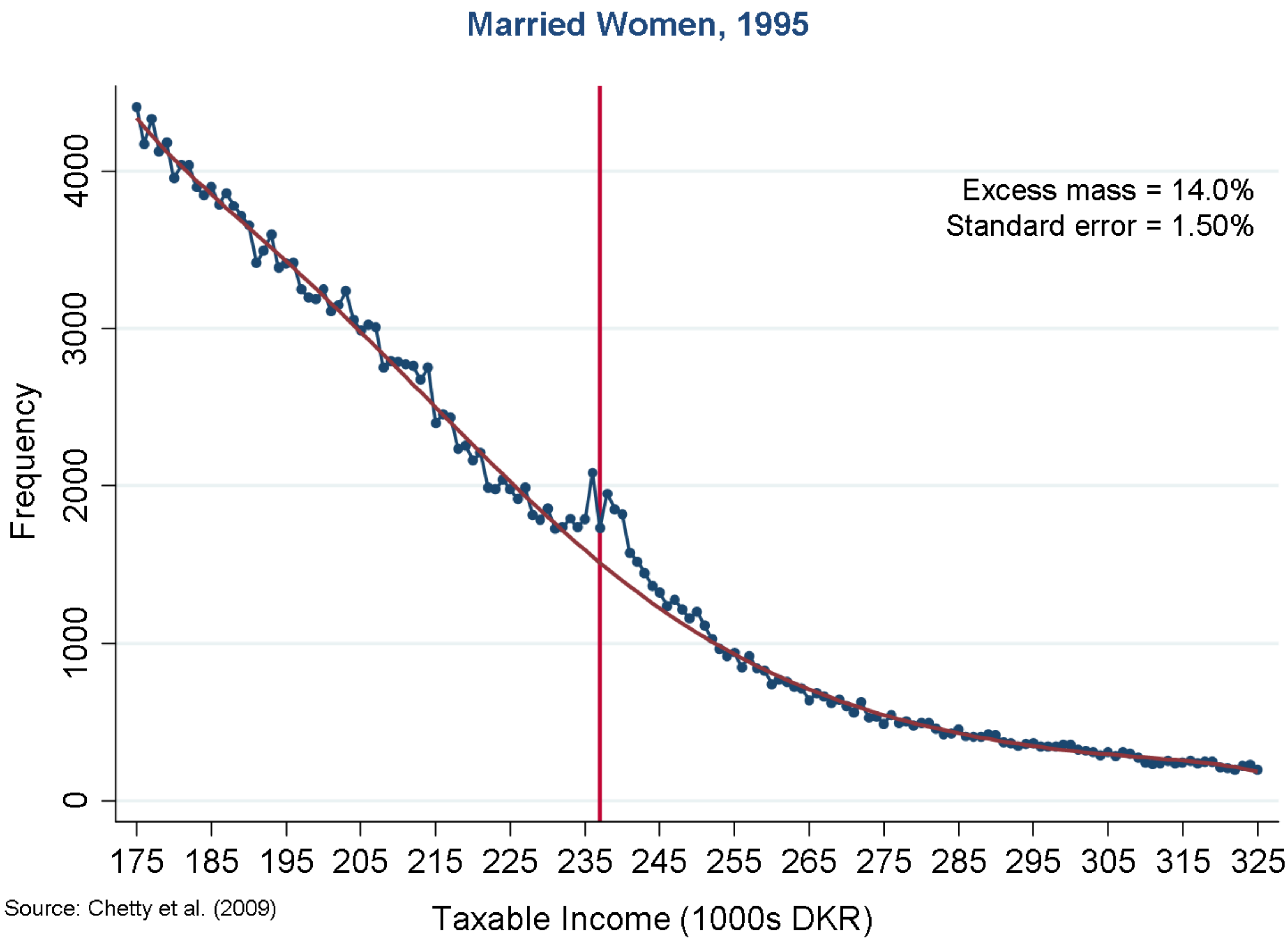



Military
Frequency
2000
4000
6000
8000
10000
Excess mass = 1.25%
Standard error = 0.61%
-50
-40
-30
-20
-10
0
10
20
30
40
50
Chetty et al. 2009
Taxable Income Relative to Top Bracket Cutoff (1000s DKr)

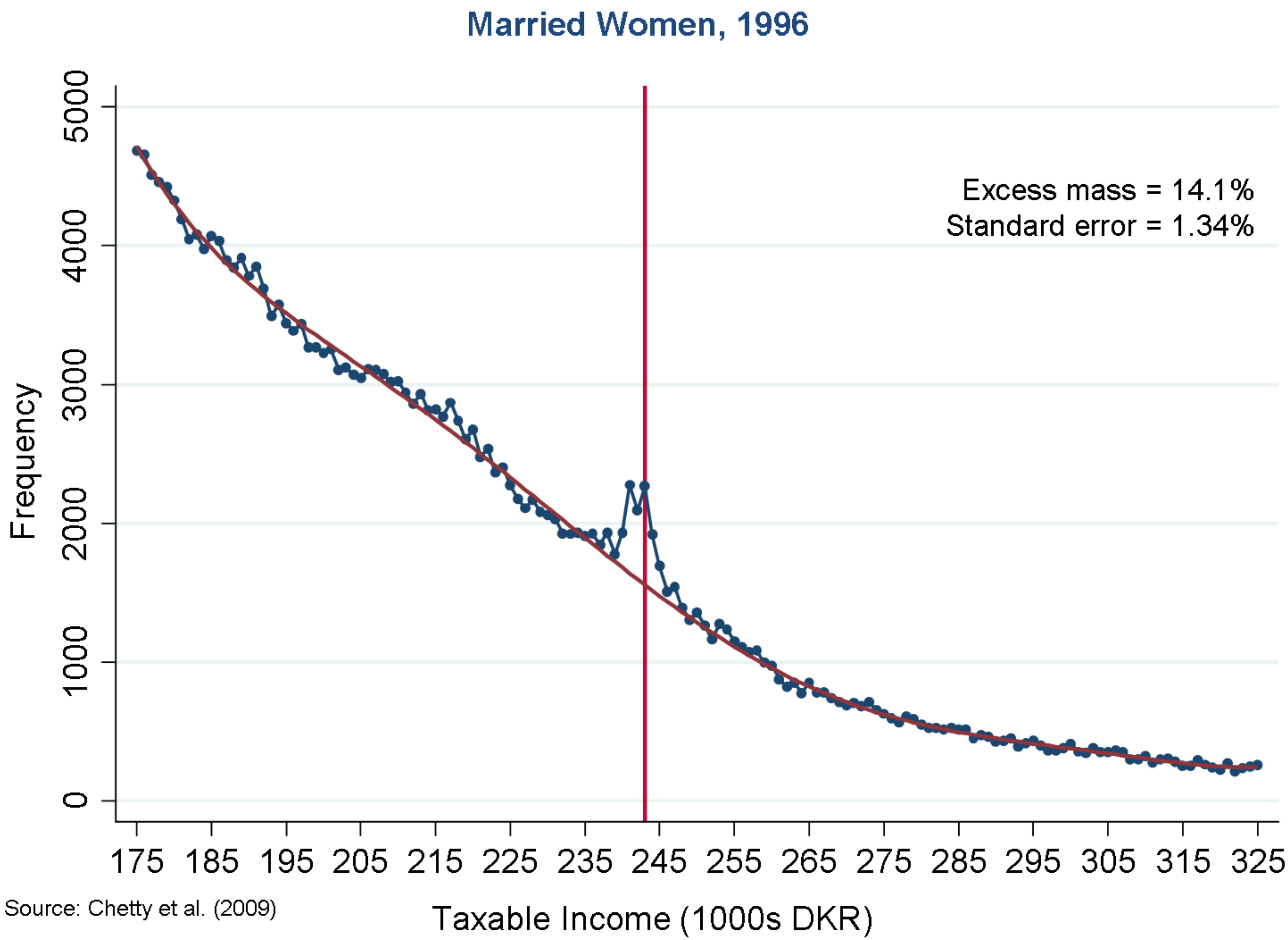
Married Women, 1994



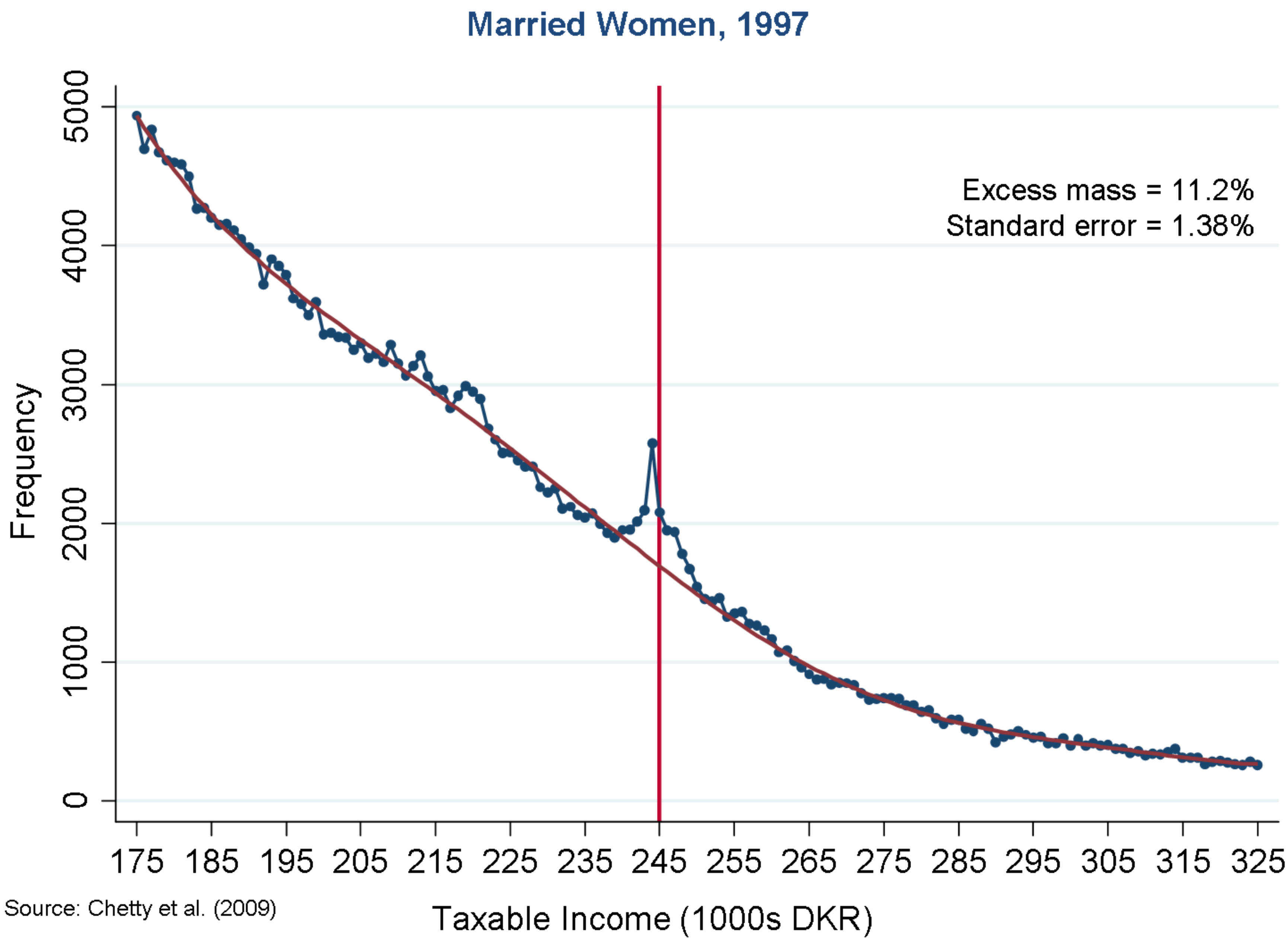
Source: Chetty et al. (2009)



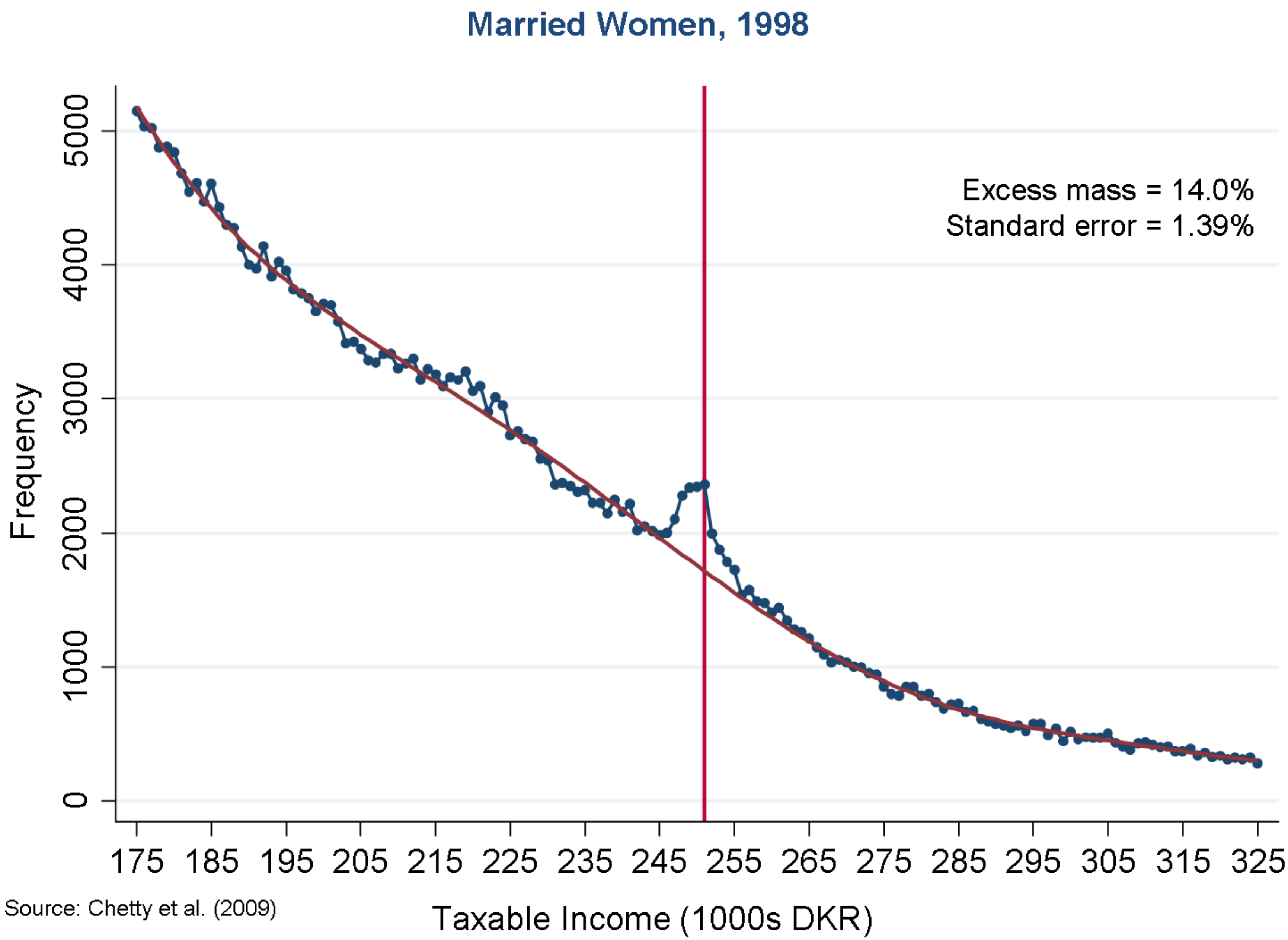
Source: Chetty et al. (2009)



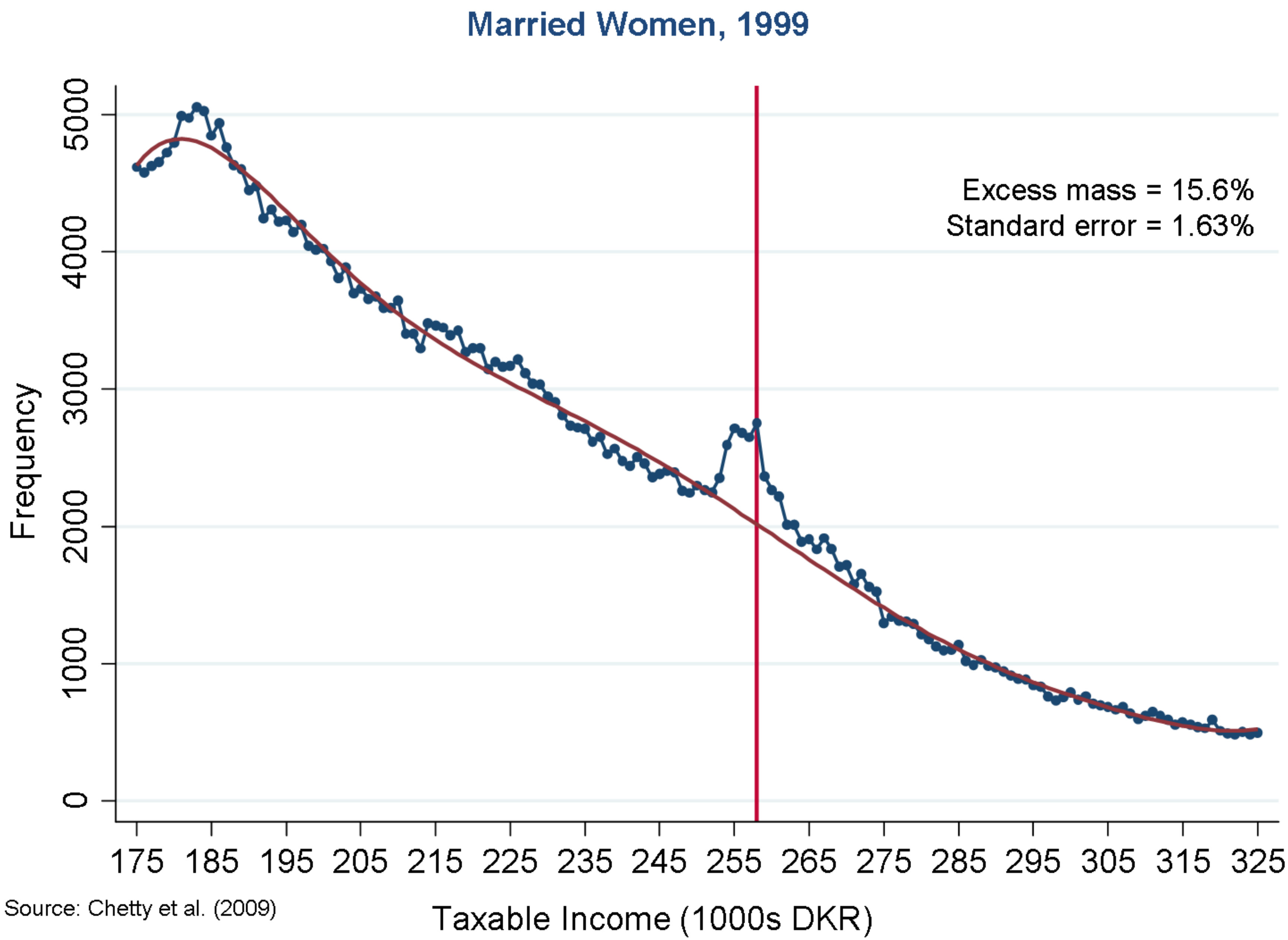
Source: Chetty et al. (2009)



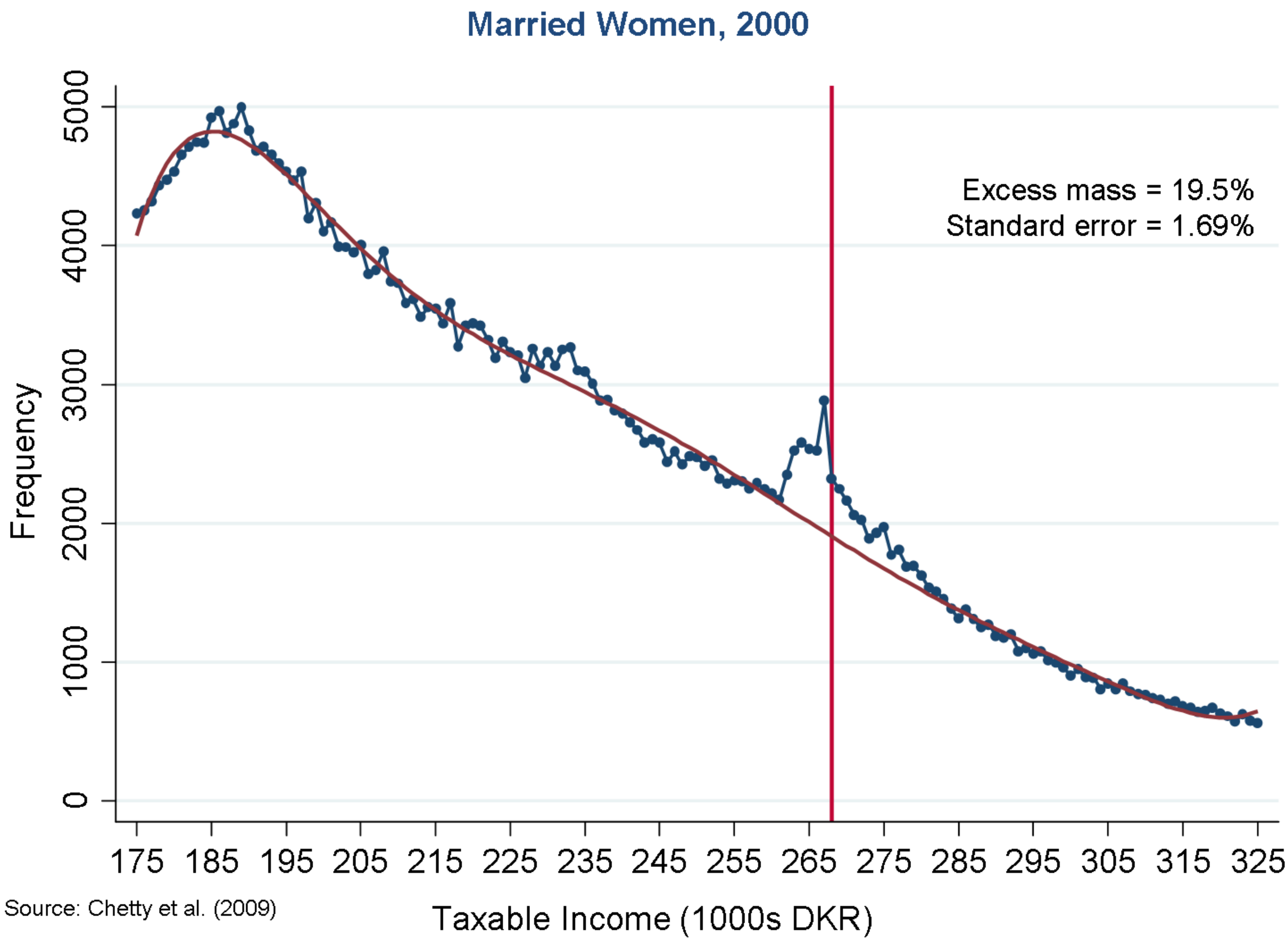
Source: Chetty et al. (2009)



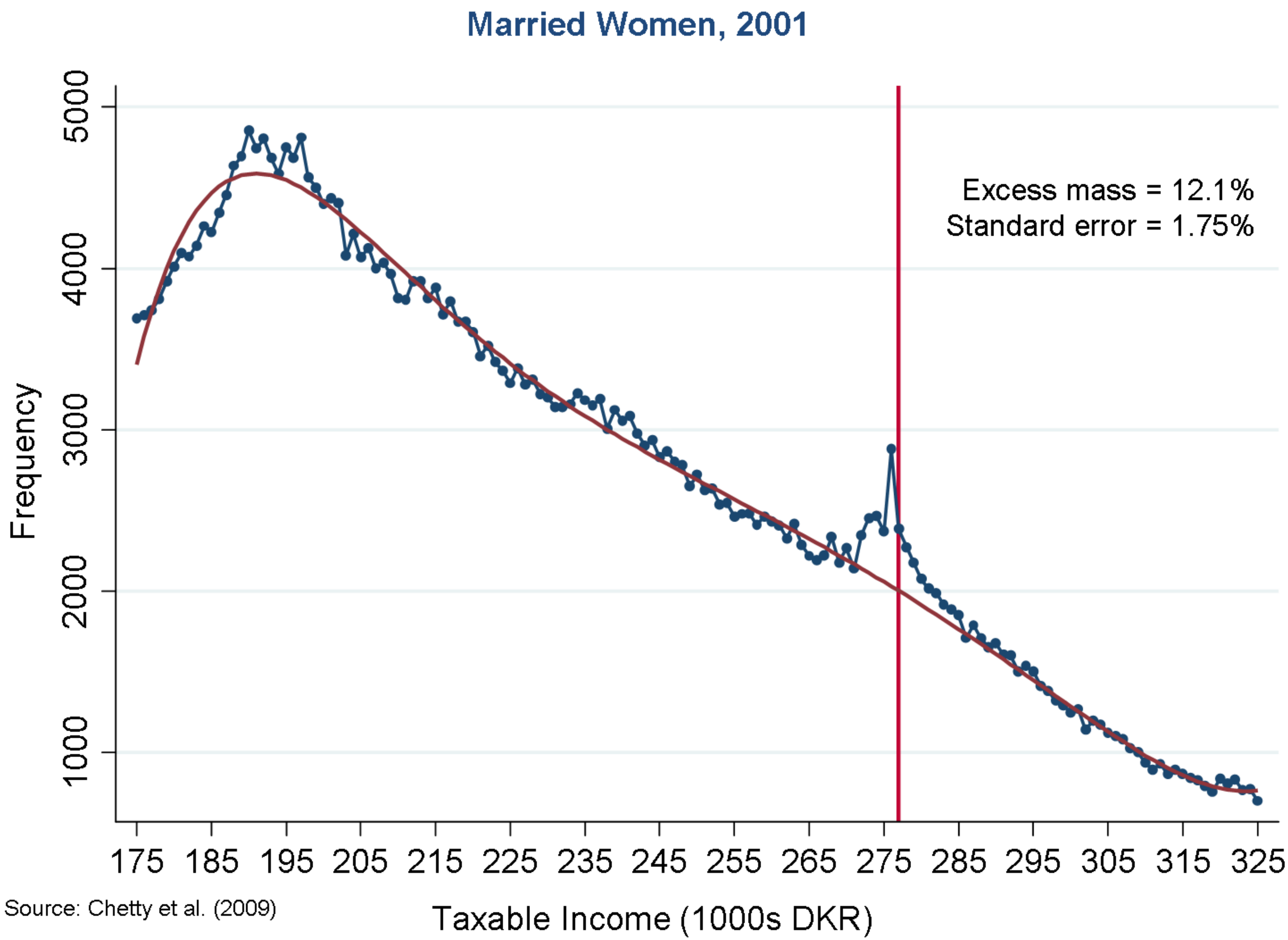
Source: Chetty et al. (2009)



