Evaluating Economic Policy: Theoretical and Empirical Lessons From the Literature

James J. Heckman,
University of Chicago
Econ 312, March 27, 2006

I. Introductory Remarks

1. The Evaluation of Economic Policy is a Central Problem in the Study of Economics

    a. The Seminal Text in the Field of Economics - Adam Smith's *An Inquiry Into the Nature and Causes of the Wealth of Nations* - Was About Policy Evaluation

b. Smith Was Very Clear in His Vision About What Good Policy Evaluation Is All About and His Lessons Remain Relevant Today.

i. Understand the Basic Economic Forces (Model the Economy)

ii. Then Understand How Policy Acts to Enhance or Thwart These Economic Forces

iii. Then Evaluate the Consequences for Winners and Losers (Conduct Welfare Economics).

2. An Essential Activity in Policy Evaluation is the Construction of Counterfactual States, Understanding What the World Would Be Like If Alternative Policies Had Been Tried.

    a. This Requires Models of the Economy that are Stable (Parameters Invariant to Intervention).

"The existence of a problem in knowledge depends on the future being different from the past, while the possibility of a solution of the problem depends on the future being like the past."

— Frank Knight (1921, p. 313)

b. With Stable Models, We Examine How Changes in Inputs Affect Outcomes, How Policies Affect Outcomes and Welfare.

3. Smith and Many of His Successors Focused on Understanding Economic Forces *Qualitatively*. How the Social System Worked. How Competition, Self-Interest and Trade Fostered the Creation of Wealth and Elevated Well Being.

4. For More Than 150 Years, the Field of Economics Was Concerned With the Qualitative Evaluation of Policy.

5. Theory Developed that Captured Central Features of the Economic System in a Qualitative Way.

6. In Fact, Until the 20th Century, With Rare Exceptions, the Theory Was Qualitative. Data Were Scarce. Economics Was Not Able to Make Quantitative Distinctions. The Magnitudes of the Effects of Policies Were Unknown

"When you cannot measure, your knowledge is meager and unsatisfactory"

—— Lord Kelvin (1889)

7. Our Knowledge of the Quantitative Importance of Various Economic Forces and the Effects of Policies is a Fairly Recent Development and Has Made Economics Into a Scientific Activity.

8. Milestones:

a. Development of Impartial Data Collection (Kuznets: Goldsmith),

Data Free of Political Manipulation, A Touchstone to Which All Parties in a Debate Can Turn—Verifiable By All Sides.

Transparency, Verifiability of Data and Independence of Statistical Collection Agencies is Central to Successful Economic Policy Making. This Feature is Rare in Most Countries Around the World.

9. What is Needed is an Effective, Objective Cost Benefit Analysis that Measures the Full Costs and Benefits of Public Policy—the Distribution of Benefits.

10. Examples of the Benefits of a Transparent and Verifiable Data Base

a. i. Welfare Reform in the U.S.

ii. Understanding the Incentive Consequences of Unemployment Insurance

iii. Understanding Unionism, Severance Payments, Costs of Regulation.

iv. Understanding the Effectiveness, or Lack of Effectiveness, of "Active Labor Market Policy"

b. Allows For Objectivity (A Touchstone For All Sides)

c. Helps Take Politics Out of the Economic Policy Process, Resolves the Part of a Problem that is Objective. Reduces the Zone of Disagreement.

11. For Evaluation of Many Policies: Health, Education, Job Training, Road Improvements, It is Much More Informative to Use Microdata—Data on Individuals, Families, Firms and Institutions.

12. Macrodata—Too Few Degrees of Freedom, Too Many Programs and Activities Going on to Isolate the Impact of Any One Activity.

13. Honest Data on Outcomes are Necessary But By No Means Sufficient, Need Cost Data, Discount Rates, *etc*.

14. What are the Requirements of a Sound Policy Evaluation System?

    a. Good Data (Objective)

    b. Adequate Methodology For Analyzing the Data

15. Need Tools to Construct Counterfactual Models.

— Stable Models, Invariant to Policy Change

16. Consider Person ω and Two Policies "1" and "0".

We Ideally Seek to Estmate the Policy Counterfactual Gain

$$Y_1(\omega) - Y_0(\omega)$$

To Person ω From Moving From "0" To "1".

17. If “1” is the New Policy, Never Previously Experienced, We Need a Model to Forecast It. This Involves Accounting for All of the Interactions Among Agents.

    Very Hard Problem at the Macro Level. Need Macro Models That are Stable.

From Macro Means Cannot Compute Distributions of Impacts.

If We Have Experienced “1” Before, Might Look to Economic History to Find “Reasonably Close” Episodes.