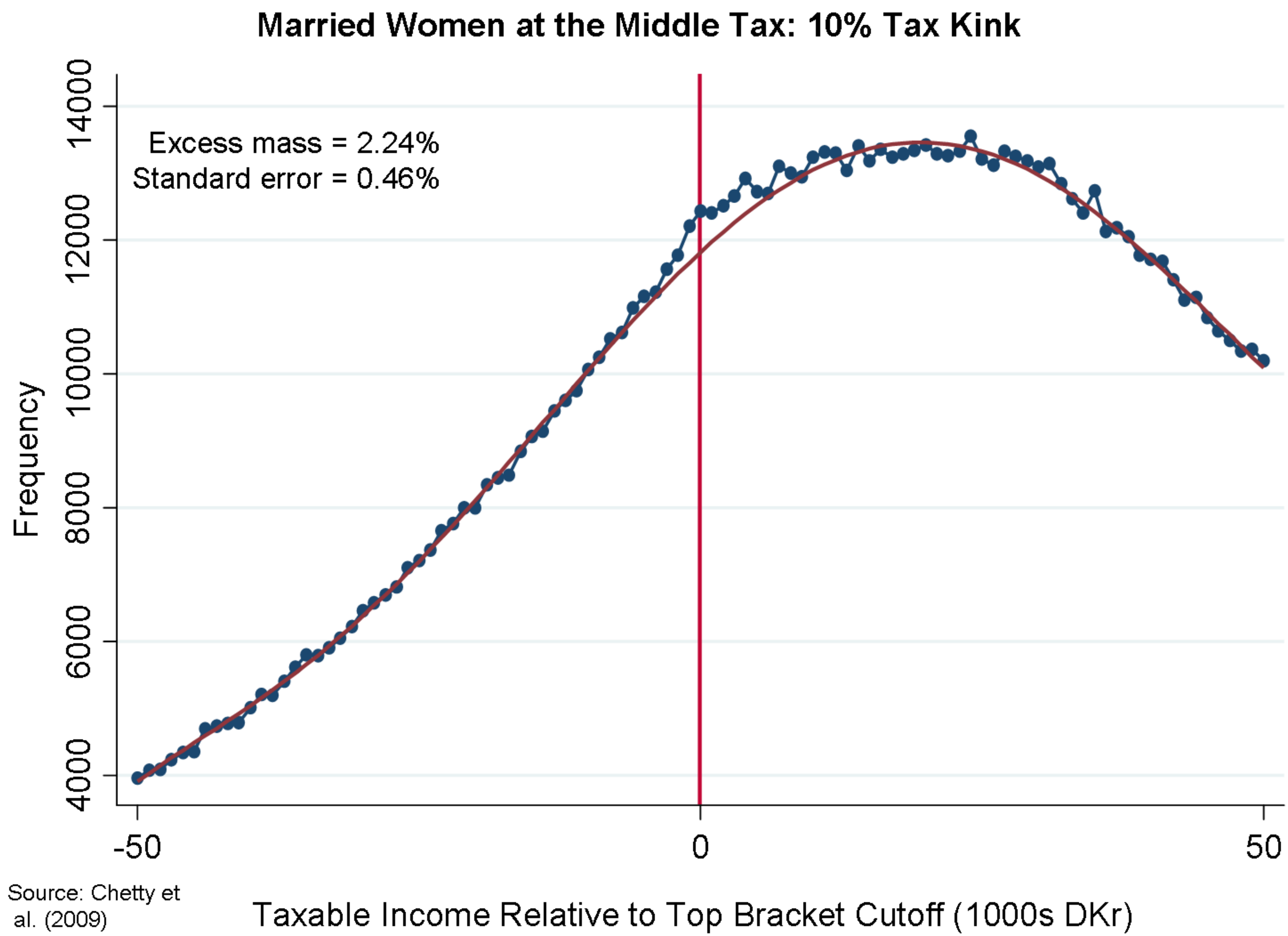
Source: Chetty et al. (2009)

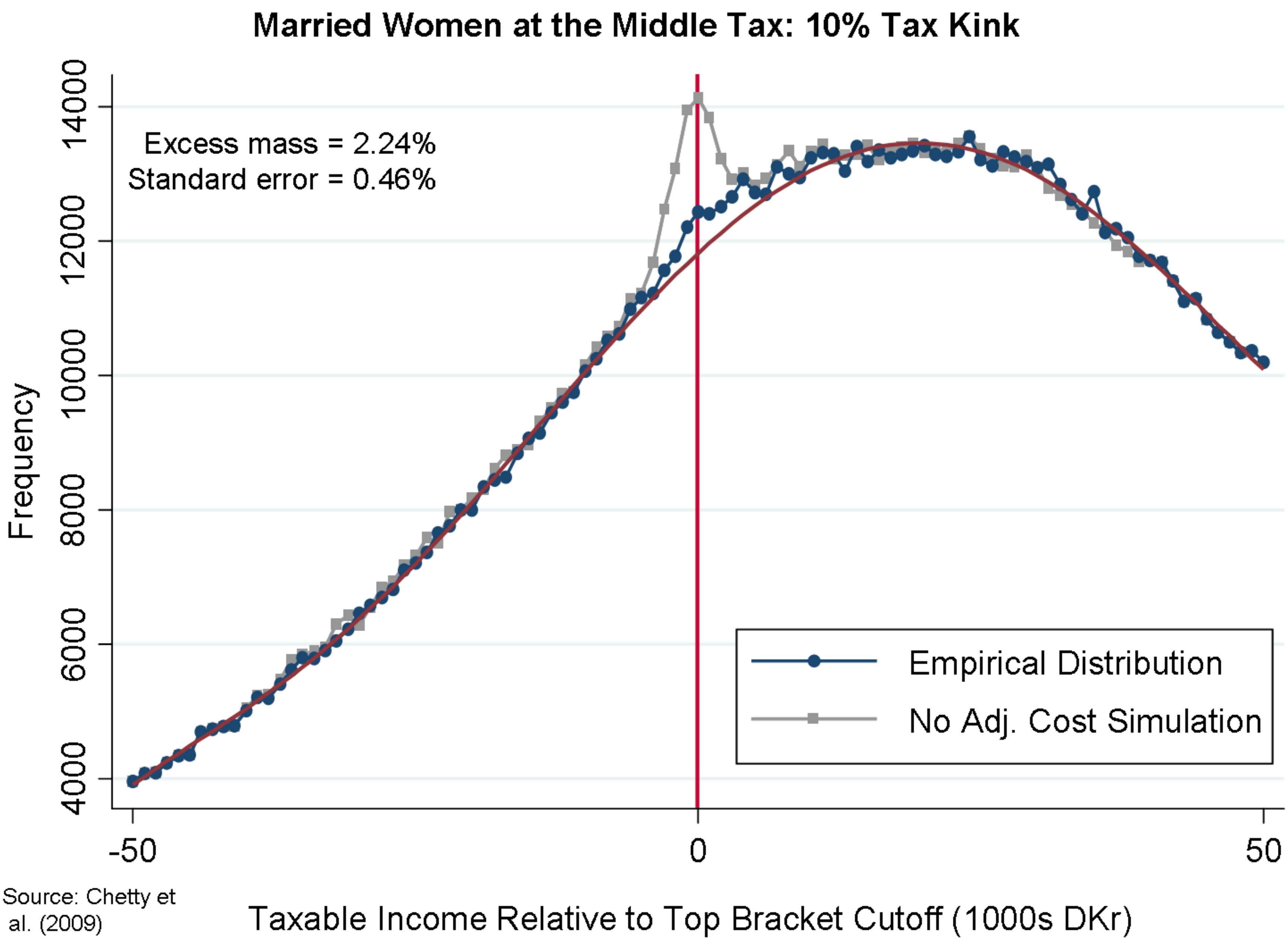


Married Women, 2000
Excess mass = 19.5%
Standard error = 1.69%
Frequency
0
1000
2000
3000
4000
5000
175 185 195 205 215 225 235 245 255 265 275 285 295 305 315 325
Taxable Income (1000s DKR)
Source: Chetty et al. (2009)



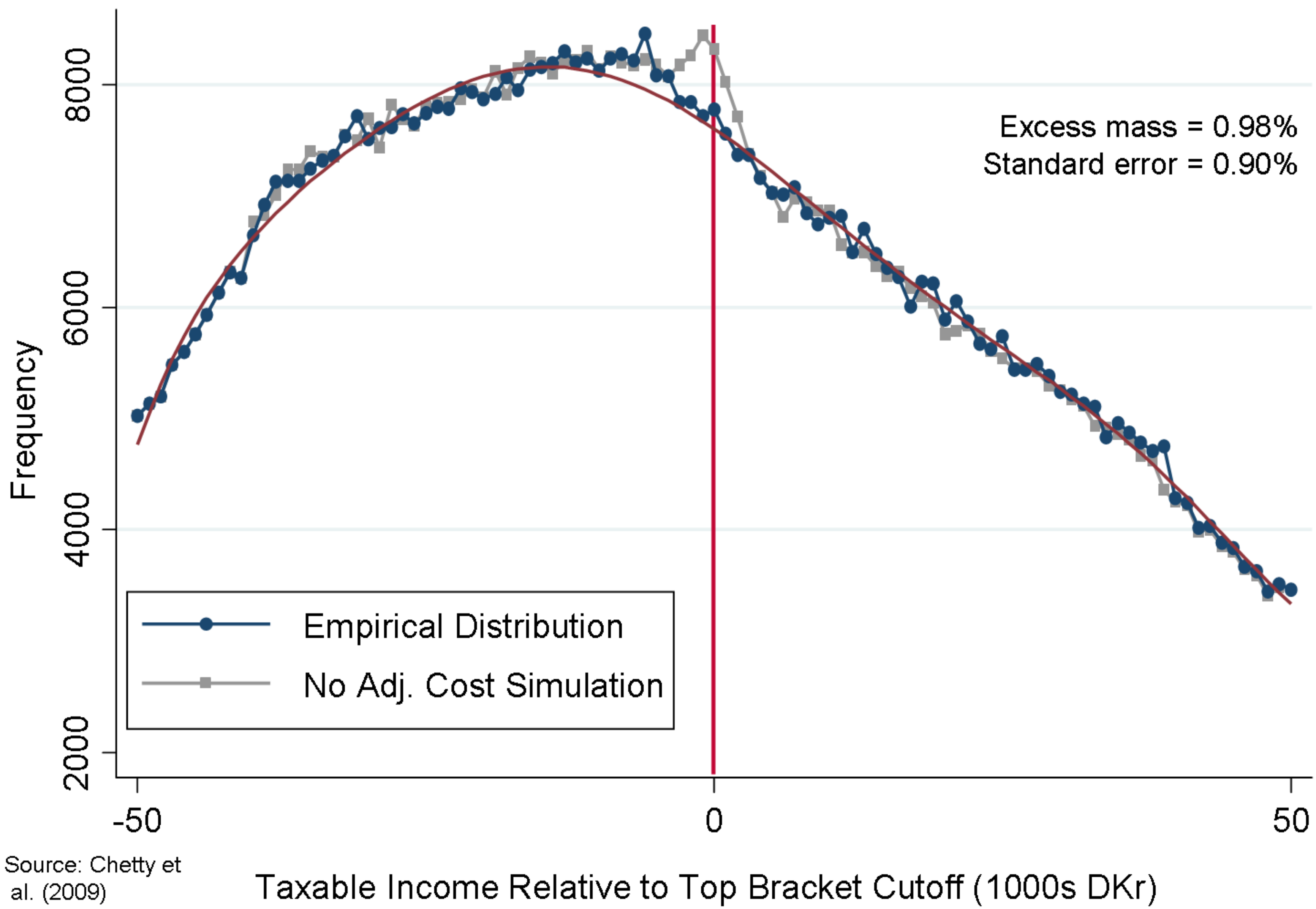
Source: Chetty et al. (2009)


Married Women at the Middle Tax: 10% Tax Kink
Excess mass = 2.24%
Standard error = 0.46%
Frequency
4000
6000
8000
10000
12000
14000
-50
0
50
Taxable Income Relative to Top Bracket Cutoff (1000s DKr)
Source: Chetty et al. (2009)

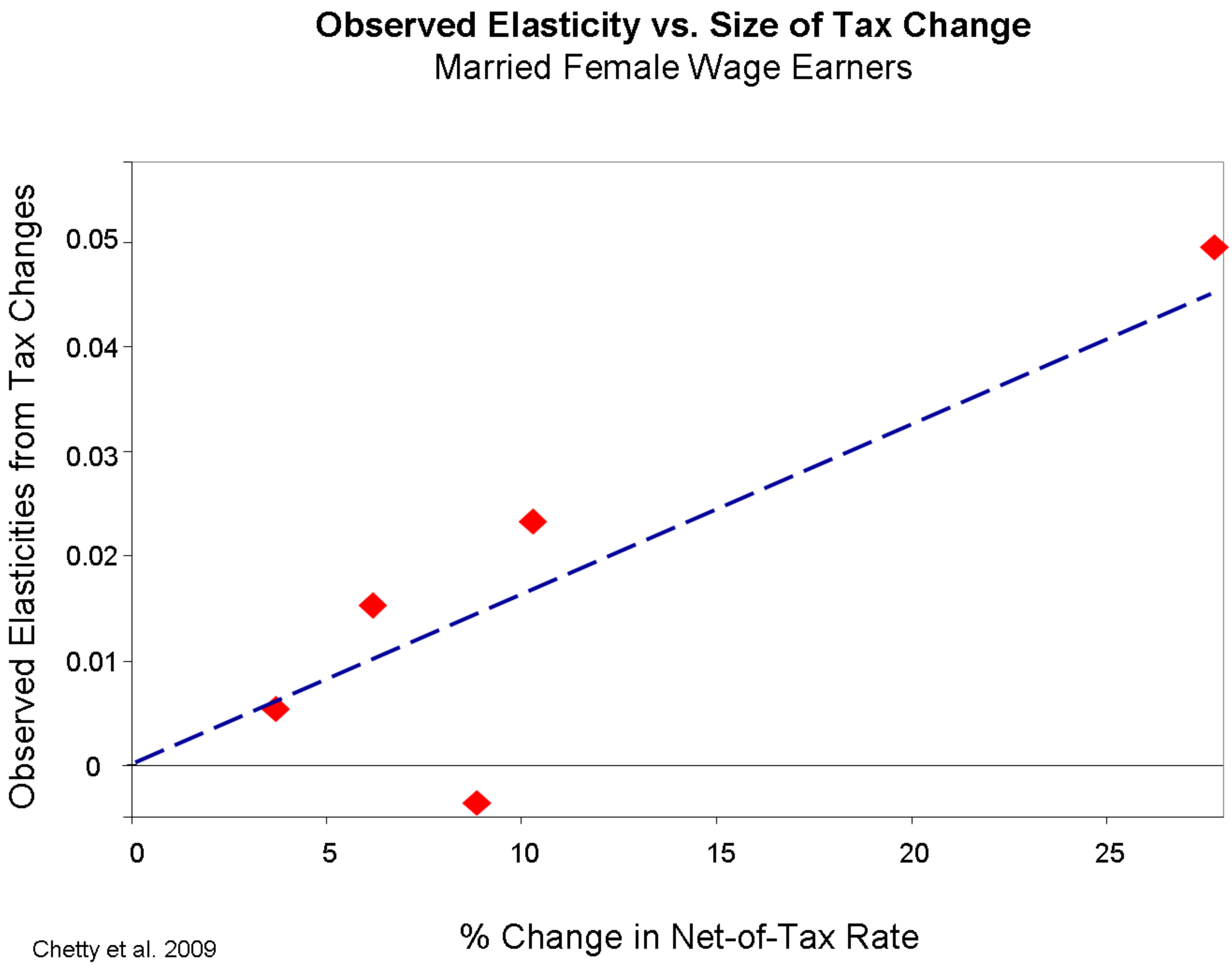


Source: Chetty et al. (2009)

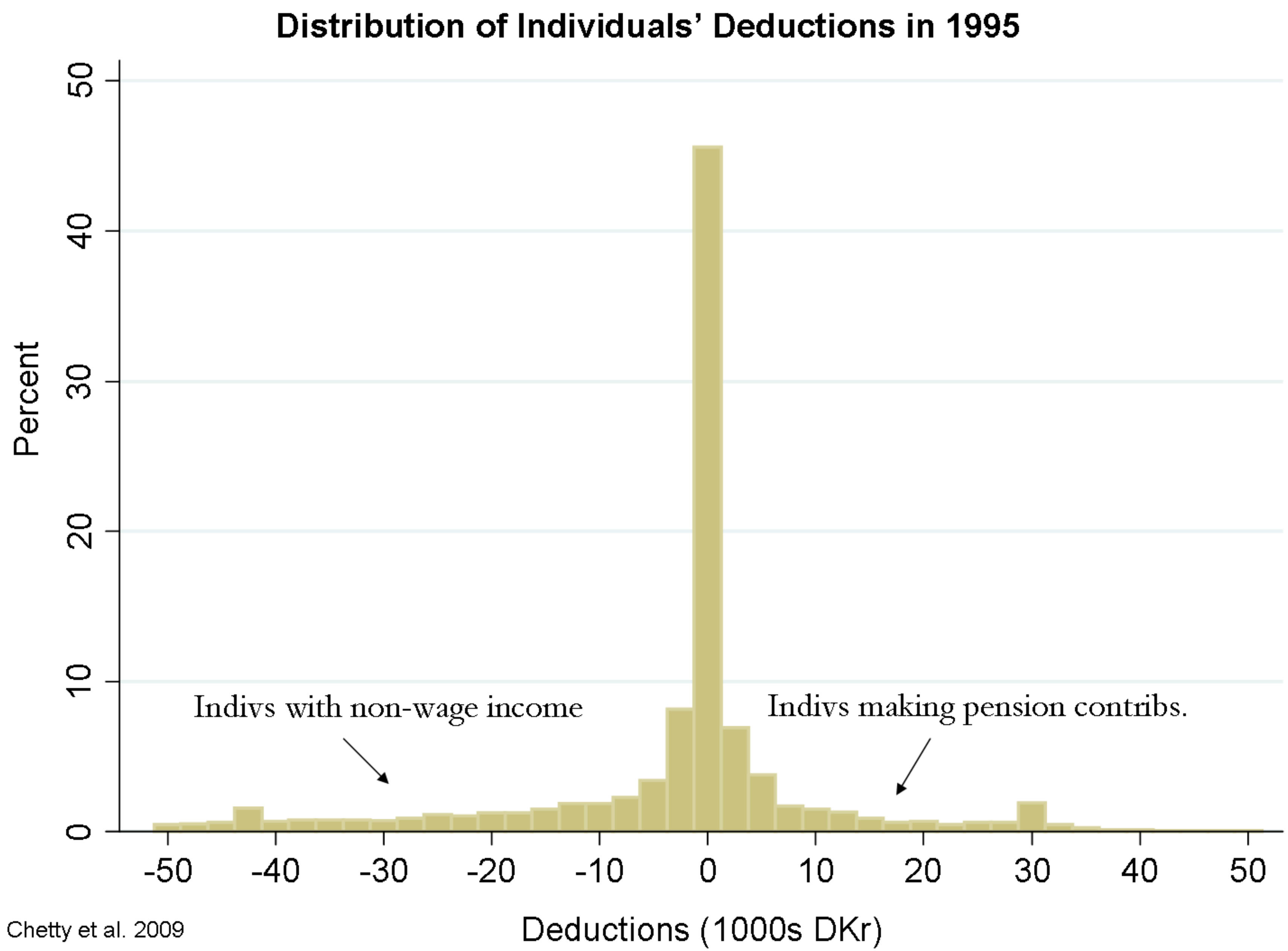
Married Women at the Middle Tax: 6% Tax Kink

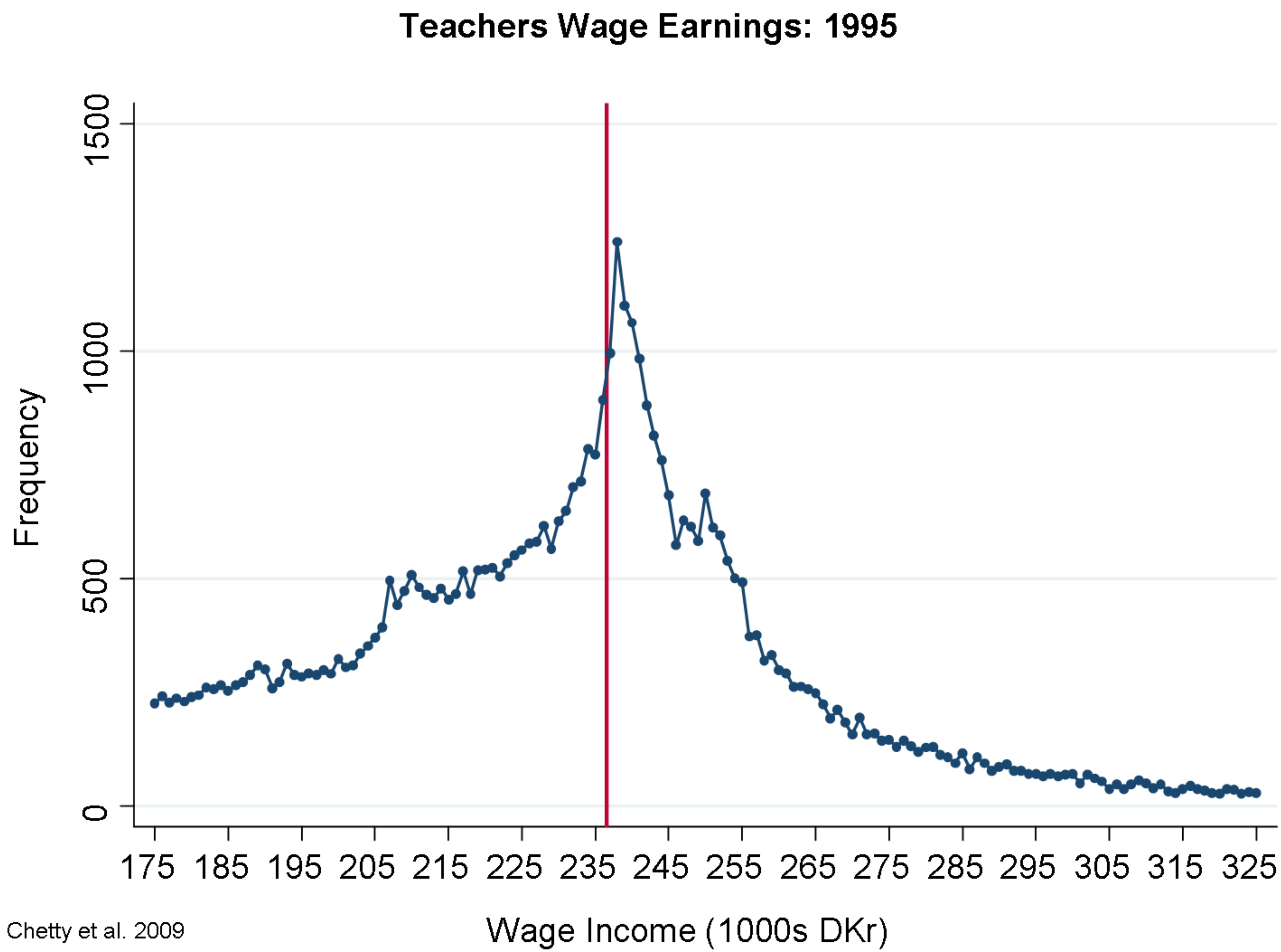


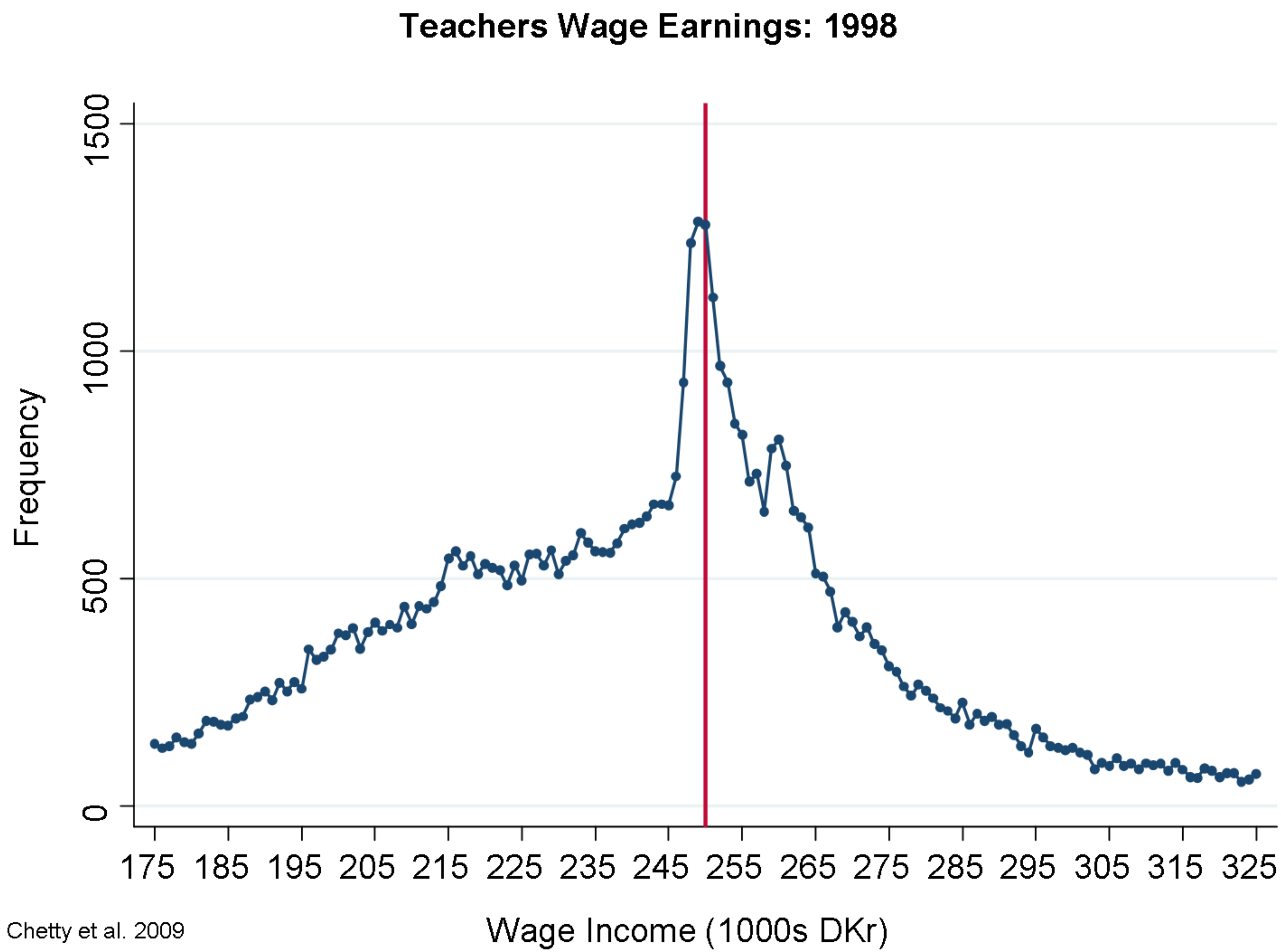
Source: Chetty et al. (2009)