18. A Major Solution to this Problem in the Modern Literature: Understanding Aggregates By Understanding Individuals, Households and Firms, the Real Economic Actors at the Micro Level.

a. Failure of the Representative Agent Paradigm Both in Theory and in the Sense of Statistical Fit

b. Understanding Heterogeneity and Diversity and the Role of Uncertainty are Central Tasks

c. Implications and Findings From the Literature

i. Understanding Aggregate Labor Supply as Choice at the Extensive Margin.

ii. Understanding Wealth Holding and Wealth Creation (Savings and Critical role of Entrepreneurship), and The Role of Credit Markets as in the Work of Townsend on Thailand.

iii. Unemployment as a Heterogeneous Phenomenon, Differs Across People.

d. Heterogeneity is Empirically Important.

e. Problems with Self Selection, Measurement Error and Reverse Causation Remain in Using Micro Data.

f. Need to Aggregate the Components Back to General Equilibrium.

19. Social Experiments: The Application of Clinical Trials Models to Economic Data

— Widely Thought to Solve Selection Problems

— Will Present Evidence on Their Strengths and Limitations

20. Integration of Micro and Macro: A Modern Synthesis.

This is the future. Use economics to synthesize the evidence. Data are very valuable. But blind empiricism does not work. Need to organize the evidence with models. Almost all policies are new or are applied to new environments. Need models to provide reliable guides to forecasting the effects of new policies.

21. Simple Example of the Micro Structural Approach.

Effect of Taxes on Labor Supply

$$H = H(W(1-\tau), X, \varepsilon) \qquad (1)$$

H = Hours of Work

W = Wage

τ = Tax Rate

X = Observed Characteristics

ε = Unobserved Variables (By Analyst)

$$W(1-\tau) = \text{After Tax Wage}$$
$$\text{Let} \quad \omega = (W, X, \varepsilon)$$
$$\text{Let} \quad H = H_\tau(\omega)$$

Policy Change

Implement a Tax τ' on Wages.
Old Tax τ (May Be Zero).

Wage After New Tax $= W(1 - \tau')$

Wage Under Old Tax $= W\,(1 - \tau)$

If We Know (1), Then Even If Taxes Have Never Been Experienced We Can Forecast Their Effect If The Relationship Is Stable.

Values of H at τ' and at τ Can Be Generated. Structural Theory Naturally Generates its Own Treatment and Control Groups.

Problems:

a. Determining (1) May Be Hard (Early Literature on Labor Supply in the 1960s).

b. Extrapolation: H May Be Estimated Over a Different Region Than What Policy Will Affect.

c. General Equilibrium (τ May Affect W Through Markets).

Benefits:

If We Have (1), Can Generate Welfare Measures, Evaluate More Complicated Policies. Can Go Out of Sample to Consider a Variety of Policies Never Tried. (*e.g.* Nonlinear Taxes, Child Care Bonuses and the Like).

22. In Place of Structural Parameters, Various "Treatment Effects" are Often Defined

Treatment Effect: Looks Only at Policies in Place and Does Not Generalize

Perfectly fine if the treatments are exactly the policies. But suppose they are not How to generalize? That is the role of theory and the role of structural econometrics.

$$H(W(1-\tau'), X, \varepsilon) - H(W(1-\tau), X, \varepsilon)$$

$$Y_1(\omega) - Y_0(\omega)$$

Y_1 is H for τ'

Y_0 is H for τ.

$E\left(Y_1 - Y_0 \mid X\right)$ Average Effect

$D = 1$ Denotes "Treated"

$E\left(Y_1 - Y_0 \mid X, D = 1\right)$

Effect for Those Allocated to "1"

$E\left(Y_1 - Y_0 \mid \text{Marginal}\right)$

Effect for People at Margin

(Marginal Gains or Losses are Required for Cost Benefit Analysis.)

Distribution of Effects

Widely Used Methods in Health, Education and Job Training Evaluations.

a. Assume Partial Coverage, Policies Studied Not Universal

Some People Are Treated

$$(D = 1) \quad (\mathcal{T} = \tau')$$

Others Are Not Treated

$$(D = 0) \quad (\mathcal{T} = \tau)$$

b. Also People May Be Followed Over Time:

$(Y_{1t}(\omega), Y_{0t}(\omega))$ (Program records)

t is Time Period.

c. Focusing on Means, There are Three Prototypical Comparisons Made With This Approach:

i. Program Comes at Time Period p

Before-After: For Persons $D = 1$

$$E\left(Y_{1t} - Y_{0t'} \mid X, D = 1\right)$$

$$t > p > t'$$

Potential Problems: Any Individual or Economy-Wide Trends Allocated to Effect.

Focuses on Treatment on Treated as the Parameter of Interest.

ii. Cross Section Comparison

$$t > p$$

$$E\left(Y_{1t} \mid D = 1, X\right) - E\left(Y_{0t} \mid D = 0, X\right)$$

Compares Different People

Potential Problem: Persons Who Have $D = 0$ Different in Y_{0t} Than Those Who Have $D = 1$

$$E\left(Y_{0t} \mid D = 1\right) \neq E\left(Y_{0t} \mid D = 0\right)$$

(Selection Bias)

Bias for Treatment on Treated is

$$E\left(Y_{0t} \mid D = 1\right) - E\left(Y_{0t} \mid D = 0\right)$$

d. Difference in Differences

$$t > p > t'$$

Potential Problem: Common Trends May Not Operate on $D = 1$ Group and $D = 0$ Group.