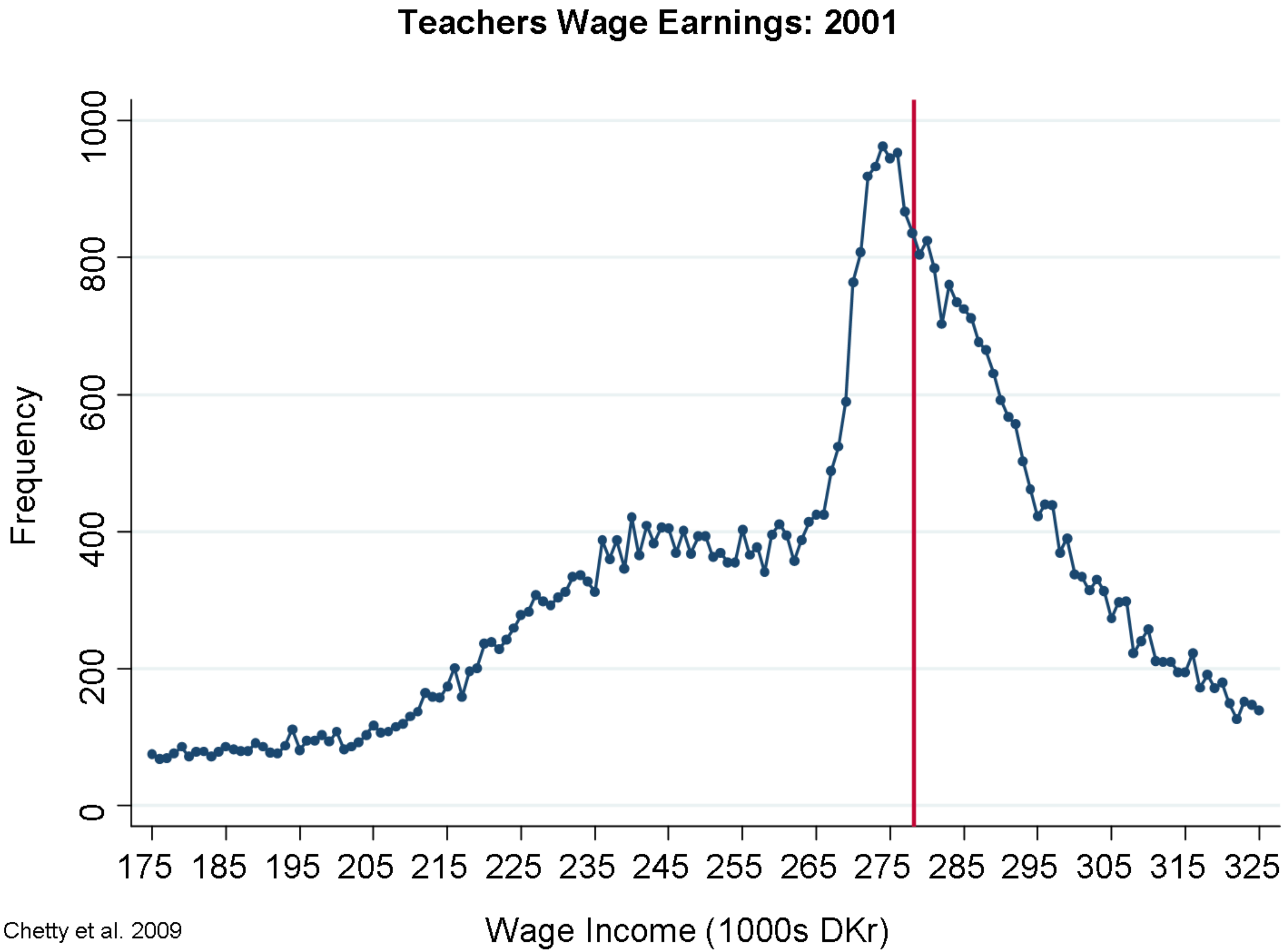


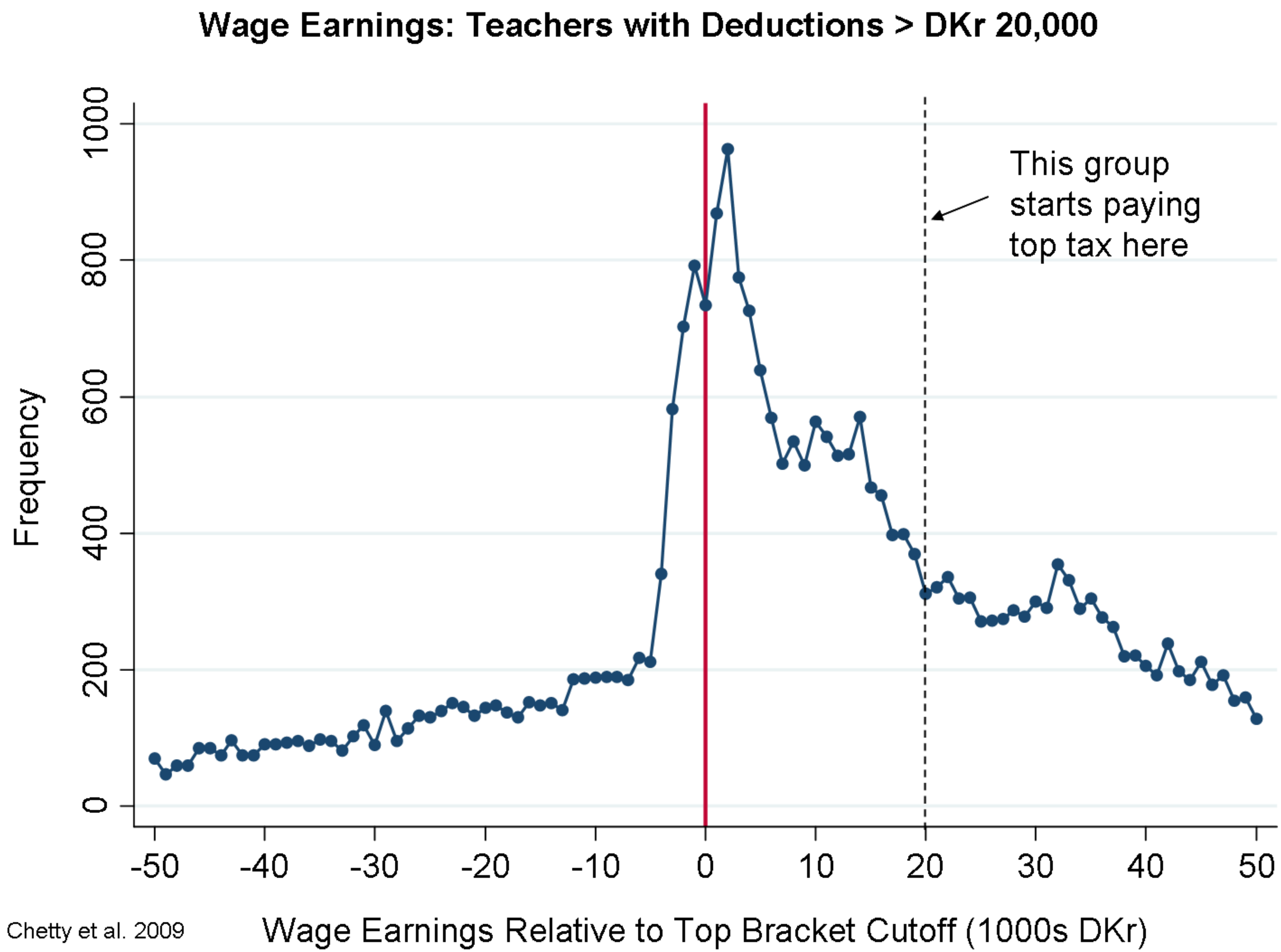
Observed Elasticity vs. Size of Tax Change
Married Female Wage Earners
Observed Elasticities from Tax Changes
0.05
0.04
0.03
0.02
0.01
0
0
5
10
15
20
25
% Change in Net-of-Tax Rate
Chetty et al. 2009

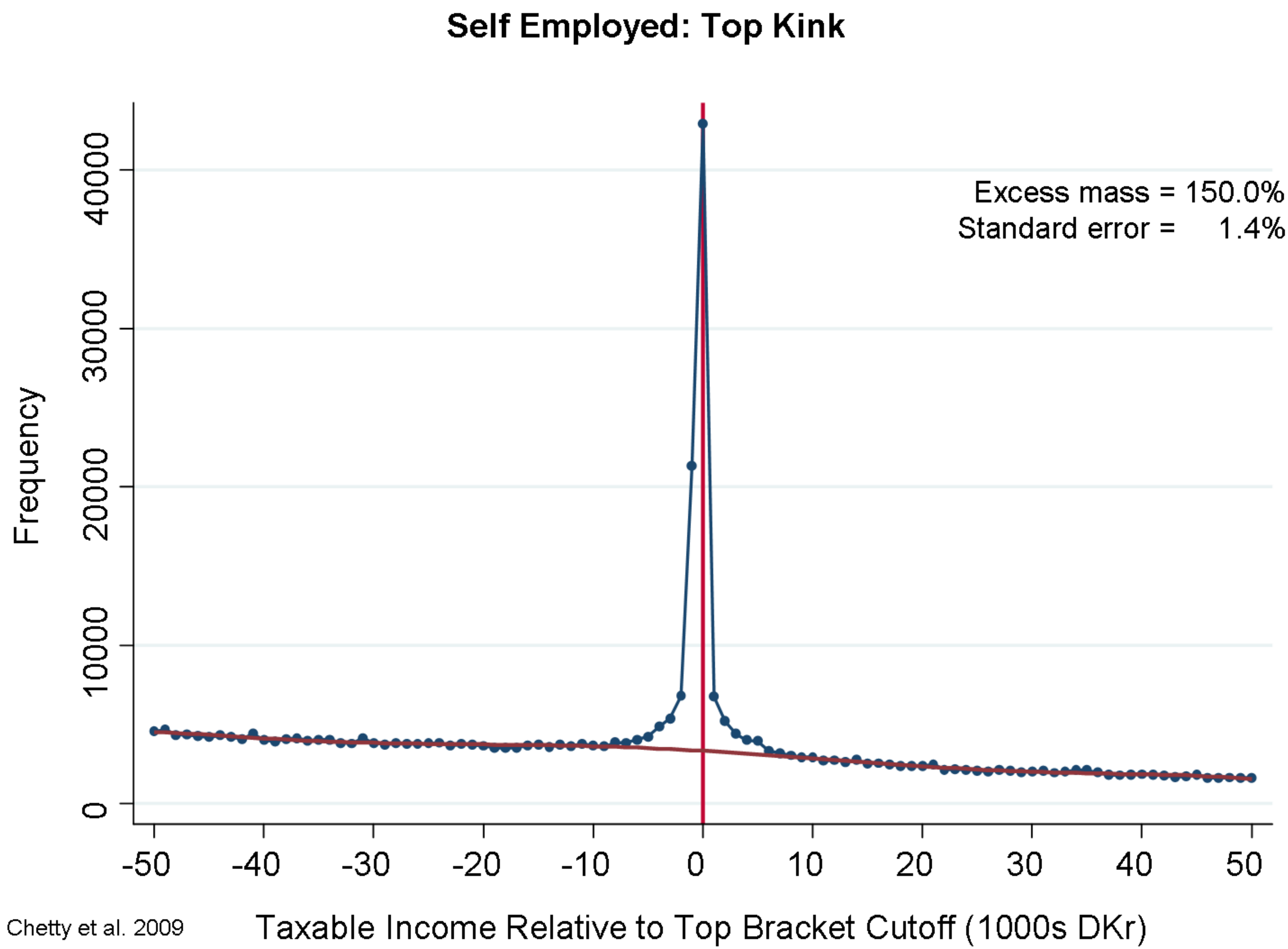

Distribution of Individuals' Deductions in 1995
Percent
0
10
20
30
40
50
Indivs with non-wage income
Indivs making pension contribs.
-50
-40
-30
-20
-10
0
10
20
30
40
50
Deductions (1000s DKr)
Chetty et al. 2009

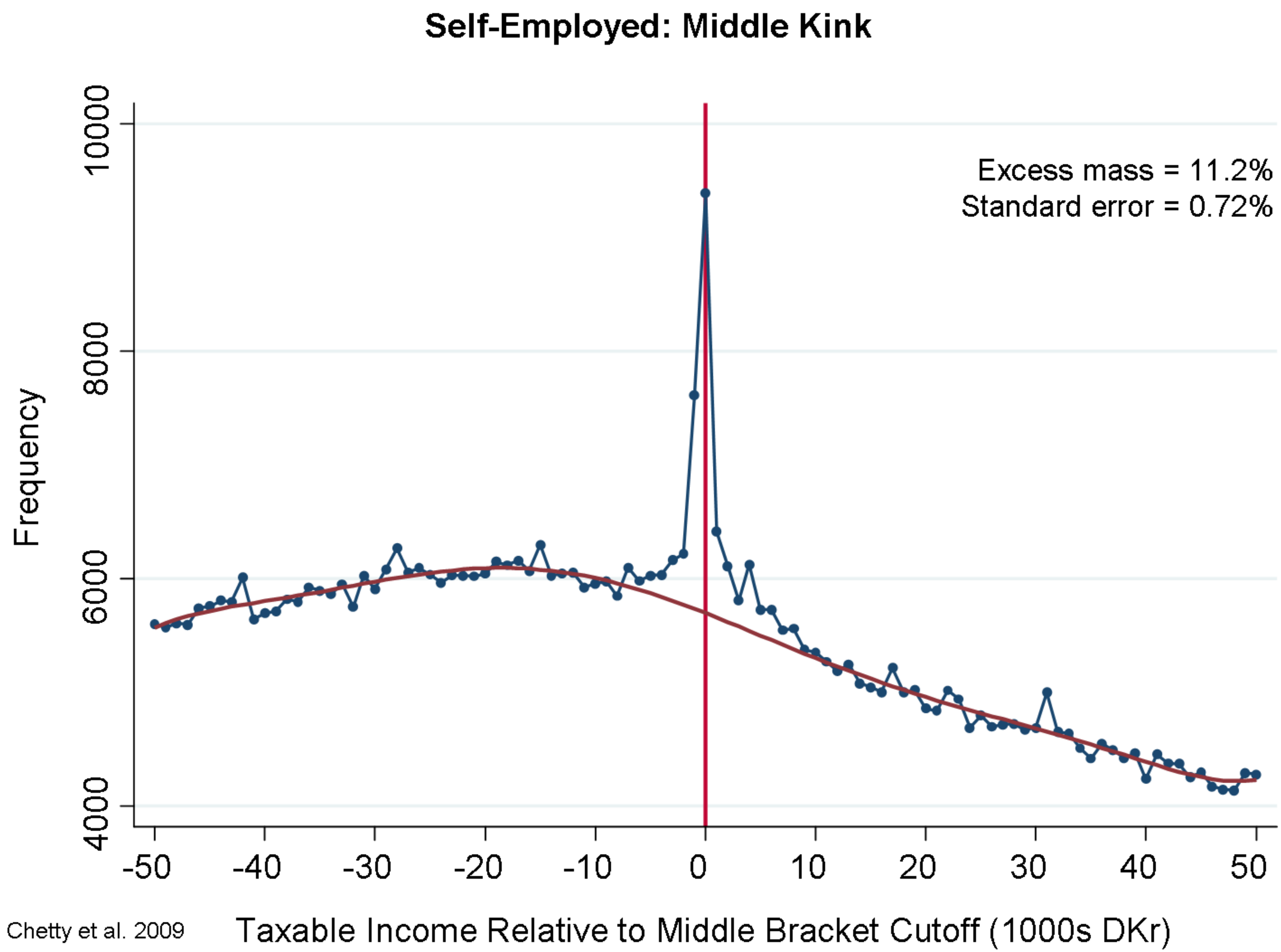

Teachers Wage Earnings: 1995
Frequency
0
500
1000
1500
175
185
195
205
215
225
235
245
255
265
275
285
295
305
315
325
Wage Income (1000s DKr)
Chetty et al. 2009

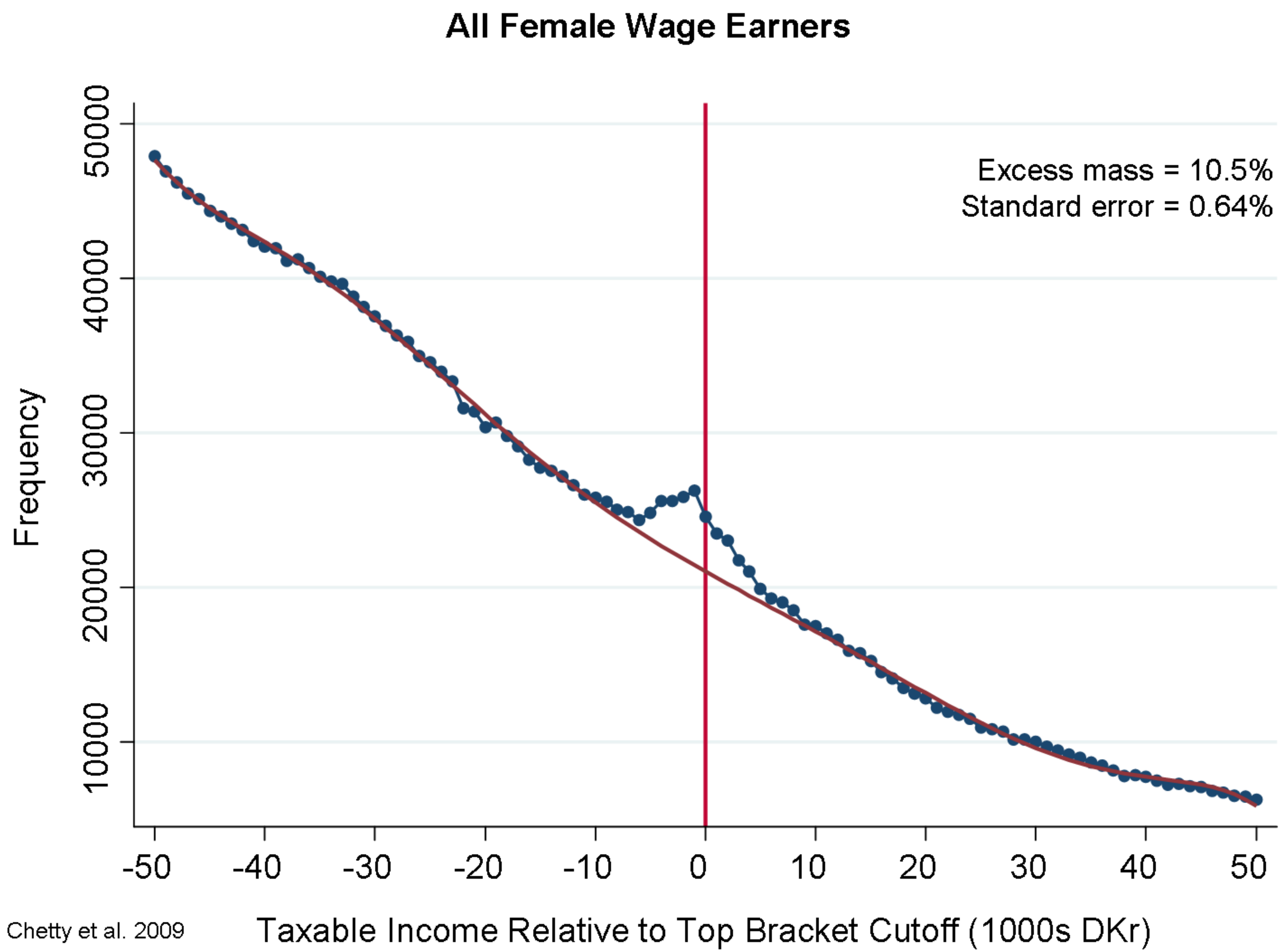

Teachers Wage Earnings: 1998
Frequency
0
500
1000
1500
175 185 195 205 215 225 235 245 255 265 275 285 295 305 315 325
Wage Income (1000s DKr)
Chetty et al. 2009


Teachers Wage Earnings: 2001
Frequency
0
200
400
600
800
1000
175 185 195 205 215 225 235 245 255 265 275 285 295 305 315 325
Wage Income (1000s DKr)
Chetty et al. 2009


Wage Earnings: Teachers with Deductions > DKr 20,000
This group
starts paying
top tax here
Frequency
0
200
400
600
800
1000
-50
-40
-30
-20
-10
0
10
20
30
40
50
Wage Earnings Relative to Top Bracket Cutoff (1000s DKr)
Chetty et al. 2009


Self Employed: Top Kink
Excess mass = 150.0%
Standard error = 1.4%
Frequency
0
10000
20000
30000
40000
-50
-40
-30
-20
-10
0
10
20
30
40
50
Taxable Income Relative to Top Bracket Cutoff (1000s DKr)
Chetty et al. 2009


Self-Employed: Middle Kink
Excess mass = 11.2%
Standard error = 0.72%
Frequency
10000
8000
6000
4000
-50
-40
-30
-20
-10
0
10
20
30
40
50
Taxable Income Relative to Middle Bracket Cutoff (1000s DKr)
Chetty et al. 2009


All Female Wage Earners
Excess mass = 10.5%
Standard error = 0.64%
Frequency
10000
20000
30000
40000
50000
-50
-40
-30
-20
-10
0
10
20
30
40
50
Chetty et al. 2009
Taxable Income Relative to Top Bracket Cutoff (1000s DKr)


All Male Wage Earners
Excess mass = 3.26%
Standard error = 0.29%
Frequency
10000
20000
30000
40000
50000
-50
-40
-30
-20
-10
0
10
20
30
40
50
Chetty et al. 2009
Taxable Income Relative to Top Bracket Cutoff (1000s DKr)

Chetty et al. 2009: Results

1) Search costs attenuate observed behavioral responses substantially

2) Firm responses and coordination critical for understanding behavior: individual and group elasticities may differ significantly

3) NLBS models may fit data better if these factors are incorporated

4) Standard method of estimating elasticities using small tax reforms on same data yields close-to-zero elasticity estimate

Natural Experiment Labor Supply Literature

Literature exploits variation in taxes/transfers to estimate Hours and Participation Elasticities

1) Return to simple model where we ignore non-linear budget set issues

2) Large literature in labor/Public economics estimates effects of taxes and wages on hours worked and participation

3) Now discuss some estimates from this older literature

Negative Income Tax (NIT) Experiments

1) Best way to resolve identification problems: exogenously increase the marginal tax rate 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 τ (50%-80%) [based on family earnings]

5) Analysis by Rees (1974), Munnell (1986) book, Ashenfelter and Plant JOLE'90, and others

6) Several groups, with randomization within each; approx. N = 75 households in each group

Table 1
Parameters of the 11 Negative Income Tax Programs

Program Number	G ($)	τ	Declining Tax Rate	Break-even Income ($)
1	3,800	.5	No	7,600
2	3,800	.7	No	5,429
3	3,800	.7	Yes	7,367
4	3,800	.8	Yes	5,802
5	4,800	.5	No	9,600
6	4,800	.7	No	6,857
7	4,800	.7	Yes	12,000
8	4,800	.8	Yes	8,000
9	5,600	.5	No	11,200
10	5,600	.7	No	8,000
11	5,600	.8	Yes	10,360

Source: Ashenfelter and Plant (1990)

NIT Experiments: Ashenfelter and Plant 1990

1) Present non-parametric evidence of labor supply effects

2) Compare implied benefit payments to treated vs. control households

3) Difference in benefit payments reflects aggregates hours and participation responses

4) This is the relevant parameter for expenditure calculations and potentially for welfare analysis (revenue method of calculating DWL)

5) Shortcoming: approach does not decompose estimates into income and substitution effects and intensive vs. extensive margin

6) Hard to identify the key elasticity relevant for policy purposes and predict labor supply effect of other programs

Table 3
Experimental Payment minus Predicted Control Payment for 3-Year Dual-headed Experimental Families, Attrition Families Excluded (Standard Errors in Parentheses)

G ($)	τ	Declining Tax Rate	Preexperimental Payment ($)	Payments for Year of Experiment ($) 1	2	3	Postexperimental Payment ($)
3,800	.5	No	193.78	248.46	368.95*	389.24*	138.56
			(143.45)	(149.58)	(170.75)	(182.99)	(188.20)
3,800	.7	No	124.96	185.18	317.28	218.37	−47.85
			(223.77)	(237.91)	(252.99)	(325.57)	(314.66)
3,800	.7	Yes	−33.37	68.94	158.44	324.84	29.28
			(178.05)	(176.07)	(213.59)	(230.50)	(222.42)
3,800	.8	Yes	75.40	336.06	221.54	160.83	91.52
			(229.44)	(237.18)	(245.92)	(264.53)	(261.84)
4,800	.5	No	52.02	85.17	294.55	337.23	70.22
			(192.31)	(184.85)	(201.73)	(221.73)	(219.58)
4,800	.7	No	220.76	288.33	496.85*	543.25*	178.32
			(160.04)	(169.04)	(197.88)	(204.50)	(194.03)
4,800	.7	Yes	136.99	281.98*	423.30*	348.03*	23.96
			(127.36)	(137.19)	(157.51)	(162.38)	(140.58)
4,800	.8	Yes	−16.87	305.09	417.90	317.39	121.47
			(175.54)	(209.24)	(234.32)	(274.11)	(239.59)
5,600	.5	No	−163.12	200.75	664.41*	717.15*	124.93
			(252.05)	(258.13)	(283.28)	(280.65)	(287.04)
5,600	.7	No	−59.97	23.34	386.12	744.94*	267.69
			(164.95)	(156.41)	(200.59)	(263.80)	(259.45)
5,600	.8	Yes	−27.64	−51.03	117.85	273.44	121.53
			(121.47)	(126.67)	(138.52)	(157.96)	(169.26)

NOTE.—Terms are explained in text.
* Denotes mean is more than twice its standard error.

Table 4
Experimental Payment minus Predicted Control Payment for 5-Year Dual-headed Experimental Families, Attrition Families Excluded (Standard Errors in Parentheses)

G ($)	τ	Declining Tax Rate	Preexperimental Payment ($)	Payment for Year of Experiment ($)					Postexperimenta Payment ($)
				1	2	3	4	5	
3,800	.5	No	102.24	345.68	526.02	110.30	390.07	169.82	229.70
			(185.55)	(221.42)	(241.53)	(265.28)	(307.01)	(286.76)	(309.06)
3,800	.7	No	81.16	23.30	−99.33	98.20	−16.42	−122.01	−406.46
			(309.85)	(316.06)	(330.14)	(383.52)	(388.07)	(352.95)	(314.40)
3,800	.7	Yes	6.99	490.00	176.14	23.22	324.70	−59.79	−598.09*
			(234.01)	(288.13)	(272.87)	(300.28)	(386.93)	(331.68)	(102.72)
3,800	.8	Yes	−130.30	349.73	189.80	329.94	1207.82*	1108.49*	307.38
			(271.23)	(286.56)	(280.63)	(365.58)	(463.10)	(487.83)	(453.29)
4,800	.5	No	−23.66	30.15	160.40	399.28	419.73	434.30	251.09
			(183.73)	(208.90)	(199.26)	(236.33)	(247.25)	(254.52)	(242.45)
4,800	.7	No	−129.98	25.71	−4.47	569.10	493.42	219.74	−38.46
			(185.46)	(208.14)	(211.44)	(314.73)	(357.32)	(340.60)	(228.01)
4,800	.7	Yes	75.66	224.96	387.66	340.71	−130.10	34.61	189.49
			(234.21)	(280.43)	(367.56)	(404.05)	(308.90)	(445.67)	(491.52)
4,800	.8	Yes	467.89	325.17	599.43*	398.62	537.21	506.95	346.28
			(252.40)	(276.31)	(274.39)	(280.50)	(365.56)	(351.98)	(337.43)
5,600	.5	No	−224.97	560.51	723.08*	782.53*	592.40	313.82	−53.07
			(286.39)	(298.21)	(306.90)	(327.39)	(366.88)	(387.31)	(325.66)
5,600	.7	No	−158.74	500.18	1194.68*	890.38*	825.39	435.01	588.91
			(239.17)	(311.24)	(416.25)	(391.61)	(467.76)	(609.49)	(510.52)
5,600	.8	Yes	−6.48	193.54	617.29*	906.13*	888.72	877.71	75.21
			(175.15)	(199.51)	(255.89)	(315.98)	(337.38)	(398.38)	(216.12)

NOTE.—Terms are explained in text.
* Denotes mean is more than twice its standard error.

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 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

Problems with NIT Experimental Design

Estimates from NIT not considered fully credible due to several shortcomings:

1) **Self reported earnings:** Treatments had financial incentives to under-report earnings $\Rightarrow$ Lesson: need to match with administrative records [Greenberg and Halsey JOLE'83]

2) **Selective attrition:**

After initial year, data collected based on voluntary income reports by families $\Rightarrow$ Those in less generous groups/far above break-even point had much less incentive to report $\Rightarrow$ Attrition rates higher in these groups $\Rightarrow$ No longer a random sample of treatment + controls [Ashenfelter-Plant JOLE'90]

3) Response might be smaller than real reform b/c of **General Equilibrium** effects

Social Experiments: Costs/Benefits

1) Cost of NIT experiments: around $1 billion (in today's dollars)

2) Huge cost for a social experiment but trivial relative to budget of the US federal government ($3 trillion)

3) Should the government do more experimentation? Potential benefits:

a) Narrow the standard error around estimates

b) Allow implementation of better tax and redistribution policy

[Literature on optimal experimenting in engineering and agriculture but never applied to economics, pb. is instability of parameters]

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 $\Rightarrow$ "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

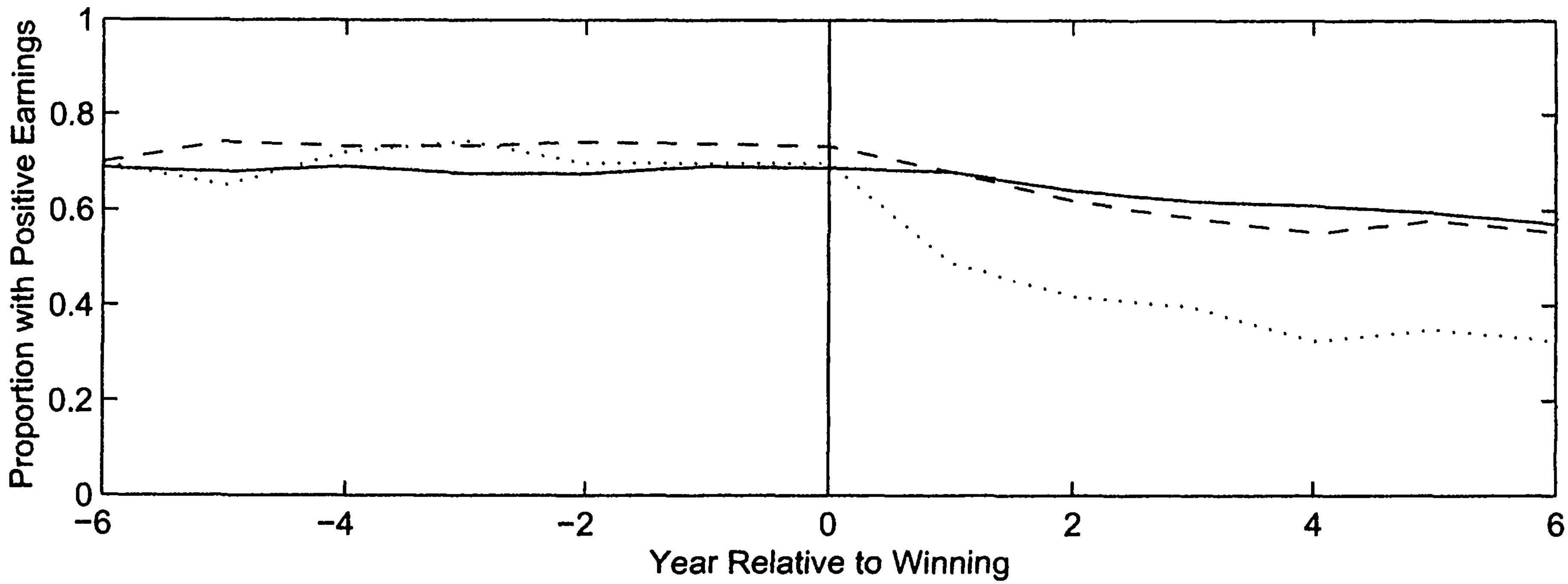


FIGURE 2. PROPORTION WITH POSITIVE EARNINGS FOR NONWINNERS, WINNERS, AND BIG WINNERS

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

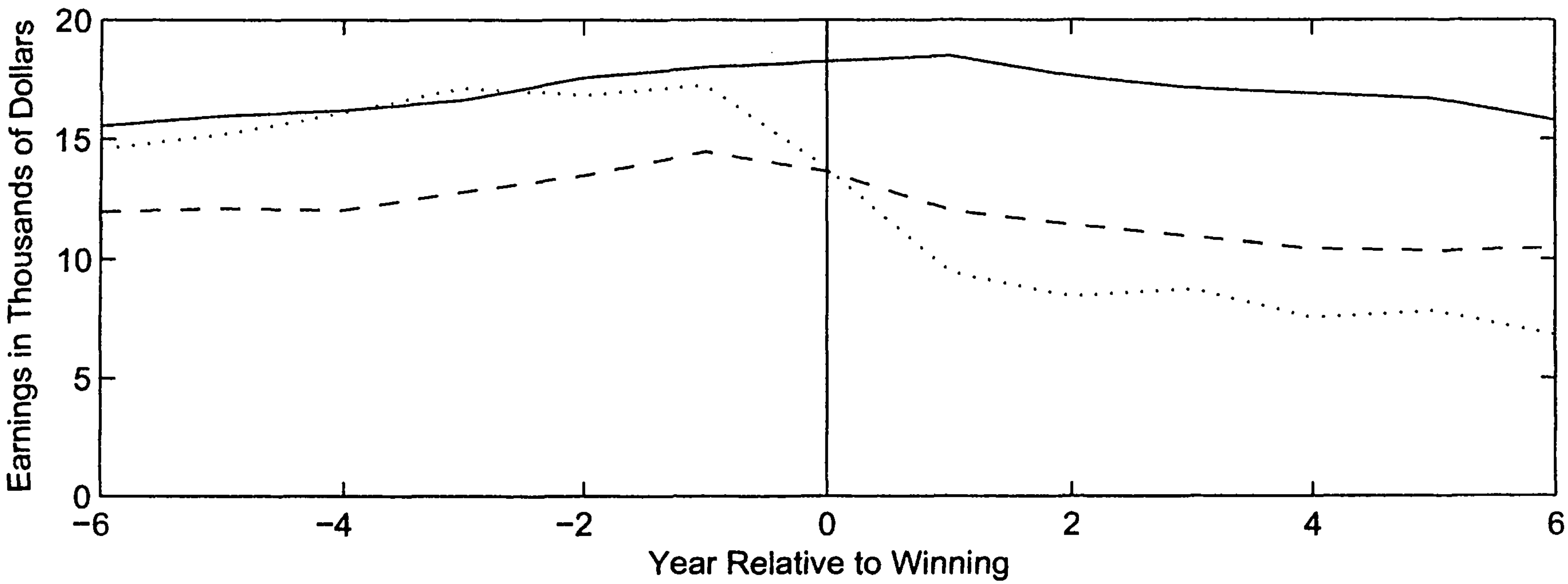


FIGURE 1. AVERAGE EARNINGS FOR NONWINNERS, WINNERS, AND BIG WINNERS

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

Instrumental Variable Methods

1) Another strategy to overcome endogeneity is instrumenting for wage rate

2) Mroz (1987): often-cited survey/meta-analysis of earlier studies

3) Uses PSID to test widely-used IV's for married women's wage

$$
\begin{aligned}
l_i &= \alpha + \beta w + \gamma X + \varepsilon \\
w &= \theta Z + \mu
\end{aligned}
$$

4) Uses Hausman specification/overidentification test to show that many instruments violate $EZ\varepsilon = 0$

Hausman Test

1) Suppose you can divide instrument set into those that are credibly exogenous (Z) and those that are questionable (Z^*)

2) Null hypothesis: both are exogenous

3) Alternative hypothesis: Z^* is endogenous

4) Compute IV estimate of β with small and large instrument set and test for equality of the coefficients

5) Note that is often a very low power test (accept validity if instruments are weak)

Mroz 1987: Setup and Results

1) Uses background variables as “credibly exogenous” instruments [Parents’ education, age, education polynomials]

2) Tests validity of labor market experience, average hourly earnings, and previous reported wages

3) Rejects validity of all three

4) Shows that earlier estimates are highly fragile and unreliable

5) Contributed to emerging view that policy variation (e.g., taxes) was necessary to really identify these elasticities properly

Tax Reform Variation (Eissa 1995)

1) Modern studies use tax changes as "natural experiments"

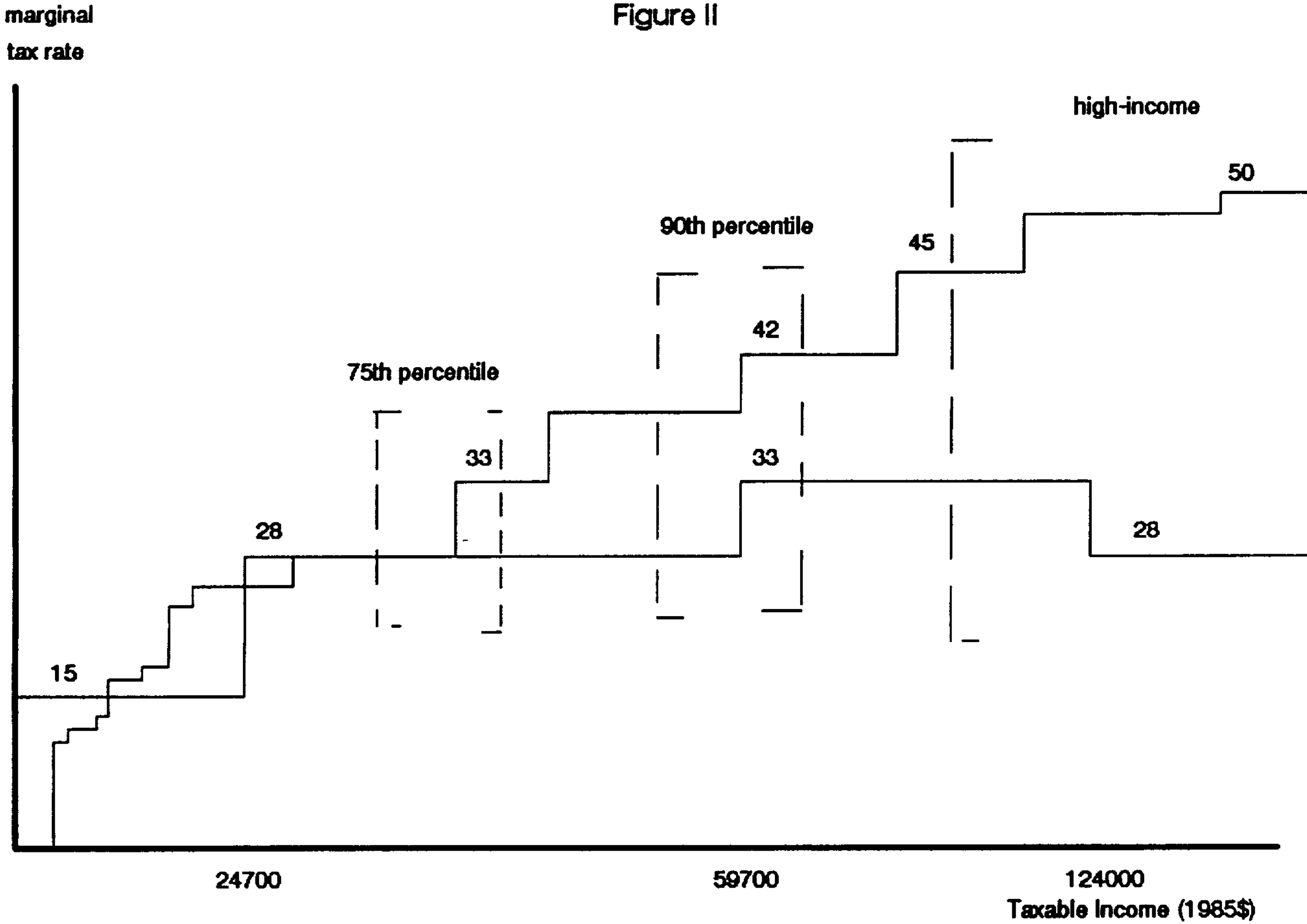
2) Representative example: Eissa (1995)

3) Uses the Tax Reform Act of 1986 to identify the effect of MTRs on labor force participation and hours of married women

4) TRA 1986 cut top income MTR from 50% to 28% from 1986 to 1988 but did not significantly change tax rates for the middle class

5) Substantially increased incentives to work of wives of high income husbands relative wives of middle income husbands

Figure II


marginal
tax rate
high-income
50
90th percentile
45
42
75th percentile
33
33
28
28
15
24700
59700
124000
Taxable Income (1985$)

Diff-in-Diff (DD) Methodology:

Step 1: Simple Difference

Outcome: LFP (labor force participation)

Two groups: Treatment group (T) which faces a change [women married to high income husbands] and control group (C) which does not [women married to middle income husband]

Simple Difference estimate: $D = LFP^T - LFP^C$ captures treatment effect if absent the treatment, LFP equal across 2 groups

Note: assumption always holds when T and C status is randomly assigned

Test for : Compare LFP before treatment happened $D_B = LFP_B^T - LFP_B^C$

Diff-in-Diff (DD) Methodology:

Step 2: Diff-in-Difference (DD)

If $D_B \neq 0$, can estimate DD:

$$DD = D_A - D_B = LFP_A^T - LFP_A^C - [LFP_B^T - LFP_B^C]$$

(A = after reform, B = before reform)

DD is unbiased if **parallel trend** assumption holds:

Absent the change, difference across T and C would have stayed the same before and after

Regression estimation of DD:

$$LFP_{it} = \beta_0 AFTER + \beta_1 TREAT + \gamma AFTER \cdot TREAT + \varepsilon$$

$\widehat{\gamma} = LFP_A^T - LFP_A^C - [LFP_B^T - LFP_B^C]$

Diff-in-Diff (DD) Methodology

DD most convincing when groups are very similar to start with [closer to randomized experiment]

Can test DD using data from more periods and plot the two time series to check parallel trend assumption

Use alternative control groups [not as convincing as potential control groups are many]

In principle, can create a DDD as the difference between actual DD and $DD^{Placebo}$ (DD between 2 control groups). However, DDD of limited interest in practice because

(a) if $DD^{Placebo} \neq 0$, DD test fails, hard to believe DDD removes bias

(b) if $DD^{Placebo} = 0$, then DD=DDD but DDD has higher s.e.

Table IIa
Marginal Tax Rate

Group	Before TRA86	After TRA86	Change	Relative Change
High	.521 (.002)	.382 (.001)	-.139 (.002)	
75th Percentile	.365 (.001)	.324 (.001)	-.041 (.001)	**-.098** **(.002)**
90th Percentile	.430 (.001)	.360 (.001)	-.07 (.001)	**-.069** **(.002)**

The marginal tax rate is calculated using family wage and salary, self-employment, interest, dividend, farm and social-security income. I assume all couples file jointly, and that all itemize their deductions. Itemized deductions and capital gains are imputed using Statistics of Income data. These figures include the secondary earner deduction, as well as social security taxes. Standard errors are in parentheses. Before TRA86 is tax years 1983-1985; After TRA86 is tax years 1989-1991.

Source: Eissa 1995

Table III
Differences-in-Differences Estimates
CPS Married Women Before and After TRA86

A: Labor Force Participation

Group	Before TRA86	After TRA86	Change	Difference-in-Difference
High	0.464 (.018) [756]	0.554 (.018) [718]	0.090 (.025) {19.5%}	
75th Percentile	0.687 (.010) [3799]	0.740 (.010) [3613]	0.053 (.010) {7.2%}	**0.037 (.028)** **{12.3%}**
90th Percentile	0.611 (.010) [3765]	0.656 (.010) [3584]	0.045 (.010) {6.5%}	**0.045 (.028)** **{13%}**

Source: Eissa 1995

B: Hours Conditional on Employment

Group	Before TRA86	After TRA86	Change	Difference-in-Difference
High	1283.0 (46.3) [351]	1446.3 (41.1) [398]	163.3 (61.5) {12.7%}	
75th Percentile	1504.1 (14.3) [2610]	1558.9 (13.9) [2676]	54.8 (20.0) {3.6%}	**108.6 (65.1) {9.4%}**
90th Percentile	1434.1 (16.4) [2303]	1530.1 (15.9) [2348]	96.0 (22.8) {6.8%}	**67.3 (64.8) {6.2%}**

Each cell contains the mean for that group, along with standard errors in (), number of observations in [], and % increase in {}. Means are unweighted.

Source: Eissa 1995

Eissa 1995: Results

1) Participation elasticity around 0.4 but large standard errors

2) Hours elasticity of 0.6

3) Total elasticity (unconditional hours) is $0.4 + 0.6 = 1$

Eissa 1995: Caveats

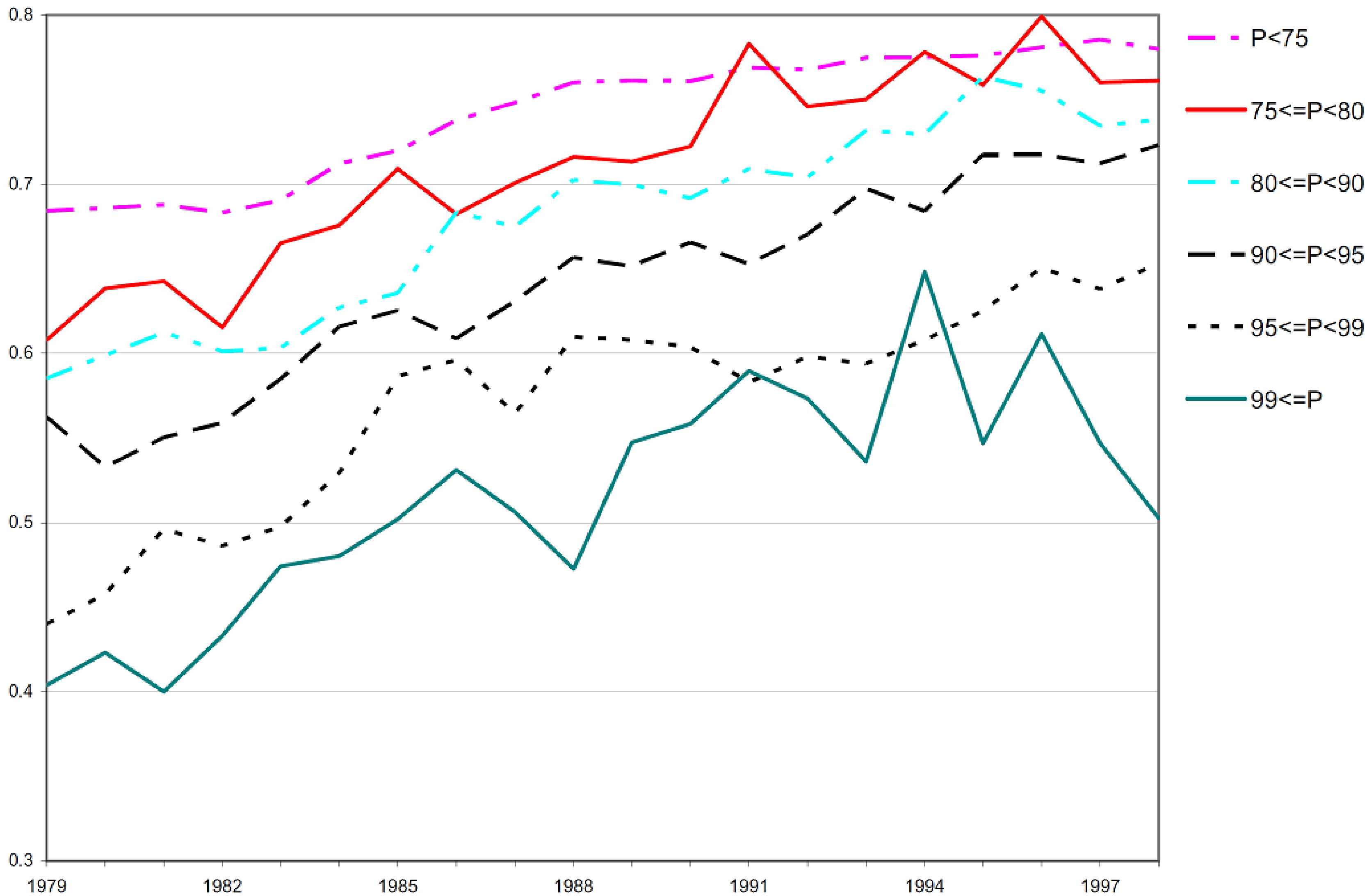
1) Does the common trends assumption hold? Potential story biasing the result:

Trend toward "power couples" and thus DD might not be due to taxes: In 1983-1985, professionals had non-working spouses, In 1989-1991, professionals married to professionals [and no change for middle class]

2) LFP before the reform is very different across T and C groups $\Rightarrow$ DD sensitive to functional form assumption [such as levels vs logs]

3) Liebman and Saez (2006) plot full time-series CPS plot and show that Eissa's results are not robust using admin data (SSA matched to SIPP) [unfortunately, IRS public tax data does not break down earnings within couples]

Figure 10
Fraction of Married Women with Positive Annual Earnings by Income Group
in March CPS



Notes: Groups are based on other household income (husband's earnings plus asset income) as described in Eissa (1995). Group 1 <=75th percentile. Group 75 is >75th percentile and <= 80th percentile. Group 80 is >80th and <=90th. Group 90 is >90th and <=95th. Group 95 is >95th and <=99th. Group 99 is >99th.

Source: Liebman and Saez (2006)

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

Overall Costs of Anti Poverty Programs

1) US government (fed+state and local) spent \$600bn in 2008 on income-tested programs

a) About 4% of GDP but 15% of \$4 Trillion govt budget (fed+state+local).

b) About 50% is health care (Medicaid)

2) Only \$150 billion in cash (1% of GDP, or 25% of transfer spending)

1996 Welfare Reform

1) Largest change in welfare policy

2) 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

3) Variation across states because Fed govt. gave block grants with guidelines

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

Monthly Welfare Case Loads: 1963-2000

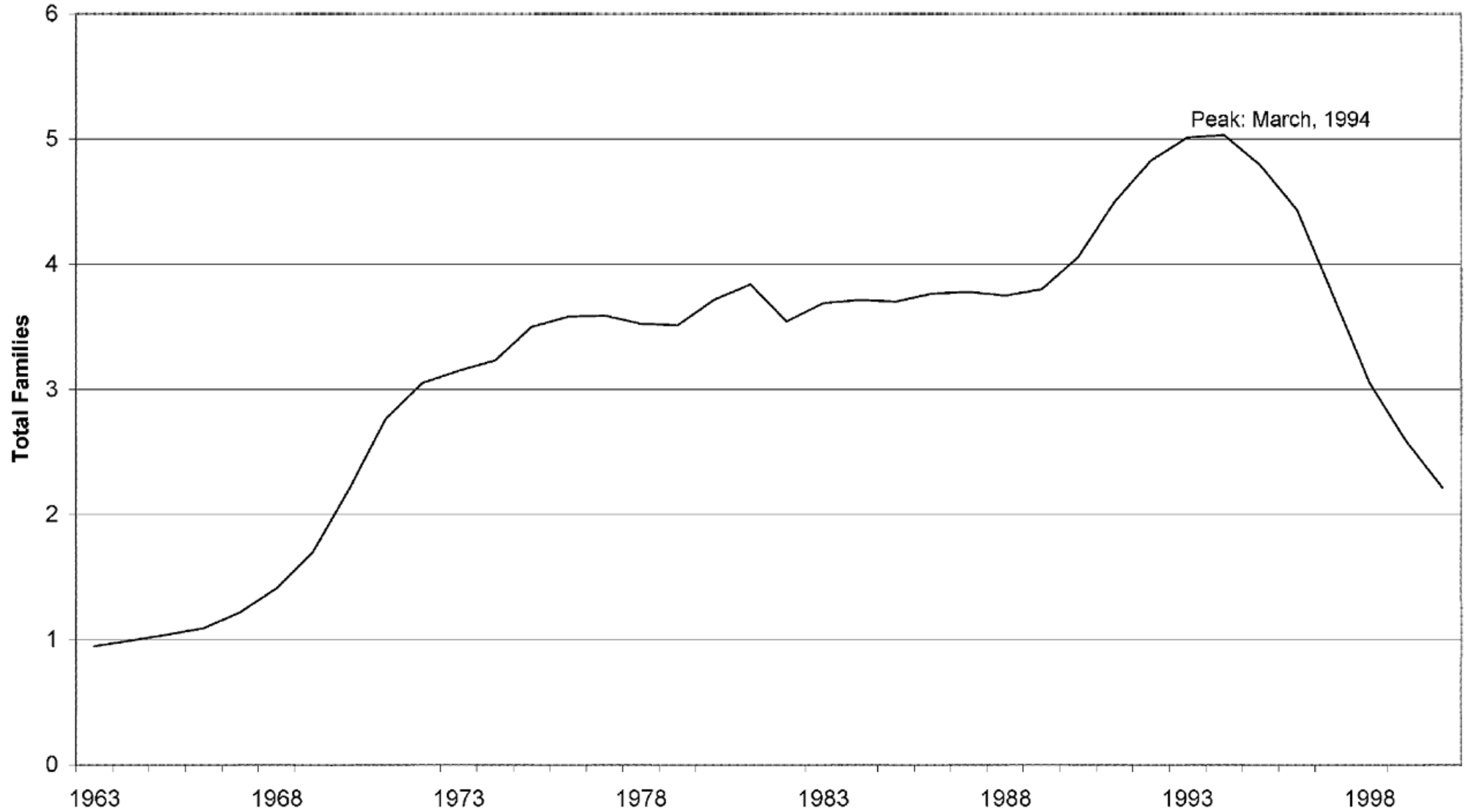


Fig. 1. Average monthly AFDC/TANF caseloads (1963 – 2000) (in millions).

Source: Meyer and Sullivan 2004

Welfare Reform: Two Empirical Questions

1) Incentives: did welfare reform actually increase labor supply? Test whether EITC expansions affect labor supply

2) Benefits: did removing many people from transfer system reduce their welfare? How did consumption change?