Bias $=$

$$E\left(Y_{0t} \mid X, D = 1\right) - E\left(Y_{0t'} \mid X, D = 1\right)$$
$$- \left[E\left(Y_{0t} \mid X, D = 0\right) - E\left(Y_{0t'} \mid X, D = 0\right)\right]$$

Growth May Be Different Among Groups.

Why Focus on Treatment on the Treated?

23. Problems With the Clinical Trials Analogy: Atheoretical Experiments and Quasi-experiments Compare "Treated" With "Untreated".

Negative Income Tax Experiments

a. i. Some Get τ

ii. Others Get τ'

Compare Means

$$E(\mathrm{H} \mid \quad \mathcal{T} = \tau) - E(H \mid \quad \mathcal{T} = \tau')$$

b. These Do Not Directly Recover Structural Parameters for Welfare and the Like.

c. Silent About Other Policies (Do Not Extrapolate)

d. Do Not Solve General Equilibrium Problems

e. Do Not Recover Distributions of Effects.

II. Example: Evaluation of Education and Training Policies

1. In Most Modern Societies, Expenditure on Education, Training and Skill Formation Policies are Substantial (3-5% of GDP)

2. Is Expenditure Efficient? Can It Be Improved?

3. Role of Evaluation. Construct Counterfactual States and Evaluate Policies.

4. Structural Approach—Estimate Models of Skill Acquisition

   a. Understand Mechanisms of Skill Formation (On the Job Training *vs.* Learning By Doing).

   b. Use as a Basis for Evaluation of a Wide Array of Policies.

   c. Rich Approach—Produces Invariant Parameters.

5. Treatment Effect Literature

a. Black Box Quality

b. Does a Particular Program Work?

LaLonde Study

An Important Challenge Was Raised by the LaLonde Study (1986)

Compare experimental with nonexperimental estimators using different comparison groups

Widely Interpreted

Conclusion:
"No Non-Experimental Method Works"

Table 1 —ANNUAL EARNINGS OF NSW TREATMENTS, CONTROLS, AND EIGHT CANDIDATE COMPARISON GROUPS FROM THE *PSID* AND THE *CPS-SSA*

Year	Treatments	Controls	Comparison Group[a,b]							
			PSID-1	*PSID*-2	*PSID*-3	*PSID*-4	*CPS-SSA*-1	*CPS-SSA*-2	*CPS-SSA*-3	*CPS-SSA*-4
1975	$895	$877	7,303	2,327	937	6,654	7,788	3,748	4,575	2,049
	(81)	(90)	(317)	(286)	(189)	(428)	(63)	(250)	(135)	(333)
1976	$1,794	$646	7,442	2,697	665	6,770	8,547	4,774	3,800	2,036
	(99)	(63)	(327)	(317)	(157)	(463)	(65)	(302)	(128)	(337)
1977	$6,143	$1,518	7,983	3,219	891	7,213	8,562	4,851	5,277	2,844
	(140)	(112)	(335)	(376)	(229)	(484)	(68)	(317)	(153)	(450)
1978	$4,526	$2,885	8,146	3,636	1,631	7,564	8,518	5,343	5,665	3,700
	(270)	(244)	(339)	(421)	(381)	(480)	(72)	(365)	(166)	(593)
1979	$4,670	$3,819	8,016	3,569	1,602	7,482	8,023	5,343	5,782	3,733
	(226)	(208)	(334)	(381)	(334)	(462)	(73)	(371)	(170)	(543)
Number of Observations	600	585	595	173	118	255	11,132	241	1,594	87

Table 1 (cont.)

[a] The Comparison Groups are defined as follows: *PSID*-1: All female household heads continuously from 1975 through 1979, who were between 20 and 55-years-old and did not classify themselves as retired in 1975; *PSID*-2: Selects from the *PSID*-1 group all women who received AFDC in 1975; *PSID*-3: Selects from the *PSID*-2 all women who were not working when surveyed in 1976; *PSID*-4: Selects from the *PSID*-1 group all women with children, none of whom are less than 5-years-old; *CPS-SSA* − 1: All females from Westat *CPS-SSA* sample; *CPS-SSA*-2: Selects from *CPS-SSA*-1 all females who received AFDC in 1975; *CPS-SSA*-3: Selects from *CPS-SSA*-1 all females who were not working in the spring of 1976; *CPS-SSA*-4: Selects from *CPS-SSA*-2 all females who were not working in the spring of 1976.

[b] All earnings are expressed in 1982 dollars. The numbers in parentheses are the standard errors. For the NSW treatments and controls, the number of observations refer only to 1975 and 1979. In the other years there are fewer observations, especially in 1978. At the time of the resurvey in 1979, treatments had been out of Supported Work for an average of 20 months.