3) Focus on single mothers, who were most impacted by reform

Earned Income Tax Credit (EITC) program

Hotz-Scholz (2003) and Eissa-Hoynes '06 detailed surveys

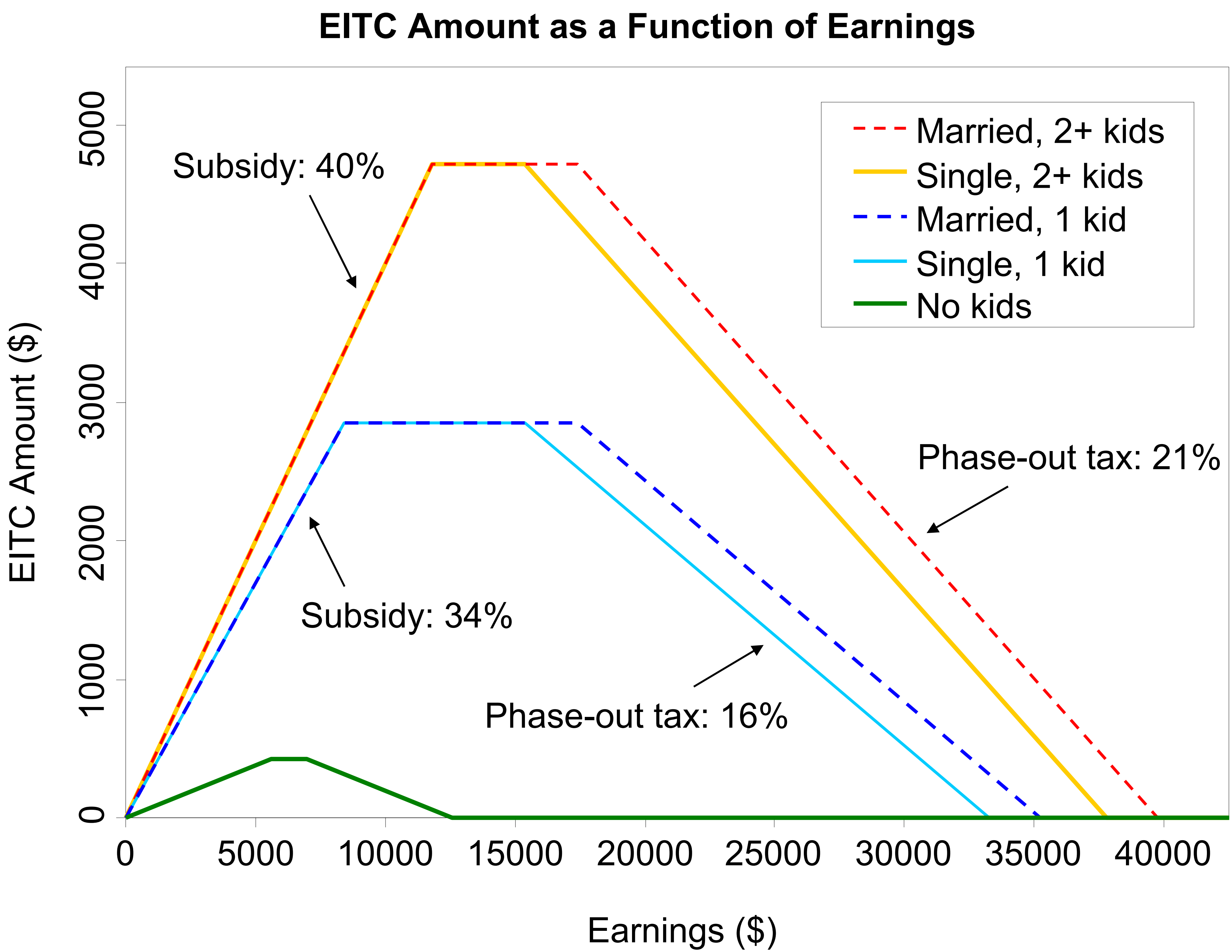
1) EITC started small in the 1970s but was expanded in 1986-88, 1994-96, 2008-09: today, largest means-tested cash transfer program [$50bn in 2009, 20m families recipients]

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) 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 Amount as a Function of Earnings
Married, 2+ kids
Single, 2+ kids
Married, 1 kid
Single, 1 kid
No kids
Subsidy: 40%
Subsidy: 34%
Phase-out tax: 21%
Phase-out tax: 16%
EITC Amount ($)
Earnings ($)
0
1000
2000
3000
4000
5000
0
5000
10000
15000
20000
25000
30000
35000
40000

Theoretical Behavioral Responses to the EITC

Extensive margin: positive effect on Labor Force Participation

Intensive margin: earnings conditional on working, mixed effects

1) Phase in: (a) Substitution effect: work more due to 40% inc. in net wage, (b) Income effect: work less $\rightarrow$Net effect: ambiguous; probably work more

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

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

Eissa and Liebman 1996

1) Pioneering study of labor force participation of single mothers before/after 1986-7 EITC expansion using CPS data

2) Limitation: this expansion was relatively small

3) Diff-in-Diff strategy:

a) Treatment group: single women with kids

b) Control group: single women without kids

c) Comparison periods: 1984-1986 vs. 1988-1990

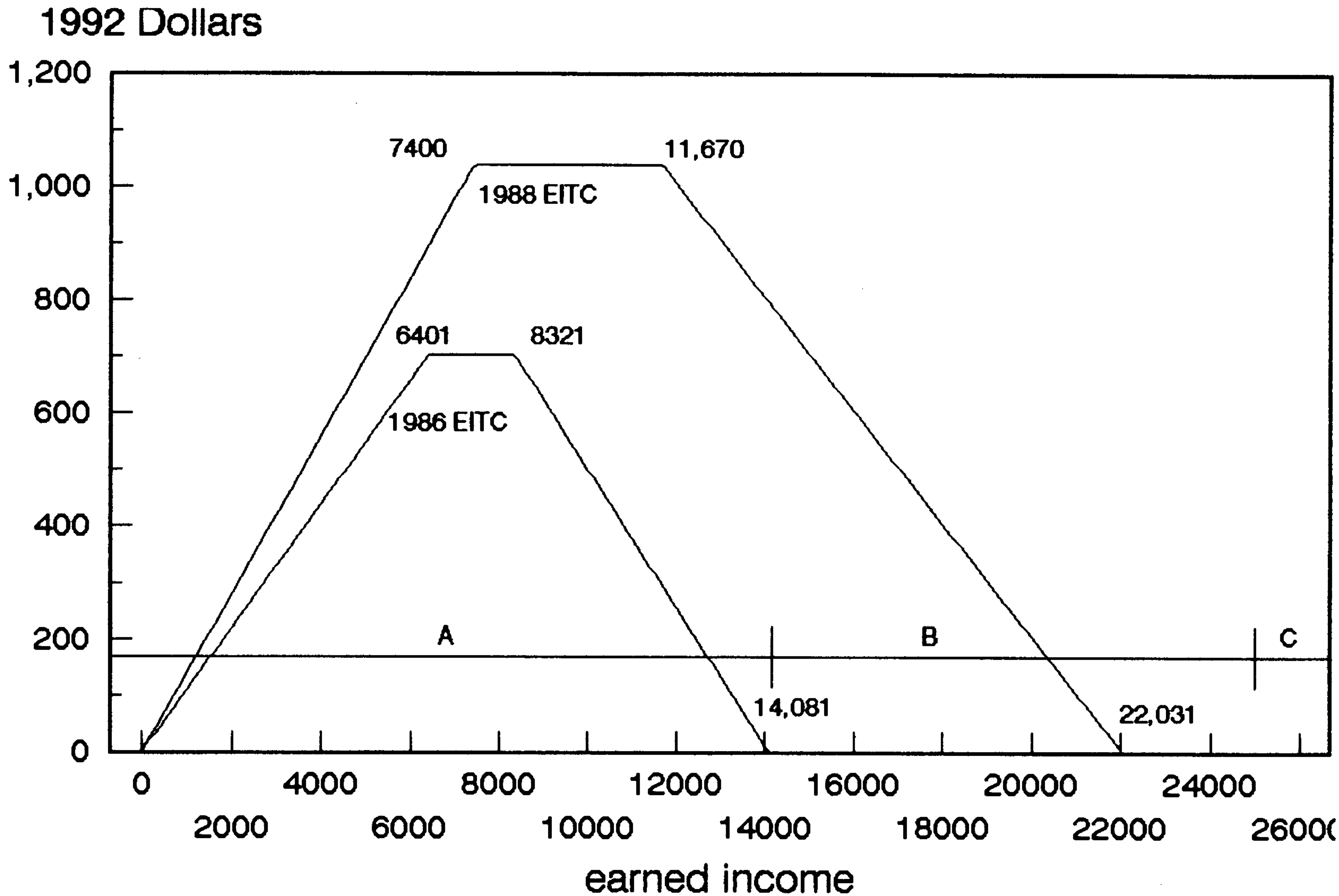


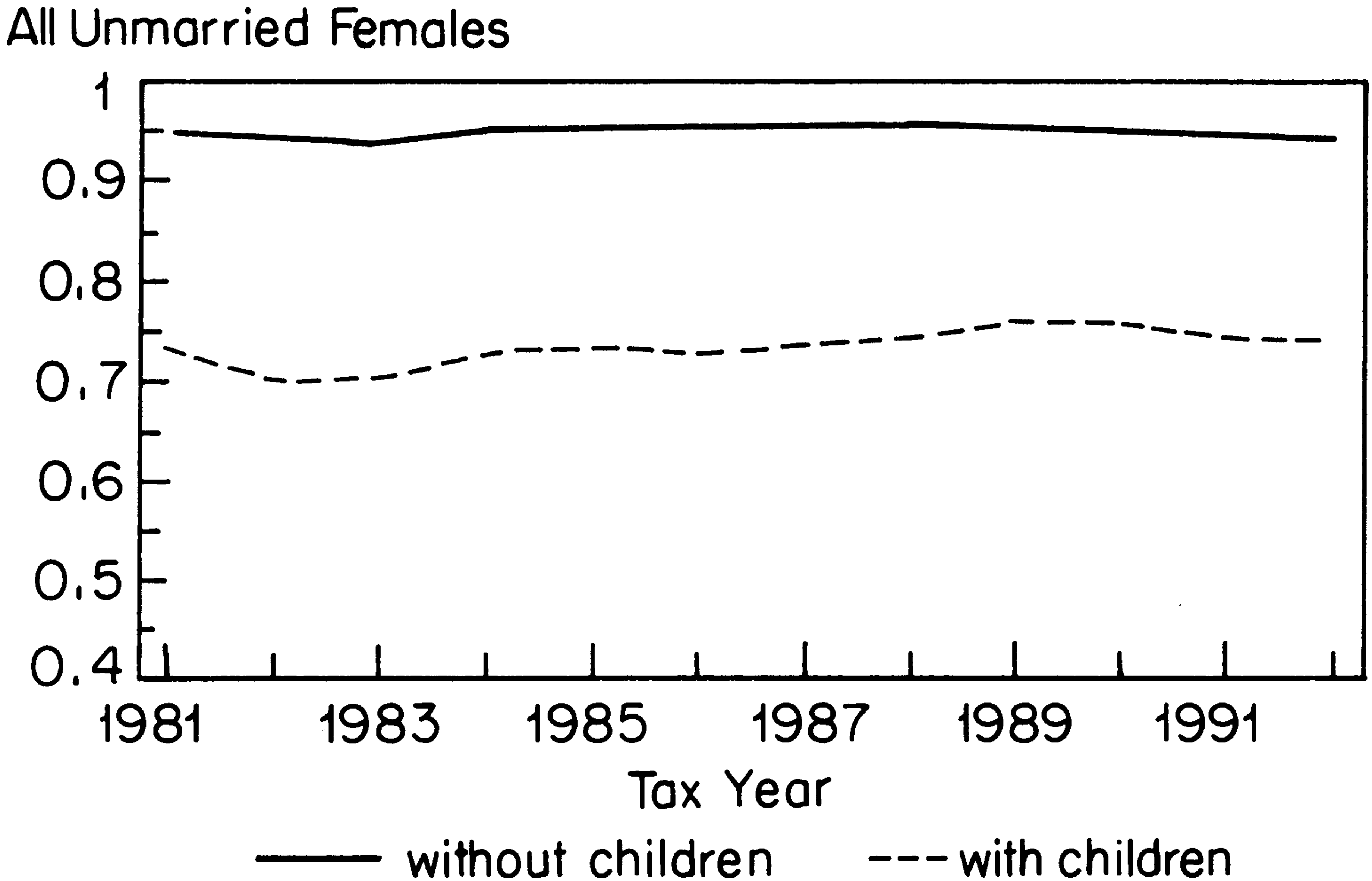
FIGURE IV
1986 and 1988 Earned Income Tax Credit

TABLE II

LABOR FORCE PARTICIPATION RATES OF UNMARRIED WOMEN

	Pre-TRA86 (1)	Post-TRA86 (2)	Difference (3)	Difference-in-differences (4)
A. *Treatment group:*				
With children [20,810]	0.729 (0.004)	0.753 (0.004)	0.024 (0.006)	
Control group:				
Without children [46,287]	0.952 (0.001)	0.952 (0.001)	0.000 (0.002)	*0.024 (0.006)*
B. *Treatment group:*				
Less than high school, with children [5396]	0.479 (0.010)	0.497 (0.010)	0.018 (0.014)	
Control group 1:				
Less than high school, without children [3958]	0.784 (0.010)	0.761 (0.009)	−0.023 (0.013)	*0.041 (0.019)*
Control group 2:				
Beyond high school, with children [5712]	0.911 (0.005)	0.920 (0.005)	0.009 (0.007)	*0.009 (0.015)*
C. *Treatment group:*				
High school, with children [9702]	0.764 (0.006)	0.787 (0.006)	0.023 (0.008)	
Control group 1:				
High school, without children [16,527]	0.945 (0.002)	0.943 (0.003)	−0.002 (0.004)	*0.025 (0.009)*
Control group 2:				
Beyond high school, with children [5712]	0.911 (0.005)	0.920 (0.005)	0.009 (0.007)	*0.014 (0.011)*

Data are from the March CPS, 1985–1987 and 1989–1991. Pre-TRA86 years are 1984–1986. Post-TRA86 years are 1988–1990. Labor force participation equals one if annual hours are positive, zero otherwise. Standard errors are in parentheses. Sample sizes are in square brackets. Means are weighted with CPS March supplement weights.



All Unmarried Females
1
0.9
0.8
0.7
0.6
0.5
0.4
1981
1983
1985
1987
1989
1991
Tax Year
without children
with children

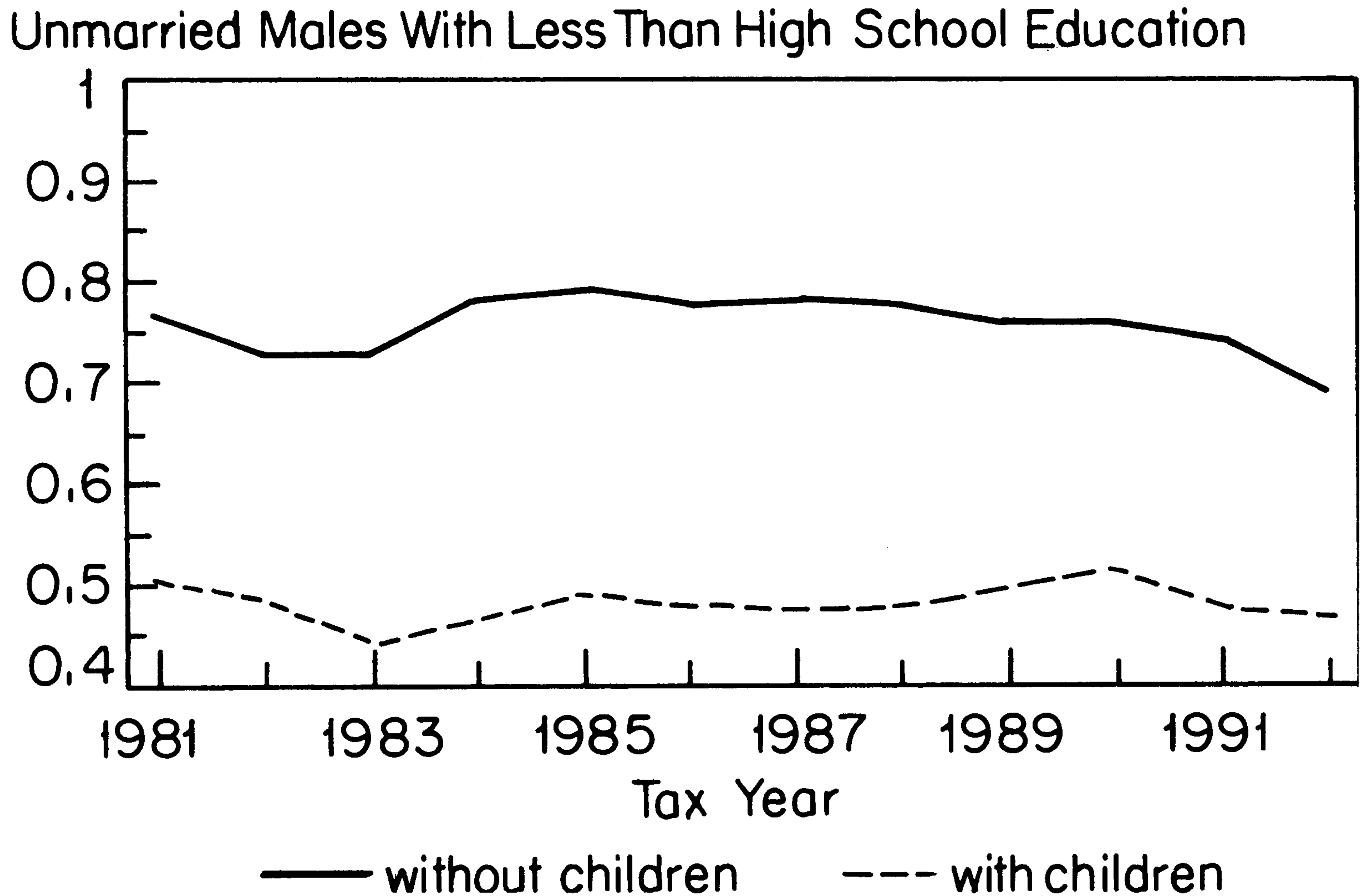


FIGURE II
Labor Force Participation Rates 1981 to 1992, Unmarried Females Ages 16–44

Eissa and Liebman 1996: Results

1) Find a small but significant DD effect: 2.4% (larger DD effect 4% among women with low education) $\Rightarrow$ Translates into substantial participation elasticities above 0.5

2) Note the labor force participation for women with/without children are not great comparison groups (70% LFP vs. +90%): time series evidence is only moderately convincing

3) Subsequent studies have used much bigger EITC expansions of the mid 1990s and also find positive effects on labor force participation of single women/single mothers

4) Conventional standard errors probably overstate precision

Bertrand-Duflo-Mullainathan QJE'04

Show that conventional standard errors in fixed effects regressions with state reform variation are too high

Randomly generated placebo state laws: half the states pass law at random date. I_{st} is one if state s has law in place at time t.

Use female wages w_{ist} in CPS data and run OLS:

$$w_{ist} = A_s + B_t + bI_{st} + \varepsilon_{ist}$$

$\widehat{b}$ significant (5% level) in 65% of cases $\Rightarrow$ ε_{ist} are not iid

Clustering by state*year cells is not enough (significant 45% of the time)

Need to cluster at state level to obtain reasonable s.e. because of strong serial correlation within states

Meyer and Rosenbaum 2001

1) Exploit the much bigger 1990s expansion in EITC

2) Document dramatic (6 pp, 10%) increase in LFP for single women with children around EITC expansion

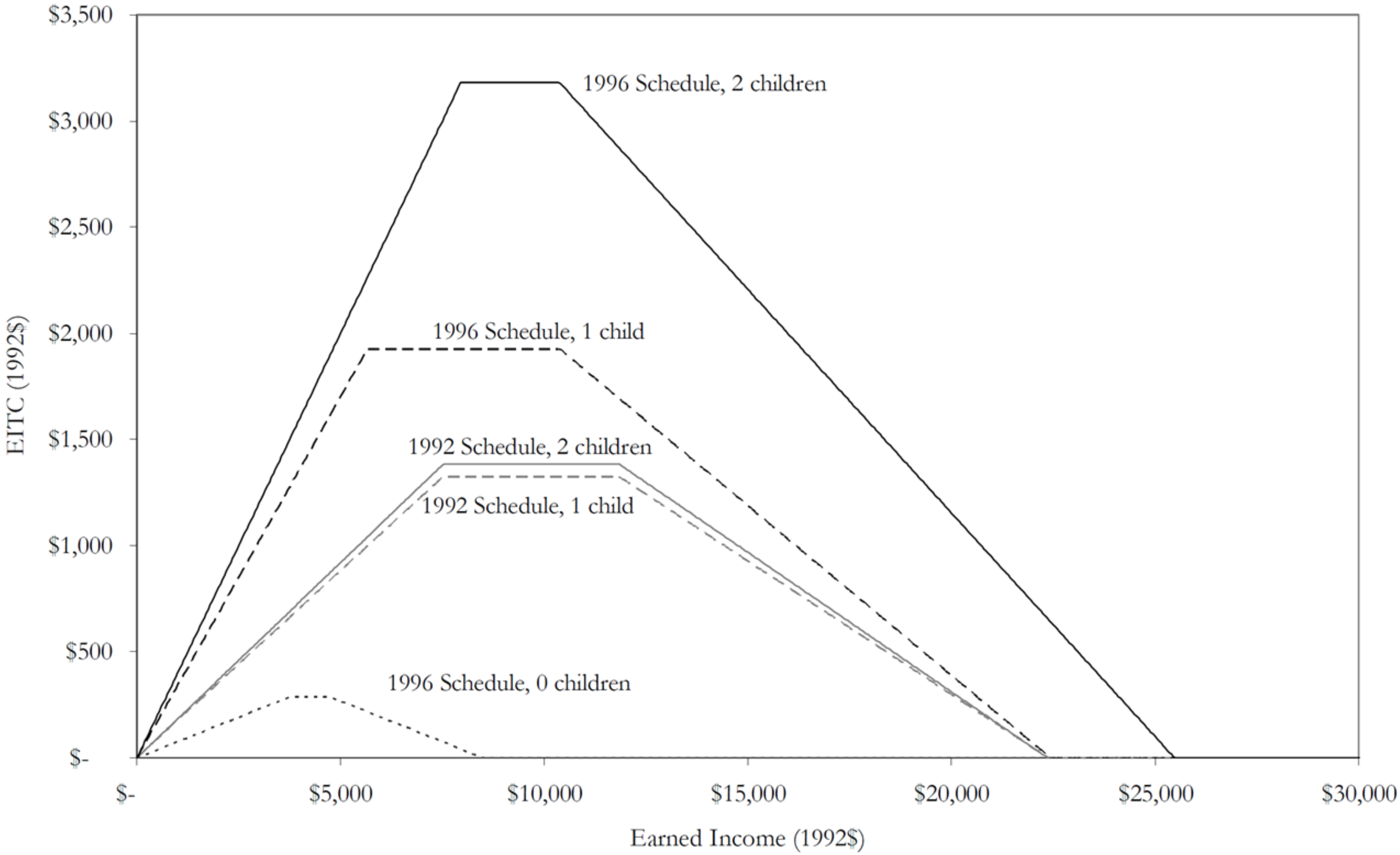
3) No change for women without children

4) Problem: expansion took place at same time as welfare reform

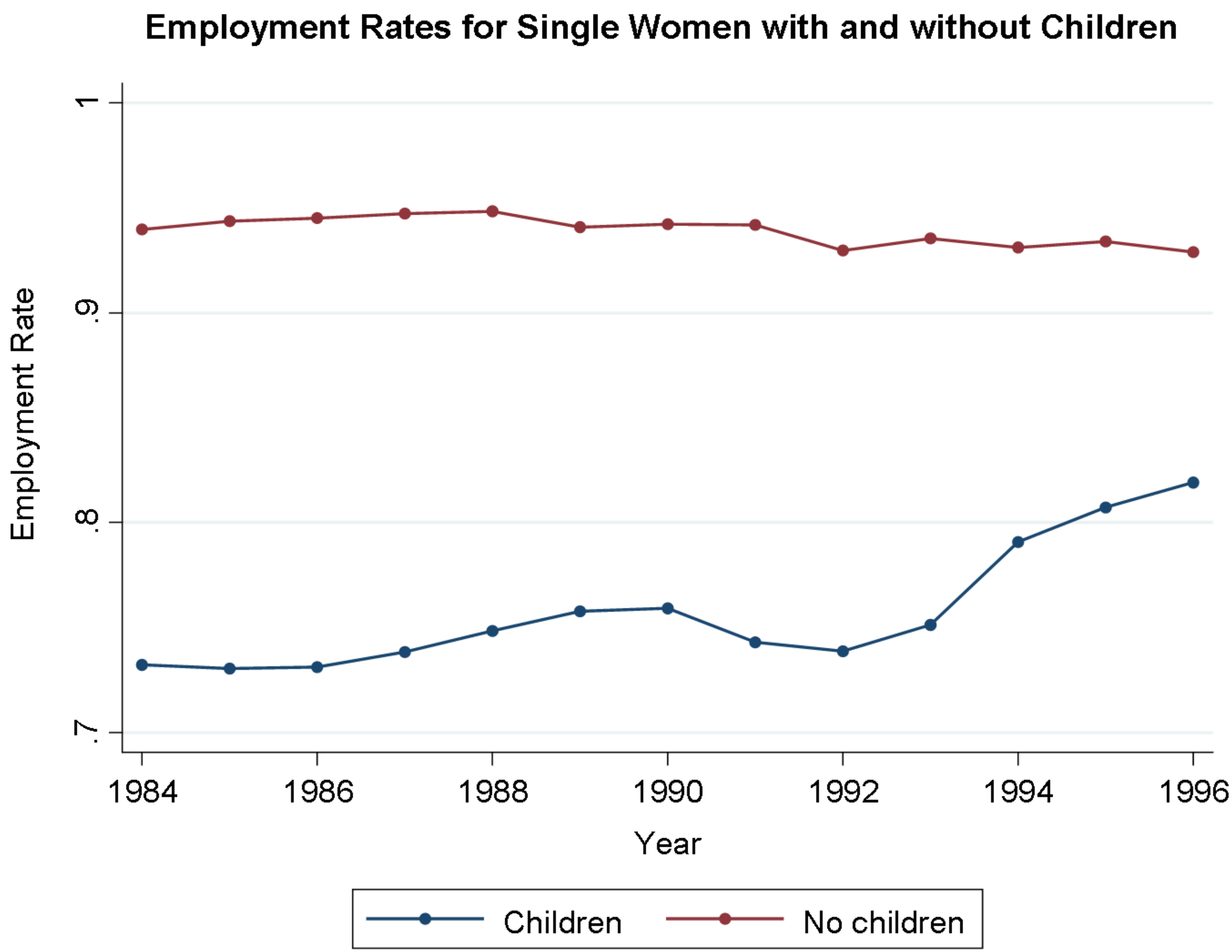
5) Try to disentangle effects of welfare waivers, changes in AFDC and state taxes, etc. using state-level variation

Bottom line: elasticity of participation w.r.t. tax/transfer incentives is significant

Figure 1. EITC Schedule, 1992 and 1996 by number of children

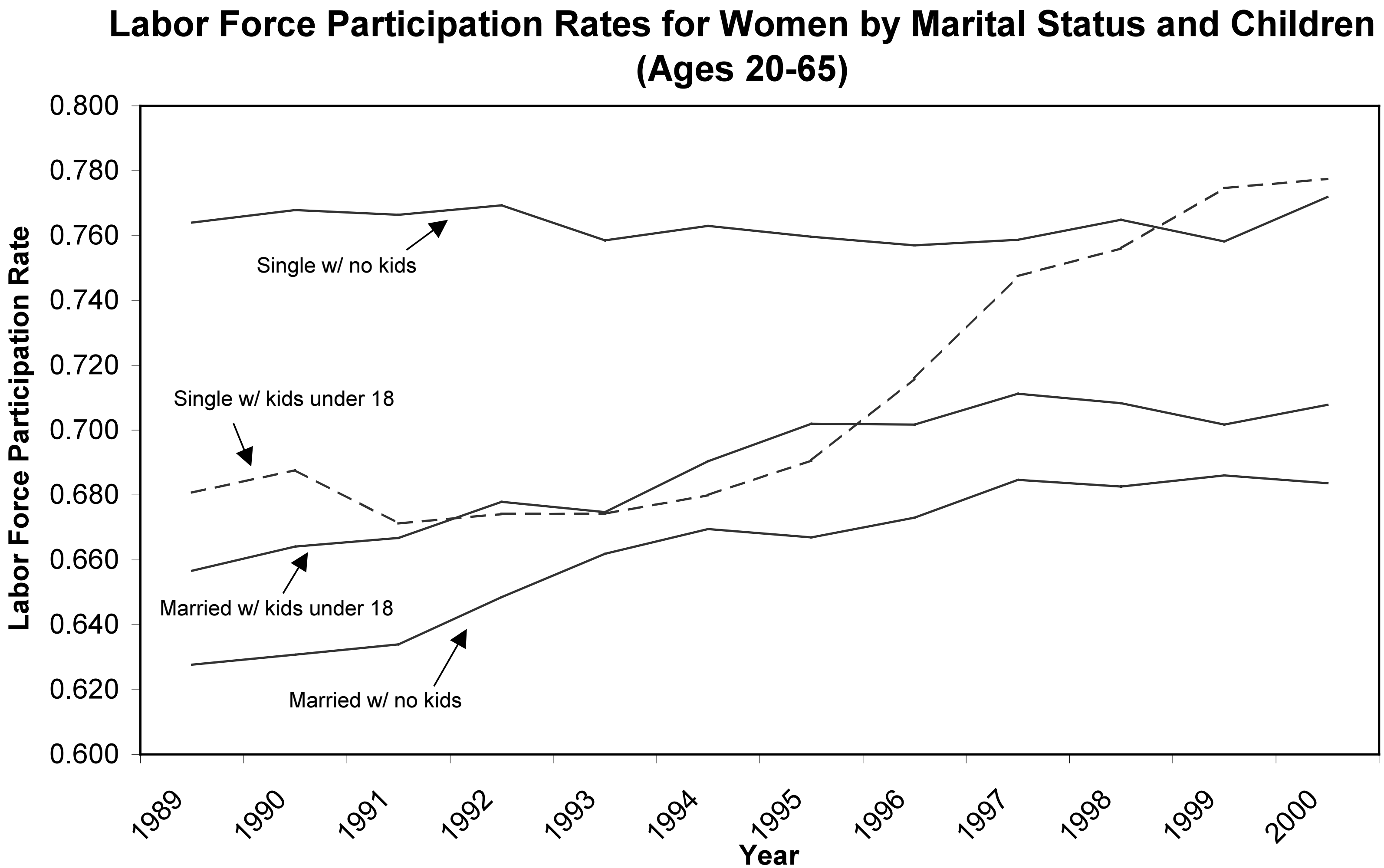


Source: Rothstein 2005



Source: Meyer and Rosenbaum 2001

Figure 4



Source: Tabulations of March Current Population Survey Data

Meyer and Rosenbaum 2001

1)Analyze the introduction of EITC and Welfare waivers for the period 1984-1996 using CPS data

2) Identification strategy: compare single mothers to single women without kids

3) Key covariates in regression model: (a) EITC, (b) AFDC benefits, (c) Medicaid, (d) Waivers, (e) Training, (f) Child Care

Meyer and Rosenbaum 2001

From 1984-1996, the extra increase in single mom's relative to single women without kids is explained by:

a) EITC expansion (60%)

b) Welfare max benefit reduction (AFDC and food stamps) (25%)

c) Medicaid if work (-10%) (insignificant and wrong sign)

d) Welfare waivers (time limits) 15%

e) Child care and training: 15%

Eissa and Hoynes 2004

1) EITC based on family rather than individual income

2) Study married couples with low earnings, recognizing that EITC *reduces* their incentive to work

3) Married women with husband earning $10-15K are in the phase-out range and face high MTR's

a) Payroll tax 15%

b) EITC phase-out 20%

c) State and federal income tax 0-20%

3) Similar identification strategy: compare those with and without kids

Eissa and Hoynes: Results

1) Conclude that EITC expansions between 1984 and 1996:

a) Increased married men's labor force participation by 0.2%

b) Reduced married women's labor force participation by about 1%

2) Implies that the EITC is effectively subsidizing married mothers to stay at home and *reducing* total labor supply for married households

Meyer and Sullivan 2004

1) Examine the consumption patterns of single mothers and their families from 1984–2000 using CEX data

2) Question: did single mothers' consumption fall because they lost welfare benefits and were forced to work?