Source: Lalonde, Robert. (1986). "Evaluating the Econometric Evaluations of Social Programs with Experimental Data." American Economic Review. 76(4): 604-620

Table 2 —Earnings Comparisons and Estimated Training Effects for the NSW AFDC Participants Using Comparison Groups From the *PSID* and the *CPS-SSA*[a,b]

Name of Comparison Group[d]	Comparison Group Earnings Growth 1975–79 (1)	NSW Treatment Earnings Less Comparison Group Earnings: Pre-Training Year, 1975 Unadjusted (2)	Pre-Training Year, 1975 Adjusted[c] (3)	Post-Training Year, 1979 Unadjusted (4)	Post-Training Year, 1979 Adjusted[c] (5)	Difference in Differences: Difference in Earnings Growth 1975–79 Treatments Less Comparisons: Without Age (6)	With Age (7)	Unrestricted Difference in Differences: Quasi Difference in Earnings Growth 1975–79: Unadjusted (8)	Adjusted[c] (9)	Controlling for All Observed Variables and Pre-Training Earnings: Without AFDC (10)	With AFDC (11)
	2,942	−17	−22	851	861	833	883	843	864	854	–
Controls	(220)	(122)	(122)	(307)	(306)	(323)	(323)	(308)	(306)	(312)	
PSID-1	713	−6,443	−4,882	−3,357	−2,143	3,097	2,657	1746	1,354	1664	2,097
	(210)	(326)	(336)	(403)	(425)	(317)	(333)	(357)	(380)	(409)	(491)
PSID-2	1,242	−1,467	−1,515	1,090	870	2,568	2,392	1,764	1,535	1,826	–
	(314)	(216)	(224)	(468)	(484)	(473)	(481)	(472)	(487)	(537)	
PSID-3	665	−77	−100	3,057	2,915	3,145	3,020	3,070	2,930	2,919	–
	(351)	(202)	(208)	(532)	(543)	(557)	(563)	(531)	(543)	(592)	
PSID-4	928	−5,694	−4,976	−2,822	−2,268	2,883	2,655	1,184	950	1,406	2,146
	(311)	(306)	(323)	(460)	(491)	(417)	(434)	(483)	(503)	(542)	(652)

Table 2 (cont.)

CPS-SSA-1	233	−6,928	−5,813	−3,363	−2,650	3,578	3,501	1,214	1,127	536	1,041
	(64)	(272)	(309)	(320)	(365)	(280)	(282)	(272)	(309)	(349)	(503)
CPS-SSA-2	1,595	−2,888	−2,332	−683	−240	2,215	2,068	447	620	665	–
	(360)	(204)	(256)	(428)	(536)	(438)	(446)	(468)	(554)	(651)	
CPS-SSA-3	1,207	−3,715	−3,150	−1,122	−812	2,603	2,615	814	784	−99	1,246
	(166)	(226)	(325)	(311)	(452)	(307)	(328)	(305)	(429)	(481)	(720)
CPS-SSA-4	1,684	−1,189	−780	926	756	2,126	1,833	1,222	952	827	–
	(524)	(249)	(283)	(630)	(716)	(654)	(663)	(637)	(717)	(814)	

[a] The columns above present the estimated training effect for each econometric model and comparison group. The dependent variable is earnings in 1979. Based on the experimental data, an unbiased estimate of the impact of training presented in col. 4 is $851. The first three columns present the difference between each comparison group's 1975 and 1979 earnings and the difference between the pre-training earnings of each comparison group and the NSW treatments.

[b] Estimates are in 1982 dollars. The numbers in parentheses are the standard errors.

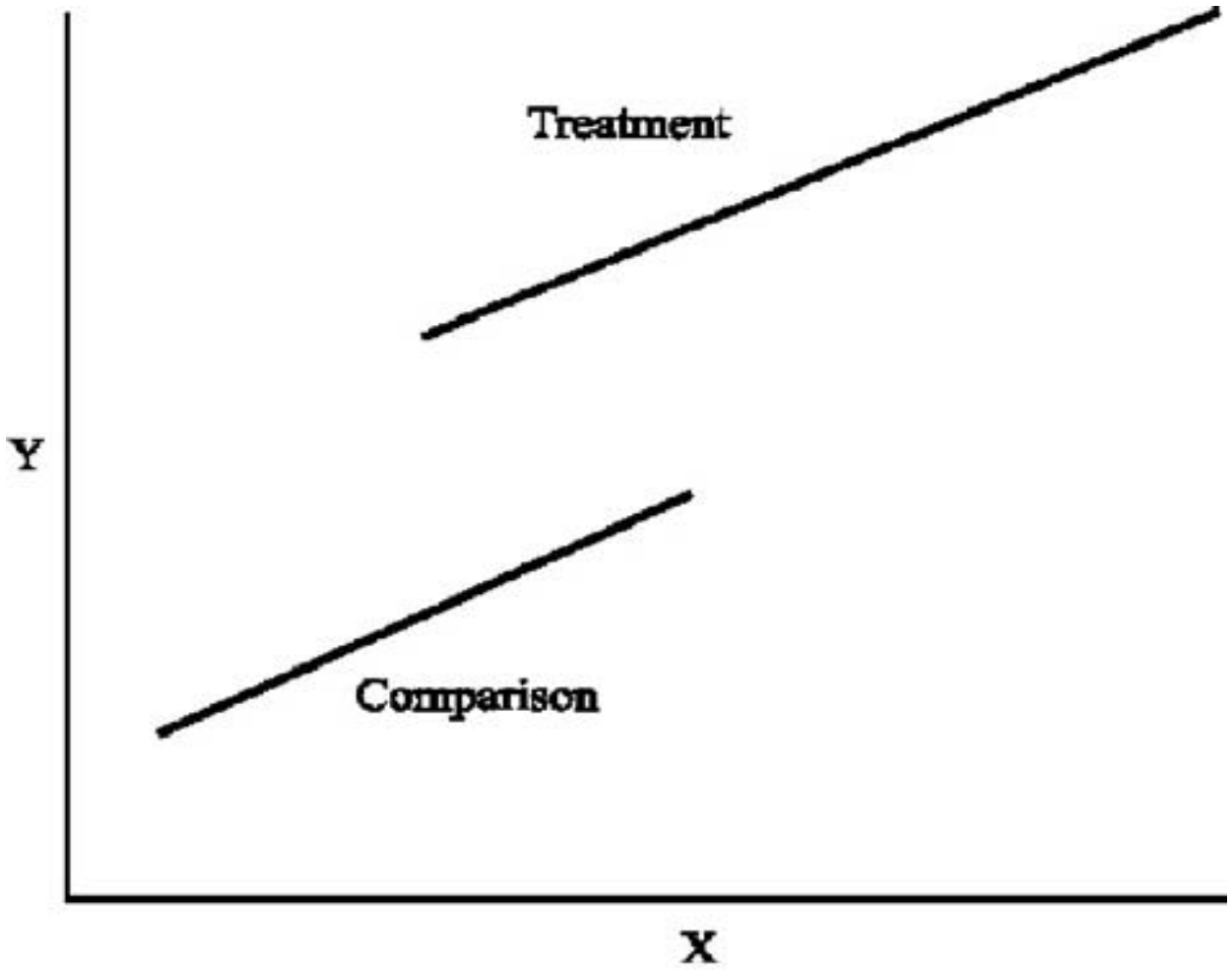
[c] The exogenous variables used in the regression adjusted equations are age, age squared, years of schooling, high school dropout status, and race.

[d] See Table 2 for definitions of the comparison groups.

Source: Lalonde, Robert. (1986). "Evaluating the Econometric Evaluations of Social Programs with Experimental Data." American Economic Review. 76(4): 604-620

Heckman-Hotz (1989)
"Specification Tests Eliminate Bad Estimators"

Heckman, Ichimura, Smith and Todd
Collect Better Data
Find Empirical Sources of Selection Bias