Total Consumption: Single Mothers 1984-2000

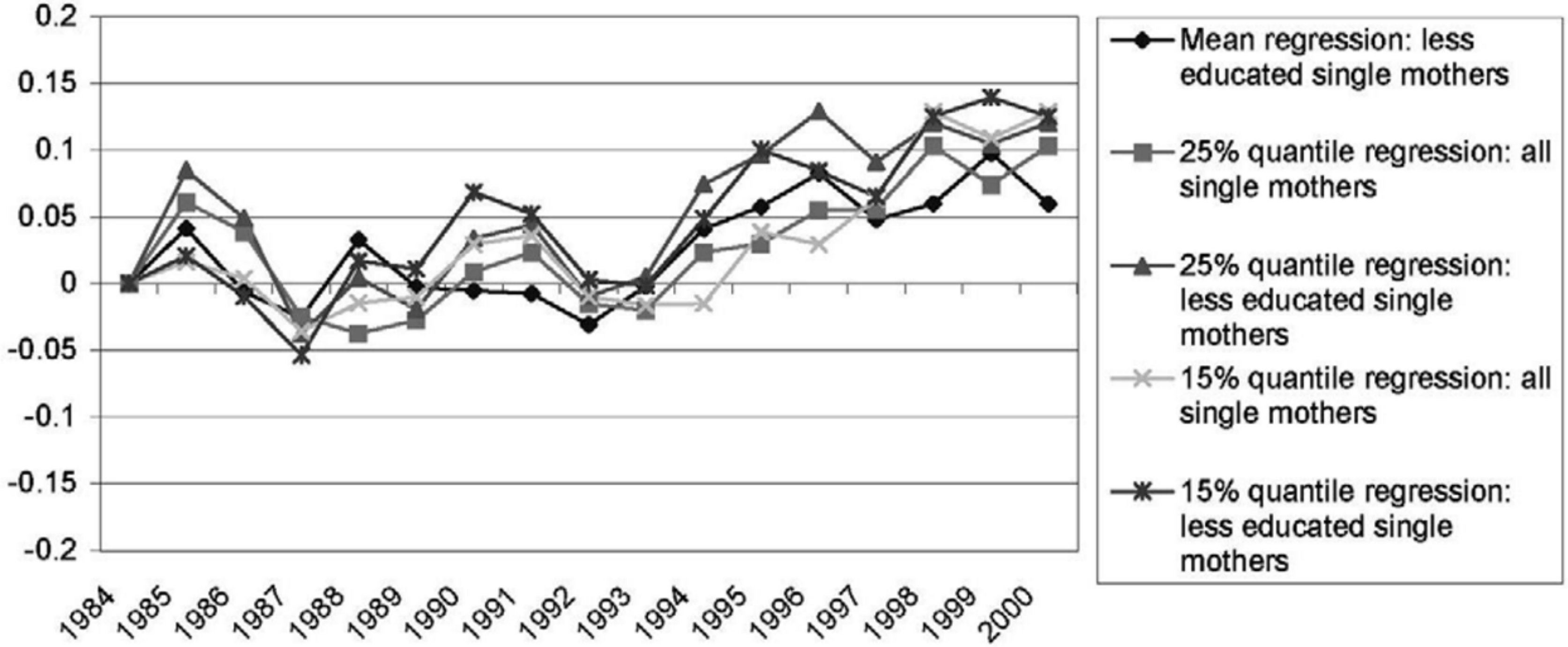


Fig. 2. Total consumption: single mothers, 1984–2000.

Source: Meyer and Sullivan 2004

Relative Consumption: single women with/without children

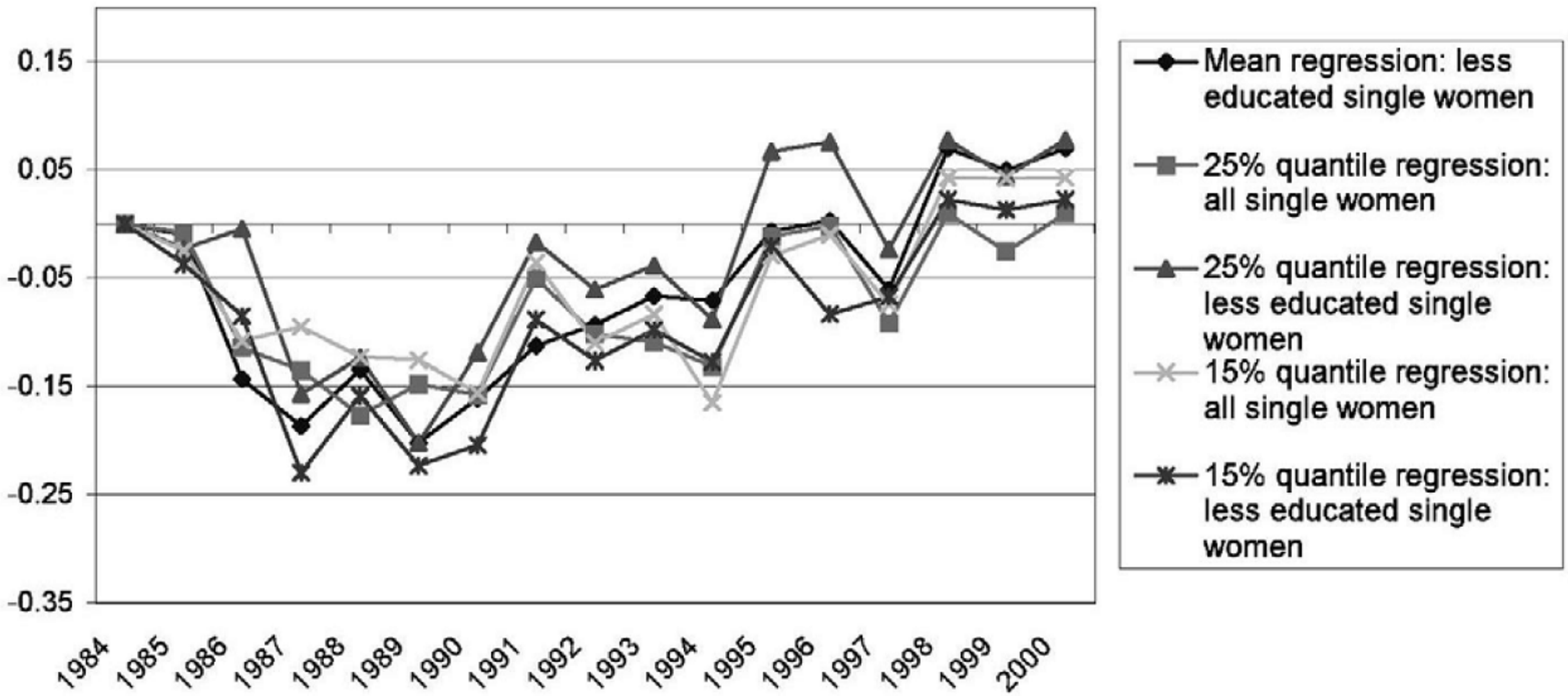


Fig. 3. Relative total consumption: single mothers vs. single women without children, 1984–2000.

Source: Meyer and Sullivan 2004

Relative Consumption: married vs. single mothers

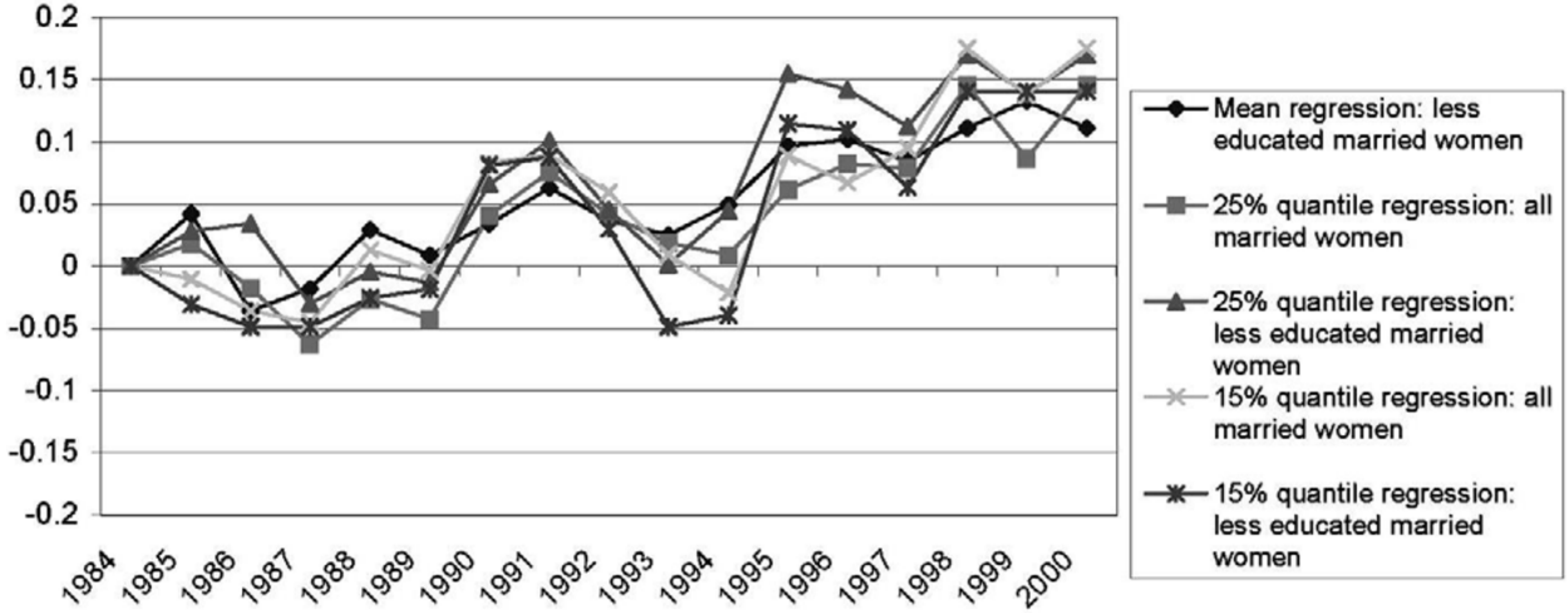


Fig. 4. Relative total consumption: single mothers vs. married mothers, 1984–2000.

Source: Meyer and Sullivan 2004

Meyer and Sullivan: Results

1) Material conditions of single mothers did not decline in recent years, either in absolute terms or relative to single childless women or married mothers

2) In most cases, evidence suggests that the material conditions of single mothers have improved slightly

3) Question: is this because economy was booming in 1990s?

4) Is workfare approach more problematic in current economy?

EITC Behavioral Studies

Strong evidence of response along extensive margin, little evidence of response along intensive margin (except for self-employed) $\Rightarrow$ 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 $\uparrow$ or $\downarrow$ 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 and Saez '09 EITC INFO

1) Randomized experiment with tax preparer H&R Block: tax pros [H&R Block employees] provide EITC information to half of 43,000 EITC filers in 2008 tax season

2) Analyze whether earnings the following year are affected by the information treatment

Key results:

1) Half of the tax pros induce treated clients to increase their EITC refunds by choosing an earnings level closer to the peak of the EITC schedule

2) Rest of tax professionals seem to increase earnings of their treated clients across the board [possible

3) Treatment effects are larger for the self-employed

Explaining EIC: 4 steps

Single With Two or More Children

The EIC (Earned Income Credit) is a tax refund that gives families as much as $4,500 per year.

We want to explain how the **EIC** works to help you decide how much to work and earn this year.

In 2006, you made $ 10,000 → you are getting an **EIC** of $ 4,000 in your refund.

1. Fill in earnings, EIC amount

- Your earnings this year (in 2007) will determine the size of your **EIC** refund next year
- The **EIC** has 3 ranges: 1) Increasing, 2) Peak, 3) Decreasing

2. Explain and dot graph

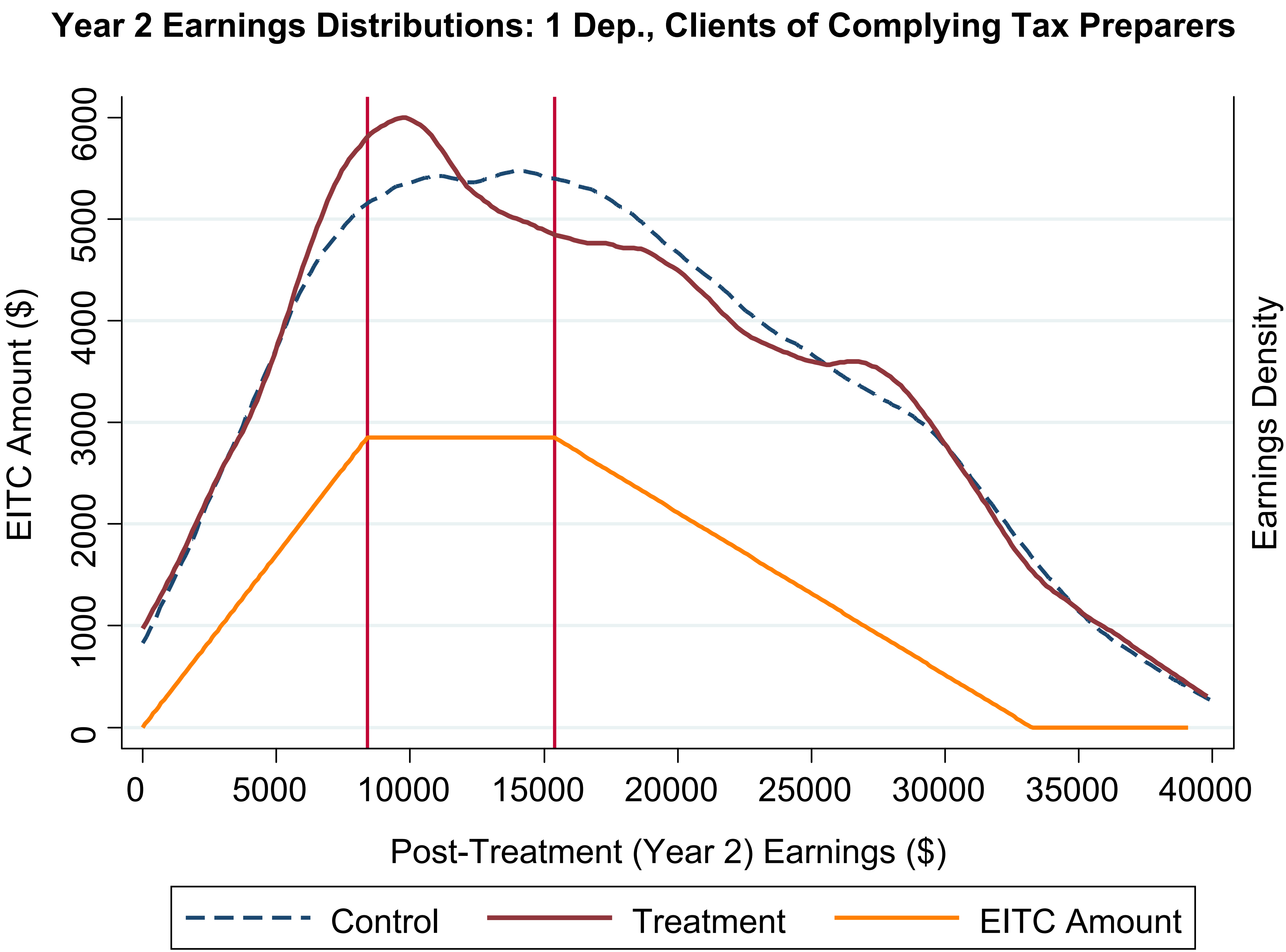
4. Take-home Message

You are in the increasing range of the **EIC**. *Think about it like this:*

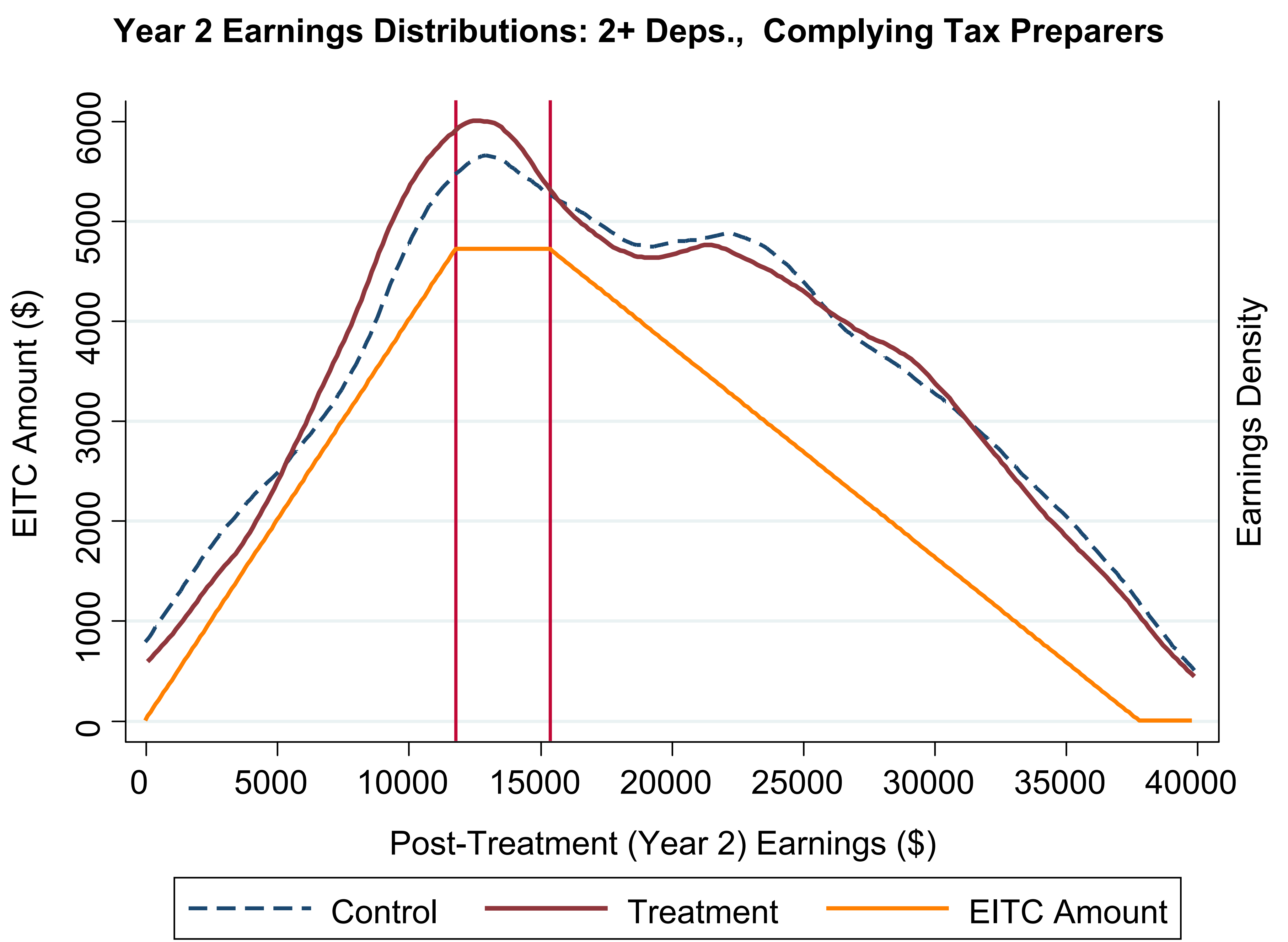
- (increasing) Suppose you earn $10 an hour, then you are really making $14.00 an hour.
- (peak) Your earnings are maxing-out the **EIC** amount
- (decreasing) If you earn $10 more, your **EIC** is reduced by $2.10

3. Table

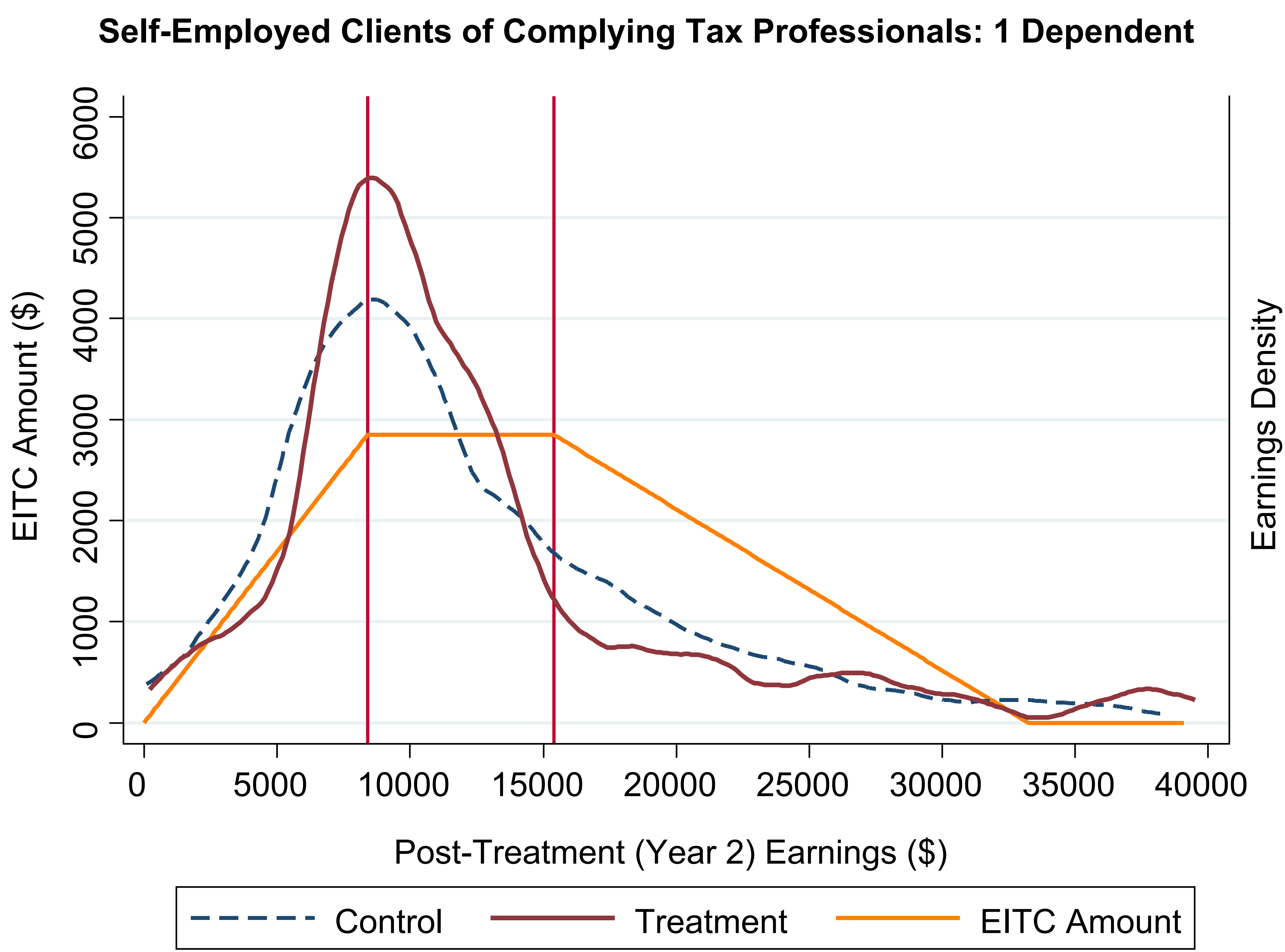
EIC Range	If you earn between	EIC refund will be	If you earn $10 more, the EIC...
Increasing	$0-$11,790	$0 up to $4,716	Increases by $4
Peak	$11,790-$15,390	$4,716	Stays the same
Decreasing	$15,390-$37,780	$4,716 down to $0	Decreases by $2.10