Sources of Selection Bias

(a) Lack of Overlap of Match Variables

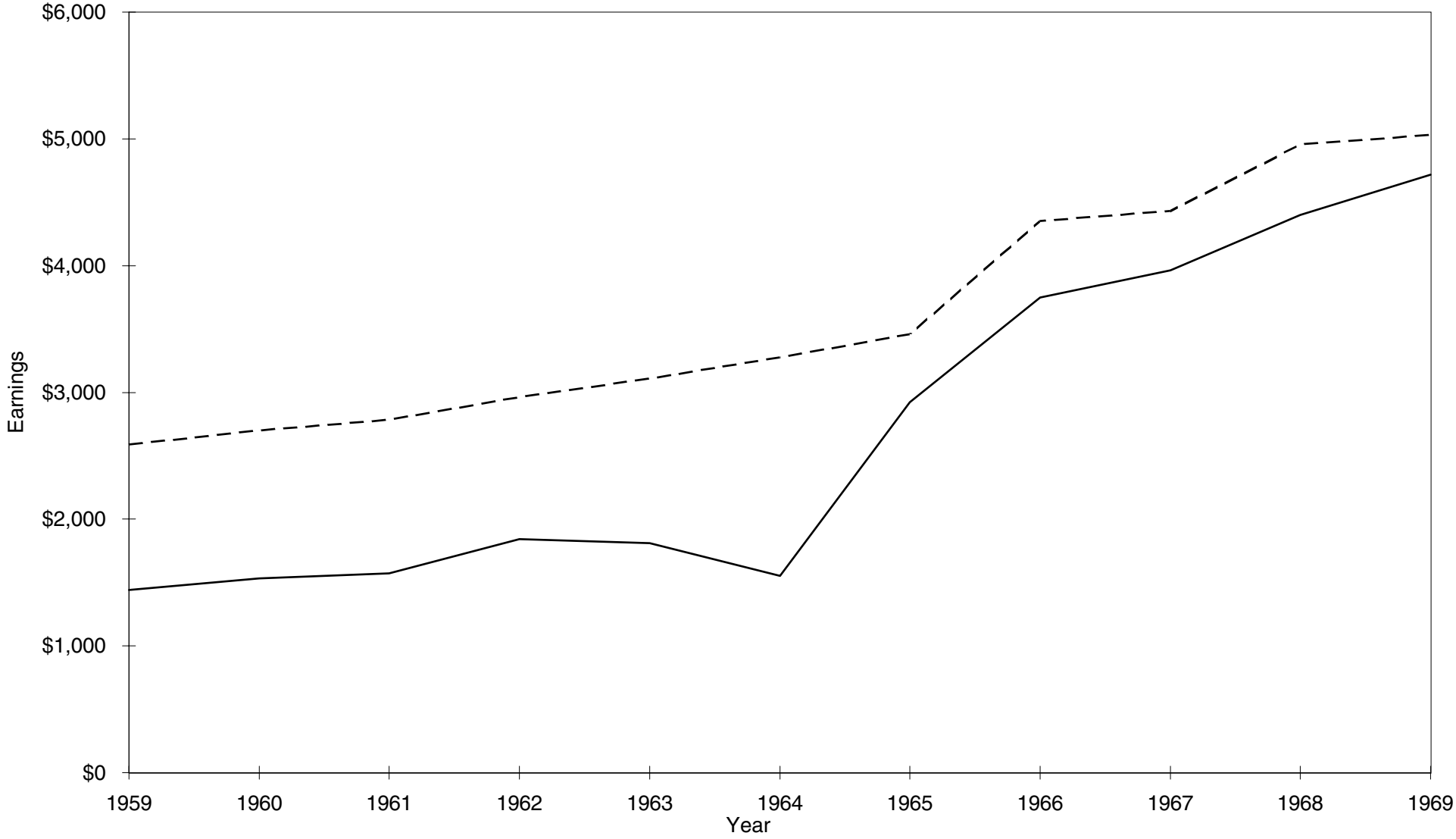
(b) Inappropriate Weighting

Examples From JTPA Model. (Ashenfelter Dip and JTPA Evidence)

- Ashenfelter Data (Figure 1)

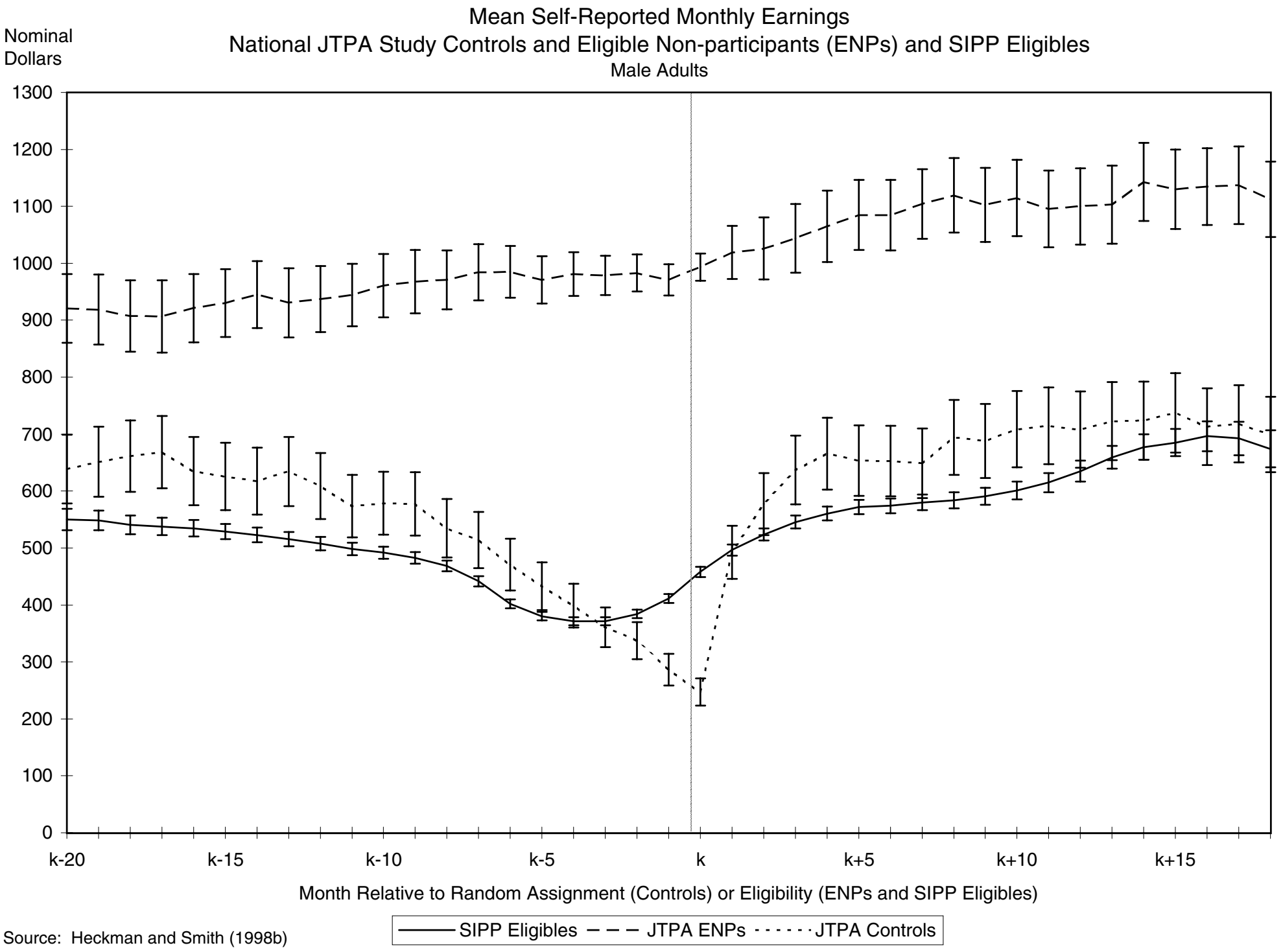
Figure 1

Mean Annual Earnings Prior, During, and Subsequent to Training for 1964 MDTA Classroom Trainees and a Comparison Group: White Males



Source: Ashenfelter (1978).

Figure 2



Source: Heckman and Smith (1998b)

Problems With Support

Real Source of Bias:

Comparing the Noncomparable

TABLE 3A
Explanatory Variables Used in Previous Studies

	CLMS-Based Studies	NSW (Supported Work) Data		NJS Data
Study	Bassi (1983, 1984)	LaLonde (1986)	Fraker and Maynard (1987)	Heckman, Ichimura, Smith and Todd (1998)
Program and Outcome Variable	CETA, 1977 and 1978 annual social security earnings	NSW, 1978 annual social security and PSID survey	NSW, 1977, 1978 and 1979 annual earnings for AFDC women and youth	NJS, monthly survey earnings in the 18 months after random assignment or measured eligibility
Local Labor Market Information	No	No	No	Yes
Age, Race, Sex	Yes	Yes	Yes	Yes
Education	?	Yes	Yes	Yes
Training History	No	No	No	Yes (partial)
Children	?	Yes	Yes	Yes
Employment Histories	?	No	No	Yes
Hours Worked	?	No	No	Yes
Unemployment Histories	?	No	No	Yes (one year)
Welfare Receipt	?	Yes	Yes	Yes (one year)
Earnings Histories	Same as Bryant and Rupp (1987)	Two years pre-program Two years post-program	Two years pre-program Two years post-program	Five years pre-program Two years post-program

Table 3A (cont.)

Matching Criteria (Criteria for Membership in Comparison Sample)	Same as Bryant and Rupp (1987); also uses a random sample from the March 1976 CPS.	PSID: household head in 1975-1979. CPS: March 1976 CPS earnings matched to SSA earnings; screens based on 1976 personal and household income.	Three samples: I: eligible in sample period; screen out in-school youth; AFDC women match on age of youngest child and welfare receipt. II: eligible in sample period; cell matching based on predictors of 1979 SS earnings including prior earnings, change in earnings, education, family income, and demographics. III: eligible in sample period; match on earnings estimated on eligible non-participant sample, age, and sex.	Within age and sec groups, match on propensity score based on site, race, age, schooling, marital status, labor force status history, number of recent jobs, training history, house-hold size, and house-hold recent earnings.

Table 3A (cont.) Source: Heckman, Ichimura and Todd (1997), Table 1.

Notes:

1. "Used" means employed in the matching process and/or included in the outcome equation.

2. CLMS = CETA Longitudinal Manpower Survey. The CLMS data matched Social Security longitudinal earnings records to CETA participants and CPS comparison group members from the March 1976 and 1977 CPS. All of the CLMS-based studies use the social security earnings data except for Bassi (1983), which also uses the CPS earnings data. All of the personal and family information available in the CPS, including short-term employment and labor force participation histories, are available in the CLMS but not necessarily used in the analyses based upon it.

3. "Matching Criteria" indicate the criteria for membership in the comparison group. This is sometimes referred to as "screening" in the literature.

4. "?" indicates that the study does not specify the variables used.

The conventional measure of selection bias

$$B = E(Y_0 \mid D = 1) - E(Y_0 \mid D = 0)$$

(LaLonde,1986, Ashenfelter, 1978)

The conventional measure of bias B

$$Y_0 = \pi_0 + \pi_1 C + \tau, \text{ and } E(\tau) = 0,$$

$$\text{where plim } \widehat{\pi}_1 = B$$

$$B =$$

$$\int_{S_{1P}} E(Y_0 \mid P, D = 1) dF(P \mid D = 1)$$

$$- \int_{S_{0P}} E(Y_0 \mid P, D = 0) dF(P \mid D = 0).$$

$$B = B_1 + B_2 + B_3$$

$$B_1 = \int_{S_{1P} \setminus S_P} E\left(Y_0 \mid P, D = 1\right) \, dF\left(P \mid D = 1\right)$$

$$- \int_{S_{0P} \setminus S_P} E\left(Y_0 \mid P, D = 0\right) \, dF\left(P \mid D = 0\right)$$