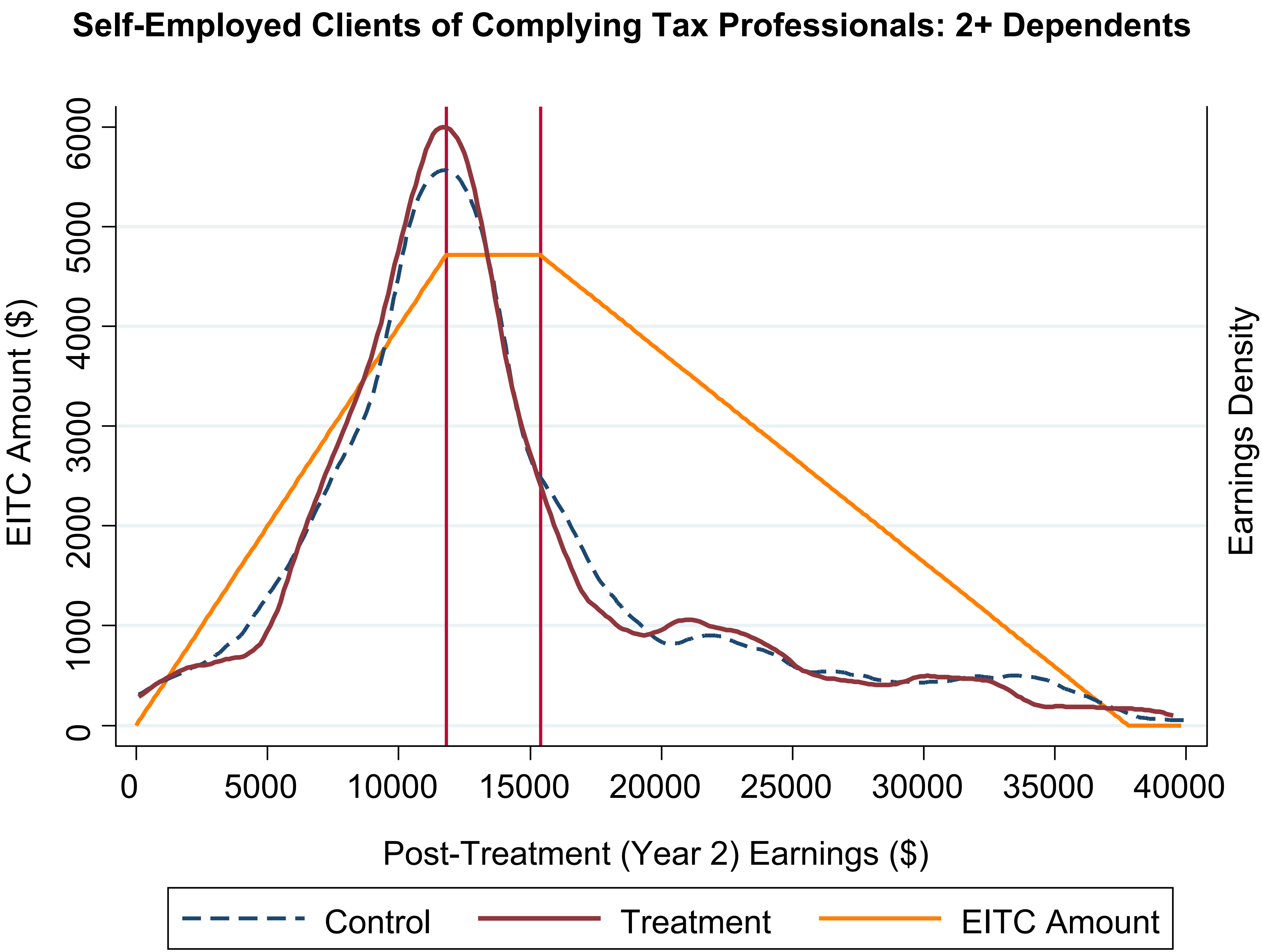


Year 2 Earnings Distributions: 1 Dep., Clients of Complying Tax Preparers
EITC Amount ($)
Earnings Density
0
1000
2000
3000
4000
5000
6000
0
5000
10000
15000
20000
25000
30000
35000
40000
Post-Treatment (Year 2) Earnings ($)
Control
Treatment
EITC Amount



Year 2 Earnings Distributions: 2+ Deps., Complying Tax Preparers
EITC Amount ($)
Earnings Density
0
1000
2000
3000
4000
5000
6000
0
5000
10000
15000
20000
25000
30000
35000
40000
Post-Treatment (Year 2) Earnings ($)
Control
Treatment
EITC Amount


Self-Employed Clients of Complying Tax Professionals: 1 Dependent
EITC Amount ($)
0
1000
2000
3000
4000
5000
6000
Earnings Density
0
5000
10000
15000
20000
25000
30000
35000
40000
Post-Treatment (Year 2) Earnings ($)
Control
Treatment
EITC Amount


Self-Employed Clients of Complying Tax Professionals: 2+ Dependents
EITC Amount ($)
0
1000
2000
3000
4000
5000
6000
Earnings Density
0
5000
10000
15000
20000
25000
30000
35000
40000
Post-Treatment (Year 2) Earnings ($)
Control
Treatment
EITC Amount

Chetty and Saez '09 IMPLICATIONS

Empirical work:

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

Cannot identify structural parameters of preferences without modelling information and salience

Normative analysis:

Information is a powerful and inexpensive policy tool to affect behavior

Should be incorporated into optimal policy design problems

ADVANCE EITC

Recipients get EITC with tax refund in a single annual refund in Feb year $t+1$ which seems suboptimal: (a) free interest loan to govt and (b) harder to smooth consumption [surveys show that primary use of tax refund is to pay overdue bills]

Tax filers have option to use Advance EITC to get part of EITC in the paycheck by filing a W5 form with employer [reverse of tax withholding]: take up extremely low (<2%)

Possible explanation: (a) Information, (b) Lack of employer cooperation, (c) Risk of owing taxes if not EITC eligible, (d) Tax filers like big refunds, (e) Inertia (default is no Advance EITC)

ADVANCE EITC

Jones AEJ-AP'10 carries a randomized experiment with large employer to encourage take-up and gets significant but very small take-up effect suggesting that (a) [Information] and (b) [Employer cooperation] cannot explain low take-up

(d) [Love of refunds] seems plausible but (1) not supplied by market absent refunds [employers could also pay part of wages as annual lumpsum], (2) A-EITC use has not $\uparrow$ with EITC expansions

(c) [Risk of owing taxes] and (e) [Inertia] are likely part of the explanation

Interesting research topic: Have big tax refunds fuelled low income credit [tax refund loans, payday loans, etc.]? Are big refunds useful forced saving mechanisms?

Other Behavioral Responses to Transfer Programs

1) Bitler, Gelbach, and Hoynes (2005): distributional effects are very important in understanding welfare programs because of nonlinearities in bc $\rightarrow$ cannot just look at means

2) Other studies have examined effects of low-income assistance programs on other margins such as family structure (divorce rate, number of kids) and find limited effects

3) Empirical work on tagging and in-kind programs is more limited and is an important area for further research

Changing Married Women Elasticities: Blau and Kahn JOLE'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

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 int + ext margin) shrank from 0.4 in 1980 to 0.2 today, (b) effect of husband earnings $\downarrow$ overtime

4) Interpretation: elasticities shrink as women become more attached to the labor force

Summary of Static Labor Supply Literature

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 incomes, retirees

3) Responses driven by extensive margin

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

b) Intensive margin (hours) elasticity close 0

Intertemporal Models and the MaCurdy Critique

1) What parameter do reduced-form regressions of labor supply on wages or taxes identify?

2) MaCurdy critique: reduced-form studies did not identify any parameter of interest in a dynamic model

3) Instead, estimate a mix of income effects, intertemporal substitution effects, and compensated wage elasticities

4) MaCurdy (1981) develops a structural estimation method (two stage budgeting) to identify preference parameters in a life-cycle model of labor supply (see Chetty '06 for simple exposition)

Life Cycle Model of Labor Supply

General model is of the form:

$$
\begin{aligned}
&U(c_0, .., c_T, l_0, .., l_T) \\
&s.t. \;\; A_0 + \sum w_t l_t/(1+r)^t \geq \sum c_t/(1+r)^t \quad (\lambda)
\end{aligned}
$$

Key Assumption for inter-temporal budget is **no credit constraints**

$\Rightarrow$ First order conditions:

$$
\begin{aligned}
U_{l_t}(c_0, .., c_T, l_0, .., l_T) + \lambda w_t/(1+r)^t &= 0 \\
U_{c_t}(c_0, .., c_T, l_0, .., l_T) - \lambda/(1+r)^t &= 0
\end{aligned}
$$

In the general case, $l_t(A_0, w_0, .., w_T)$ same as the multi-good choice – no generic results

Life Cycle Model: Time Separability

By assuming time separability can rewrite the problem as:

$$U = \sum_{t=0}^{T} \beta^t u(c_t, l_t)$$

Leads to simpler first order conditions

$$\begin{aligned} l_t &: \beta^t u_{l_t} + \lambda w_t/(1+r)^t = 0 \\ c_t &: \beta^t u_{c_t} - \lambda/(1+r)^t = 0 \end{aligned}$$

Combining yields: $-u_l = w_t u_c$

Intratemporal FOC same as in static model

Intertemporal FOC: $u_{c_t}/u_{c_{t+1}} = \beta(1+r)$

Dynamic Life Cycle Model: Policy Rules

1) $\lambda = u_{c_0}$ is the marginal utility of initial consumption

2) The two first order conditions imply that

$$\begin{aligned} l_t &= l(w_t, \lambda/(\beta(1+r))^t) \\ c_t &= c(w_t, \lambda/(\beta(1+r))^t) \end{aligned}$$

3) Current labor and consumption choice depends on current w_t

4) All other wage rates and initial wealth enter only through the budget constraint multiplier λ (MaCurdy 1981)

5) Easy to see for separable utility [$u(c)$ concave, $v(l)$ convex]:

$$\begin{aligned} u(c,l) &= u(c) - v(l) \\ \Rightarrow v'(l_t) &= \lambda w_t/[\beta(1+r)]^t \\ \Rightarrow l_t &= v'^{-1}(\lambda w_t/[\beta(1+r)]^t) \end{aligned}$$

6) Sufficiency of λ greatly simplifies solution to ITLS model

Dynamic Life Cycle Model: Frisch Elasticity

Frisch intertemporal labor supply elasticity defined as:

$$\delta = (\frac{w_t}{l_t})\frac{\partial l}{\partial w_t}|_\lambda$$

Experiment: change wage rate in one period only, holding all other wages, and consumption profile constant

Can show that $\delta > 0$: work more today to take advantage of temporarily higher wage

In separable case:

$$\begin{aligned} v'(l_t) &= \lambda w_t/[\beta(1+r)]^t \\ \Rightarrow \frac{\partial l}{\partial w_t}|_\lambda &= \frac{\lambda}{[\beta(1+r)]^t v''(l_t)} > 0 \end{aligned}$$

Frisch vs Hicksian Elasticity: Illustrative Example

Suppose that you are paid a piece rate

It takes 1 hour of work to make a piece

You usually work from 8am-12pm and 1pm-5pm.

Suppose your employer tells you that the piece rate will be twice as high only during the 12pm-1pm time slot

What do you do?

$\rightarrow$Have lunch earlier at 11am-12pm and work from 12pm-1pm

ITLS and Income Effects

Single inter-temporal budget constraint:

$$A_0 + \sum w_t l_t/(1+r)^t \geq \sum c_t/(1+r)^t$$

$\Rightarrow$ Receiving \$ M in year 0 vs. \$ $(1+r)^t \cdot M$ in year t has the same impact on labor supply

Temporary transfer has a small effect on labor in **all** periods

In reality, temporary transfers seem to have large effects on labor supply [e.g., severance payments, Card-Chetty-Weber QJE'08] $\Rightarrow$

(1) Many people are credit constrained: static labor supply model might be a better depiction of reality

(2) People might not make intertemporal choices as in ITLS model [behavioral economics]

Dynamic Life Cycle Model: Three Types of Wage Changes

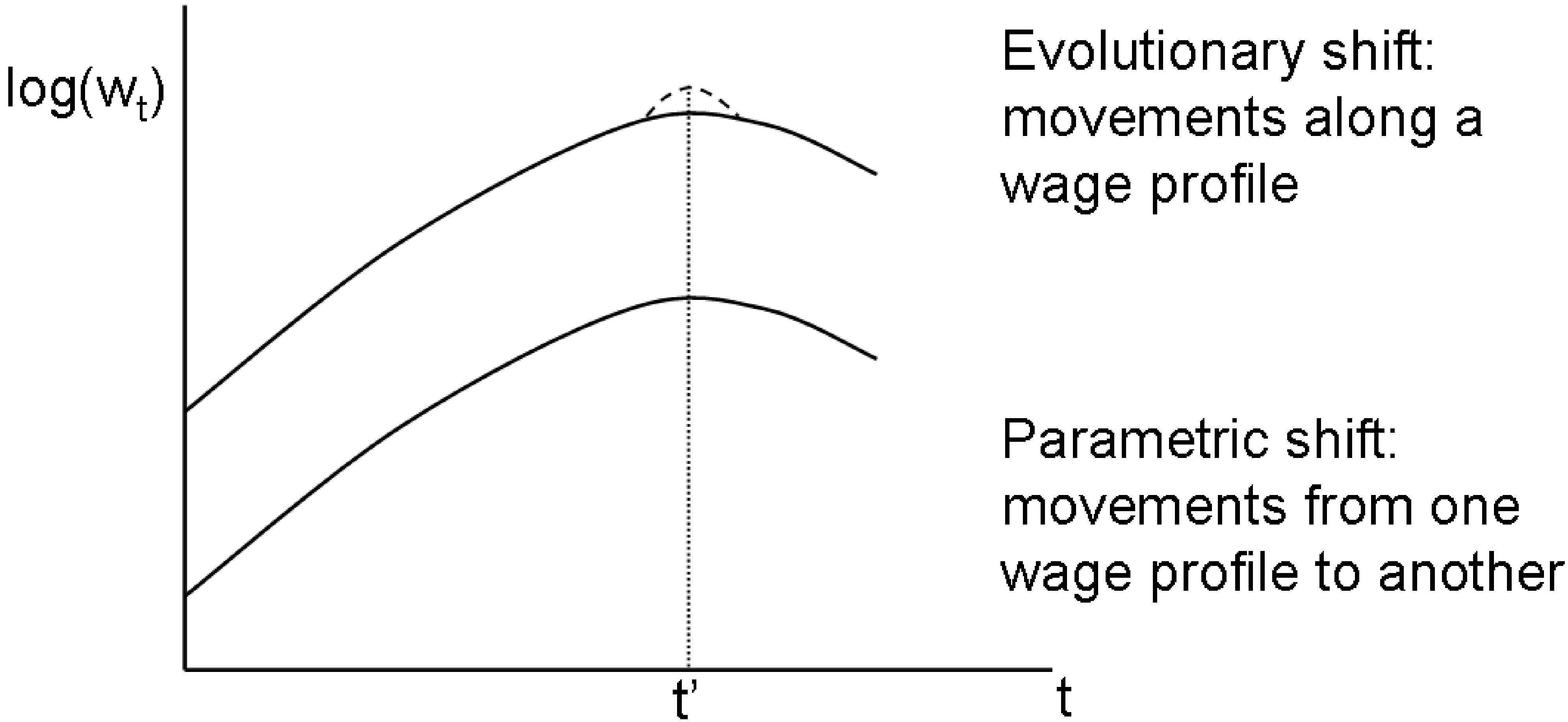
1) Evolutionary change: movements along profile (life-cycle)

2) Parametric change: temporary tax cut

3) Profile shift: changing the wage rate in all periods

a) Equivalent to a permanent parametric change

b) Implicitly the elasticity that static studies estimate with unanticipated tax changes



Source: MaCurdy 1981

Frisch vs. Compensated vs. Uncompensated Elasticities

Frisch elasticity: changing wages in a single period and keeping marginal utility of income λ constant

Compensated static elasticity: changing wages in all periods but keeping utility constant

Uncompensated static elasticity: changing wages in all periods with no compensation

Frisch vs. Compensated vs. Uncompensated Elasticities

Intertemporal substitution: Frisch elasticity $\geq$ Compensated static elasticity

Income effects: Compensated static elasticity $\geq$ Uncompensated static elasticity

Difference in Frisch and Compensated elasticities also loosely related to anticipated vs. unanticipated changes

Looney and Singhal (2007) exploit this reasoning to identify Frisch elasticity [MTR changes predictably when filers loose a child exemption]

Frisch elasticity is of central interest for calibration of macro business cycle models

MaCurdy 1983

1) Structural estimate using panel data for men and within-person wage variation

2) Find both Frisch and compensated wage elasticity of around 0.15

3) But his wage variation is not exogenous

Pencavel 2002

1) Instruments with trade balance interacted with schooling and age

2) Frisch elasticity: 0.2

3) Uncompensated wage elasticity: 0-0.2

Instruments not credibly exogenous but results closer to structurally interpretable parameters

Critique of ITLS models

• Card critique of value of ITLS model

a) Fails to explain most variation in hours over lifecycle

b) Sheds little light on profile-shift elasticities that we care about for policy

• Core "structural vs. reduced-form" divide in applied microeconomics: Trade off between credible identification and well defined theoretical framework

Blundell, Duncan, and Meghir 1998

1) Good combination of structural and reduced form methods on labor supply

2) Argue against standard DD approach, where treatment/control groups are endogenously defined

a) Reduced tax rate may pull households from low income group to high income group

b) Need group definitions that are stable over time

3) Use birth cohort (decade) interacted with education (e.g. high school or more)

Blundell, Duncan, and Meghir 1998

1) Construct group-level labor supply measures for women in couples

2) Measure how labor supply co-varies with wages rates net of taxes in the UK in 1980s

3) Importantly, tax reforms during this period affected groups very differently

4) Use consumption data as a control for permanent income

5) Can therefore obtain a structurally interpretable (λ constant) estimate

TABLE IV

ELASTICITIES: GROUPING INSTRUMENTS: COHORT AND EDUCATION

	Wage	Compensated Wage	Other Income	Group Means: Hours	Wage	Income
No Children	0.140	0.140	0.000	32	2.97	88.63
	(0.075)	(0.088)	(0.041)			
Youngest Child 0–2	0.205	0.301	−0.185	20	3.36	129.69
	(0.128)	(0.144)	(0.104)			
Youngest Child 3–4	0.371	0.439	−0.173	18	3.10	143.64
	(0.150)	(0.159)	(0.139)			
Youngest Child 5–10	0.132	0.173	−0.102	21	2.86	151.13
	(0.117)	(0.127)	(0.109)			
Youngest Child 11 +	0.130	0.160	−0.063	25	2.83	147.31
	(0.107)	(0.117)	(0.084)			

Note: Asymptotic standard errors in parentheses.

Source: Blundell, Duncan, and Meghir 1998

Blundell, Duncan, and Meghir: Results

1) Compensated wage elasticities: 0.15-0.3, depending on number of kids

2) No income effects when no kids, moderate income effects when kids present

3) Identification assumption is common trends across cohort/education groups

4) However, reforms in 80s went in opposite directions at different times $\rightarrow$ Secular trends cannot explain everything

5) See Blundell and MaCurdy (1999) for additional ITLS estimates

Intertemporal Substitution: High Frequency Studies

1) Recent literature focuses on high frequency substitution

2) Focus on groups with highly flexible and well measured labor supply such as:

a) cab drivers [Camerer et al. QJE'97, Farber JPE'05, AER-PP'08, Crawford-Meng '09]

b) stadium vendors [Oettinger JPE'99]

c) cycling messengers [Fehr-Goette AER'07]

Camerer et al. QJE'97

Examine how variation across days in wage rate for cab drivers (arising from variation in waiting times) correlates with hours worked

a) Striking finding: strong negative effect

b) Interpret this as "target earning" – strongly contradicts standard intertemporal labor supply model

c) Would suggest very counter intuitive effects for temporary tax changes, etc.

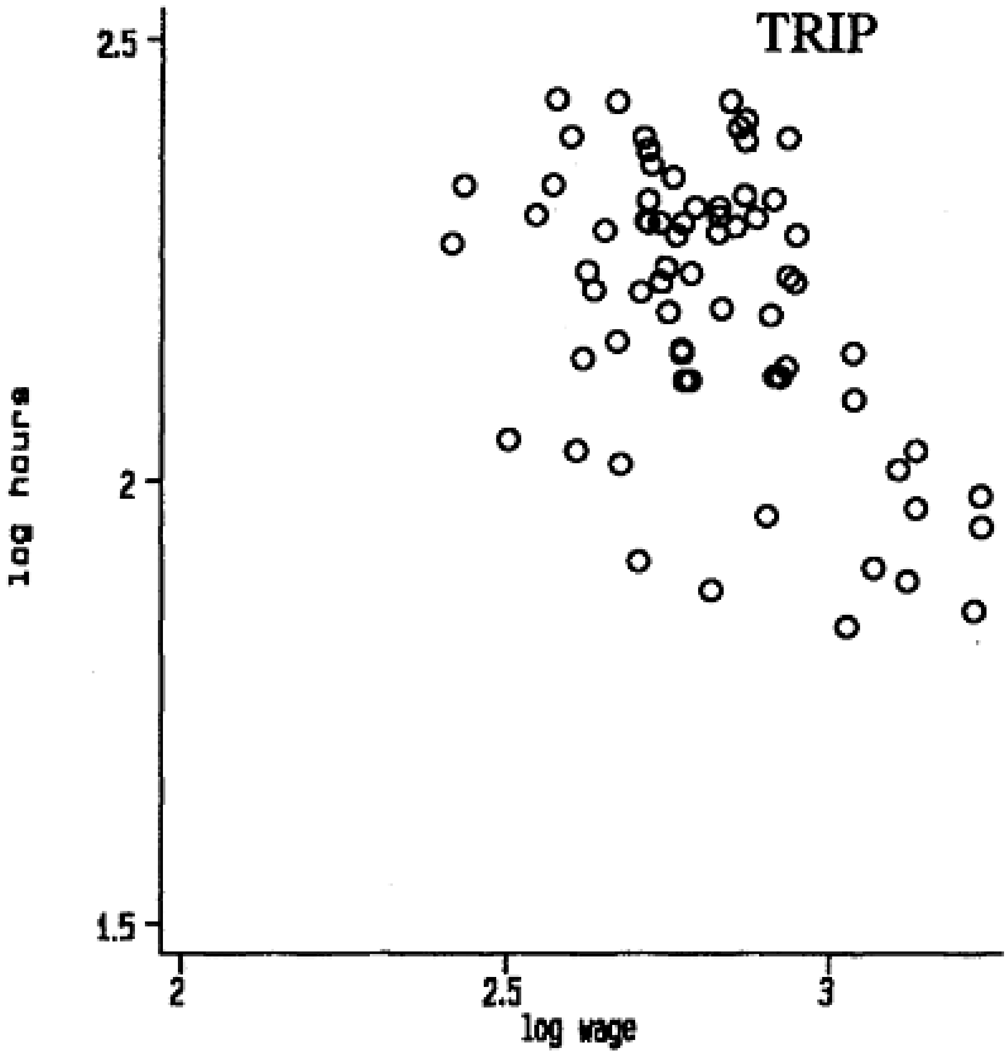


FIGURE I
Hours-Wage Relationships

Source: Camerer et al. 1997

TABLE II

OLS LOG HOURS WORKED EQUATIONS

Sample	TRIP		TLC1		TLC2
Log hourly wage	−.411	−.186	−.501	−.618	−.355
	(.169)	(.129)	(.063)	(.051)	(.051)
High temperature	.000	−.000	.001	.002	−.021
	(.002)	(.002)	(.002)	(.002)	(.007)
Shift during week	−.057	−.047	−.004	.030	—
	(.019)	(.033)	(.035)	(.042)	
Rain	.002	.015	—	—	−.150
	(.035)	(.035)			(.062)
Night shift dummy	.048	−.049	−.127	−.294	−.253
	(.053)	(.049)	(.034)	(.047)	(.038)
Day shift dummy	—	—	.000	.053	—
			(.028)	(.045)	
Fixed effects	No	Yes	No	Yes	No
Adjusted R^2	.243	.484	.175	.318	.146
Sample size	70	65	1044	794	712
Number of drivers	13	8	484	234	712

Dependent variable is the log of hours worked. Standard errors are in parentheses and are corrected for the nonfixed effects estimates in coulmns 1 and 3 to account for the panel structure of the data. Explanatory variables are described in Appendix 1.

Source: Camerer et al. 1997

Farber: Division Bias

Argues that Camerer et al. evidence of target earning behavior is driven by econometric problems

Camerer et al. regression specification:

$$h_{it} = \alpha + \beta e_{it}/h_{it} + \varepsilon_{it}$$

Camerer et al. recognize this and try to instrument using average daily wage $\bar{w}_t$ across all drivers

But there may be a random component to hours at the group level (e.g., good weather makes job more pleasant $\Rightarrow$ more hours and smaller wages at the group level)

$\Rightarrow$ Spuriously find a negative association between average daily wage and average hours

Farber: Within-Day Volatility

Farber's alternative test for target earnings: hazard model

$$Quit = f(cum\_hours, cum\_inc)$$

Result: main determinant of quitting is hours worked (fatigue), NOT cumulative income $\Rightarrow$ Rejects target earning, but does not yield ITLS estimate

Two other studies find positive ITLS:

a) Bicycle messengers (randomized experiment with 25% wage subsidy for 4 weeks): work more days and earn more when wages higher but effort per day $\downarrow$ [fatigue effect]

b) Baseball stadium vendors (work more in high attendance games)

But structural parameters estimated in these studies are not of direct interest to public finance because they are too high frequency

Macro Evidence

1) Macroeconomists estimate/calibrate 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) But highly influential in calibration of macroeconomic models

Trend-based Estimates and Macro Evidence

Long-Run: US wage rates multiplied by about 10 from 1900 to present due to economic growth

Male hours have fallen slightly and then stabilized

$\Rightarrow$ Uncompensated hours of work elasticity is small

However, taxes are rebated as transfers so can still have labor supply effect of taxes if compensated elasticity is large

Short-Run: Hours worked are strongly pro-cyclical [unemployment in recessions and overtime in booms]

Real business cycles do not have involuntary unemployment [questionable assumption] $\Rightarrow$ Variation in hours due to labor supply $\Rightarrow$ Frisch elasticity must be very large for macro-models to work

Prescott 2005

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 2

Actual and Predicted Labor Supply

In Selected Countries in 1993–96 and 1970–74

Period	Country	Labor Supply*		Differences (Predicted Less Actual)	Prediction Factors	
		Actual	Predicted		Tax Rate τ	Consumption/ Output (c/y)
1993–96	Germany	19.3	19.5	.2	.59	.74
	France	17.5	19.5	2.0	.59	.74
	Italy	16.5	18.8	2.3	.64	.69
	Canada	22.9	21.3	–1.6	.52	.77
	United Kingdom	22.8	22.8	0	.44	.83
	Japan	27.0	29.0	2.0	.37	.68
	United States	25.9	24.6	–1.3	.40	.81
1970–74	Germany	24.6	24.6	0	.52	.66
	France	24.4	25.4	1.0	.49	.66
	Italy	19.2	28.3	9.1	.41	.66
	Canada	22.2	25.6	3.4	.44	.72
	United Kingdom	25.9	24.0	–1.9	.45	.77
	Japan	29.8	35.8	6.0	.25	.60
	United States	23.5	26.4	2.9	.40	.74

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

Source: Prescott (2004)

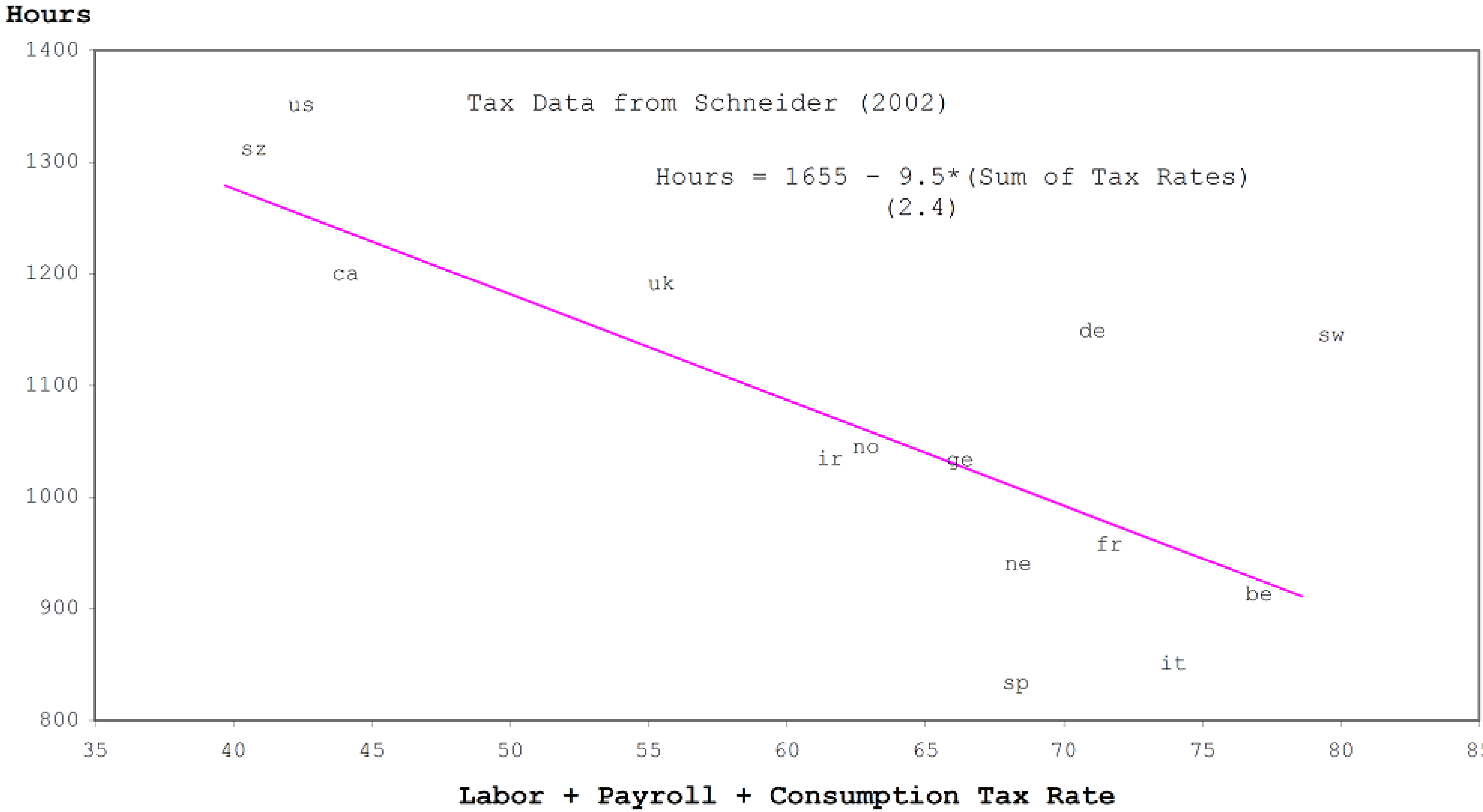
Davis and Henrekson 2005

Run regressions of hours worked on tax variables with various controls

Some panel evidence, but primarily cross-sectional

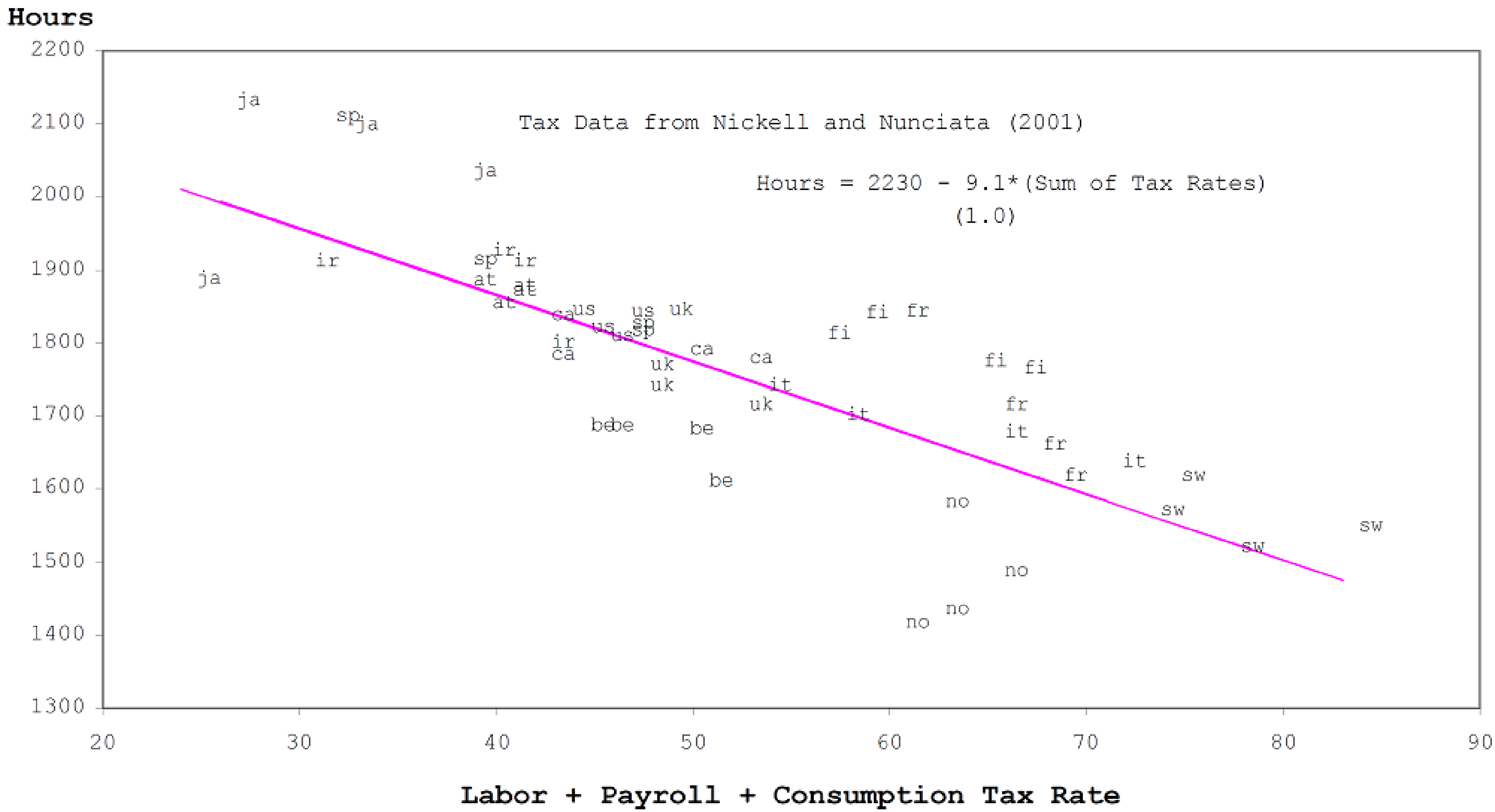
Separate intensive and extensive margin responses

Figure 1: Tax Rates and Annual Work Hours Per Adult
Sample D: 14 Countries in 1995



Source: Davis and Henrekson 2005

Figure 2: Tax Rates and Annual Hours Per Employed Person
Sample A: 13 Countries with Data for 1977, 1983, 1990 and 1995



Source: Davis and Henrekson 2005

Reconciling Micro and Macro Estimates

Recent interest in reconciling micro and macro elasticity estimates

Three potential explanations

a) Statistical Bias: regulations, culture differs in countries with higher tax rates [Alesina, Glaeser, Sacerdote 2005]

b) Extensive vs. Intensive margin [Rogerson and Wallenius 2008]

$$
\begin{aligned}
L &= Nh \\
\frac{d\log L}{d(1-\tau)} &= \frac{d\log N}{d(1-\tau)} + \frac{d\log h}{d(1-\tau)} > \frac{d\log h}{d(1-\tau)}
\end{aligned}
$$

c) Other programs: retirement, education affect labor supply at beginning and end

Blundell (Mirrlees Review)

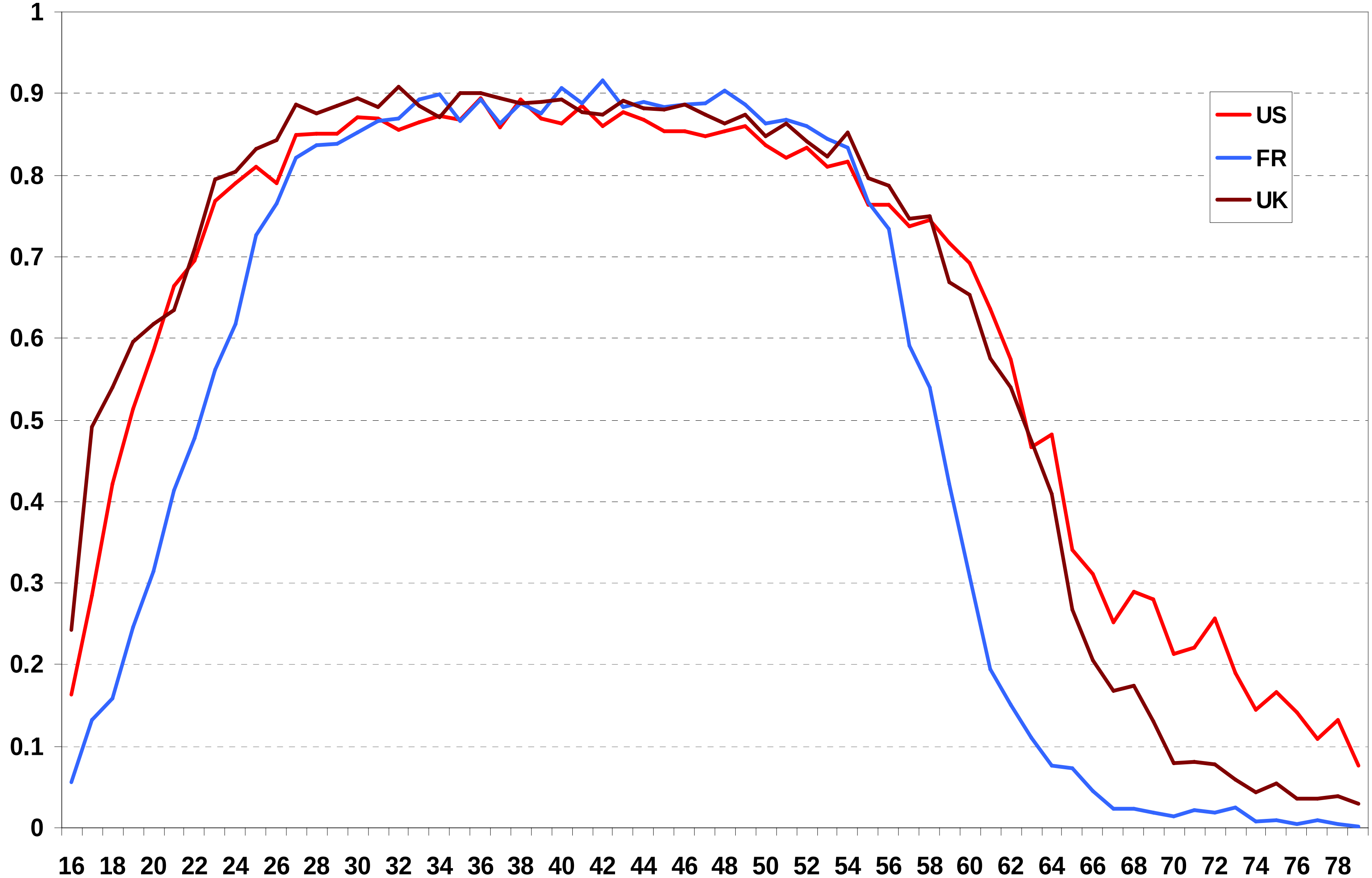
Strong evidence that variation in aggregate hours of work across countries happens among the young and the old: (a) schooling-work margin (b) presence of young children (for women), (c) early retirement

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]

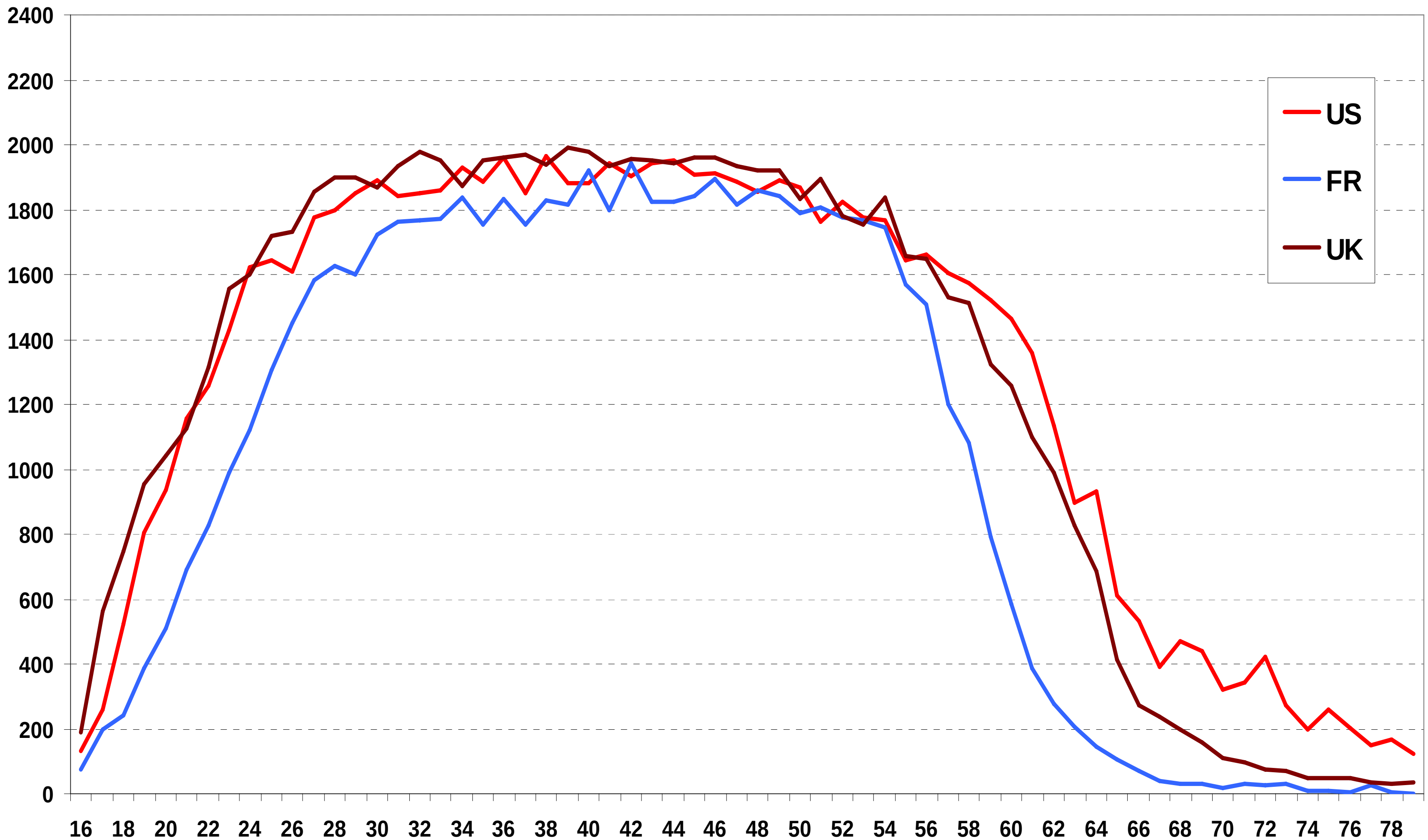
This has been done quite successfully in the case of retirement by series of books by Gruber and Wise, *Retirement around the world*

$\Rightarrow$ Need to develop a more comprehensive international / time series database of tax wedges by age and family types

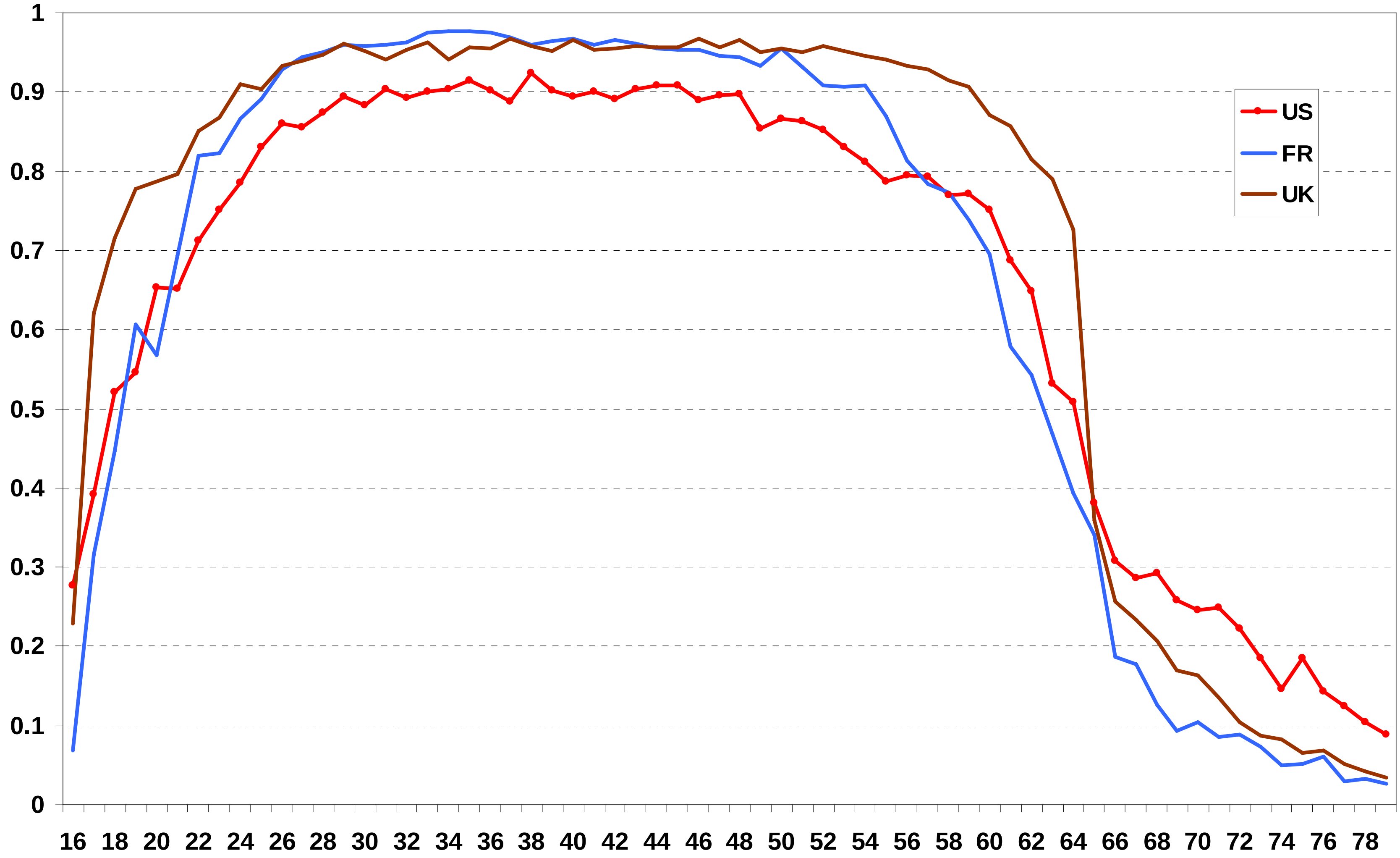
Male employment by age – US, FR and UK 2005



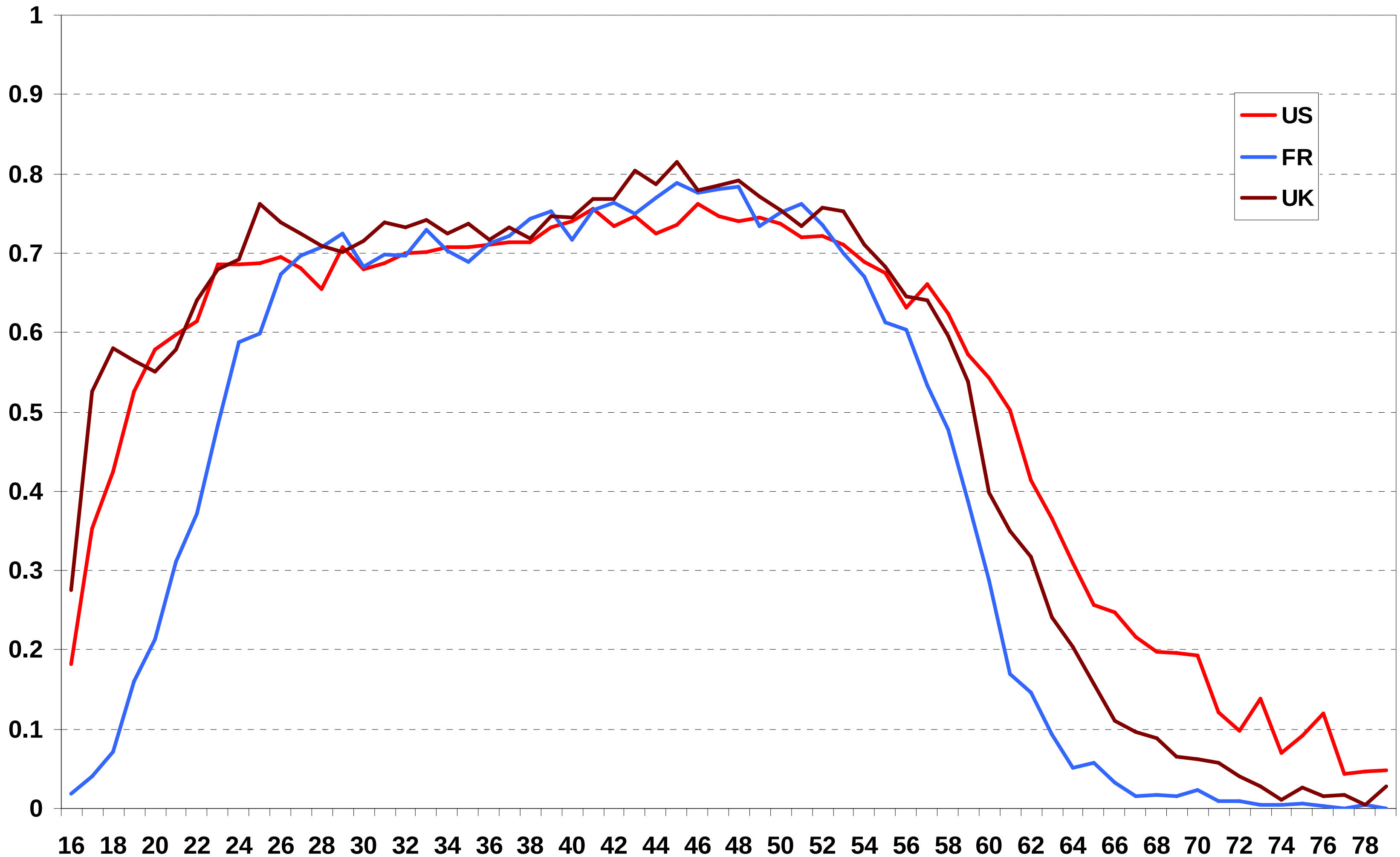
Male Hours by age – US, FR and UK 2005



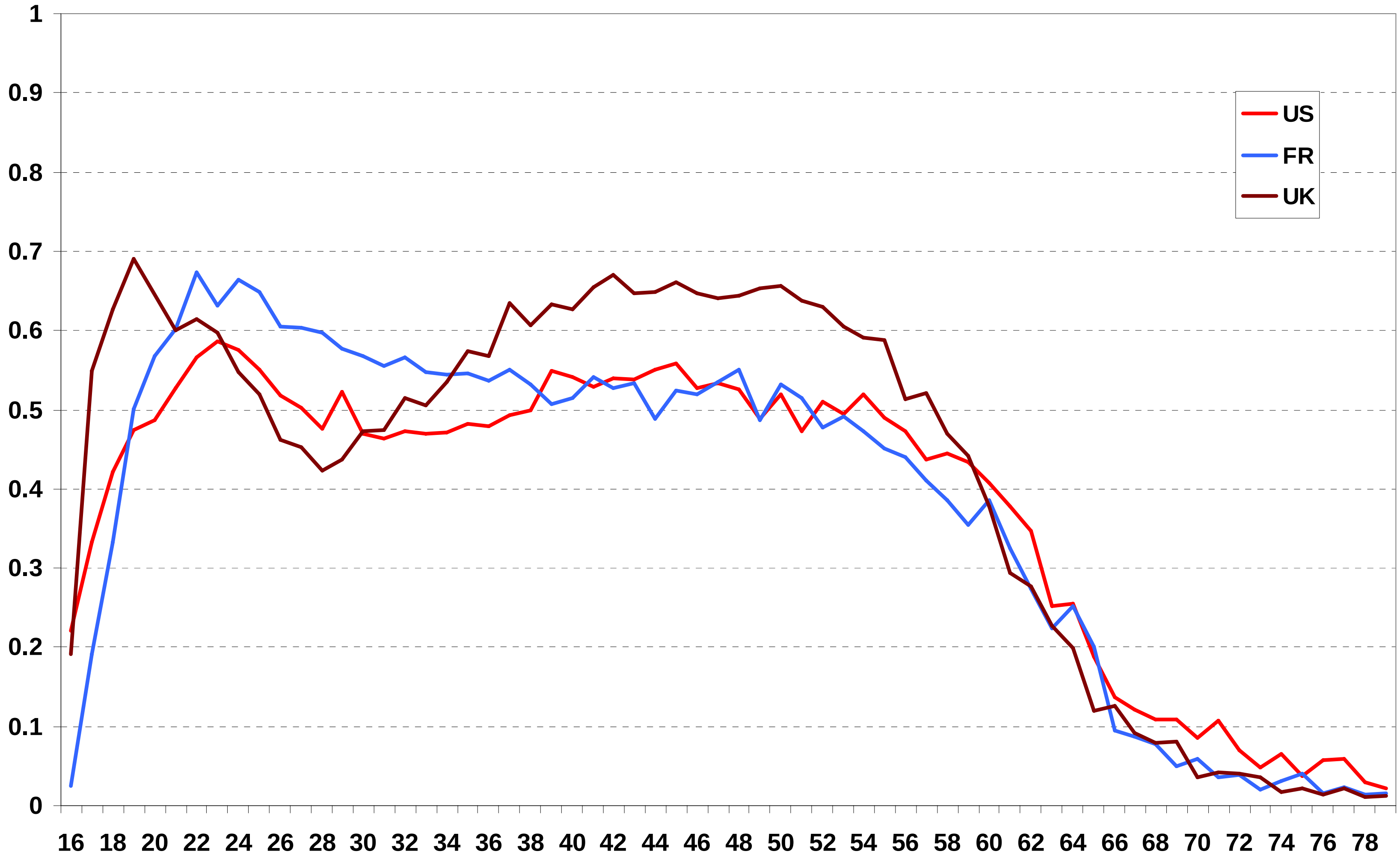
Male employment by age – US, FR and UK 1975



Female Employment by age – US, FR and UK 2005



Female Employment by age – US, FR and UK 1975



Female Hours by age – US, FR and UK 2005