(Bias From Lack of Overlap)

$$B_2 = \int_{S_P} E(Y_0 \mid P, D = 0) \cdot [dF(P \mid D = 1) - dF(P \mid D = 0)]$$

(Bias From Different Weights)

$$B_3 = PR_P\overline{B}S_P$$

$$PR_P = \int_{S_P} dF(P \mid D = 1)$$

- the proportion of the density of P given $D = 1$

  in the overlap set S_P
- $S_{1P} \backslash S_P$ is the support of P given $D = 1$

  that is not in the overlap set $S_{P,}$
- $S_{0P} \backslash S_P$ is the support of P given $D = 0$

  that is not in the overlap set S_P.

Re-examining the Conventional Measure of Selection Bias

What Do We Know About Sources of Bias?

$$S_{1P} = \{P \mid f(P \mid D = 1) > 0\}$$

the support of P for $D = 1$

$$S_{0P} = \{P \mid f(P \mid D = 0) > 0\}$$

the support of P for $D = 0$

Let $S_P = S_{0P} \cap S_{1P}$:

the mean selection bias $\overline{B}_{S_P}$

$$\overline{B}_{S_P} = \frac{\int_{S_P} B(P) dF(P \mid D = 1)}{\int_{S_P} dF(P \mid D = 1)}.$$

Estimating the Components of B

Orthogonal decomposition

$$E(Y_0 \mid X, D = 1) = E(Y_0 \mid P(X), D = 1) + V$$

$$V = E(Y_0 \mid X, D = 1) - E(Y_0 \mid P(X), D = 1)$$

$$E(V \mid P(X), D = 1) = 0.$$

$$B =$$

$$E(Y_0 \mid D = 1) - E(Y_0 \mid D = 0)$$

$$= B_1 + B_2 + B_3.$$

(Decompose Bias In Terms of $P(X)$).

Estimated Selection Bias Under Alternative Estimators and Sensitivity of Estimates to Alternative Specifications of the Outcome and Participation Equations

TABLE 4

Decomposition of Differences in Mean Earnings for Adult Participants in the U.S. National JTPA Study

[Mean monthly earnings differences between experimental controls and comparison sample of eligible nonparticipants during the 18 months following the baseline in four sites]

	Adult Males	Adult Females
Mean Difference in Earnings = B	-337 (47)	33 (26)
Nonoverlapping Support = B_1	298 (35) [-88]	106 (13) [318]
Different Density Weighting of Propensity Scores = B_2	-659 (42) [195]	-118 (20) [-355]
Selection Bias = B_3	24 (28) [-7]	45 (26) [136]
Average Selection Bias When Matching Only in Regions of Common Support	48	59
Selection Bias as a Percent of Treatment Impact	109	202
Control Group Sample Size	508	696
Comparison Group Sample Size	388	866

Source: Heckman, Ichimura, Smith, and Todd (1996), Table 1, p. 13418.

Notes:

1. The numbers in parentheses are the bootstraped standard errors. They are based on 50 replications with 100% sampling.
2. The numbers in square brackets are the percentage of the mean difference in earnings (row 1) attributable to each component of the bias.

TABLE 5
Decomposition of Differences in Mean Earnings in the U.S. National JTPA Study
[Mean monthly earnings differences during the 18 months following the baseline in four sites]

	Experimental Controls and Treatment Group Dropouts		Experimental Controls and SIPP eligibles.	
	Adult Males	Adult Females	Adult Males	Adult Females
Mean Difference in Earnings = B	29 (38)	9 (23)	-145 (56)	47 (23)
Nonoverlapping Support = B_1	-13 (12) [-45]	1 (6) [9]	151 (30) [-104]	97 (19) [206]
Different Density Weighting of Propensity Scores = B_2	3 (16) [11]	-9 (10) [-99]	-417 (44) [287]	-172 (16) [-367]
Selection Bias = B_3	38 (37) [135]	18 (26) [190]	121 (33) [-83]	122 (15) [260]
Average Selection Bias When Matching Only in Regions of Common Support	42 (40)	20 (29)	192 (57)	198 (26)
Selection Bias as a Percent of Treatment Impact	97%	68%	440%	676%

Source: Heckman, Ichimura and Todd (1997), Table 2.

Notes:

1. Bootstrap standard errors appear in parentheses. They are based on 50 replications with 100% sampling.

Table 5 (cont.)

2. The numbers in square brackets are the percentage of the mean difference in earnings (row 1) attributable to each component of the bias.

3. Treatment group dropouts (or "no-shows") are persons randomly assigned to the experimental treatment group who failed to enroll in JTPA.

4. The SIPP eligibles are persons in the 1998 SIPP full panel who were eligible in month 12 of the 24 month panel using eligibility definition "B" from Devine and Heckman (1994).

FIGURE BIAS - 6

Density of Estimated Probability of Program Participation

For Adult Male Controls and Eligible Nonparticipants in the National JTPA Study

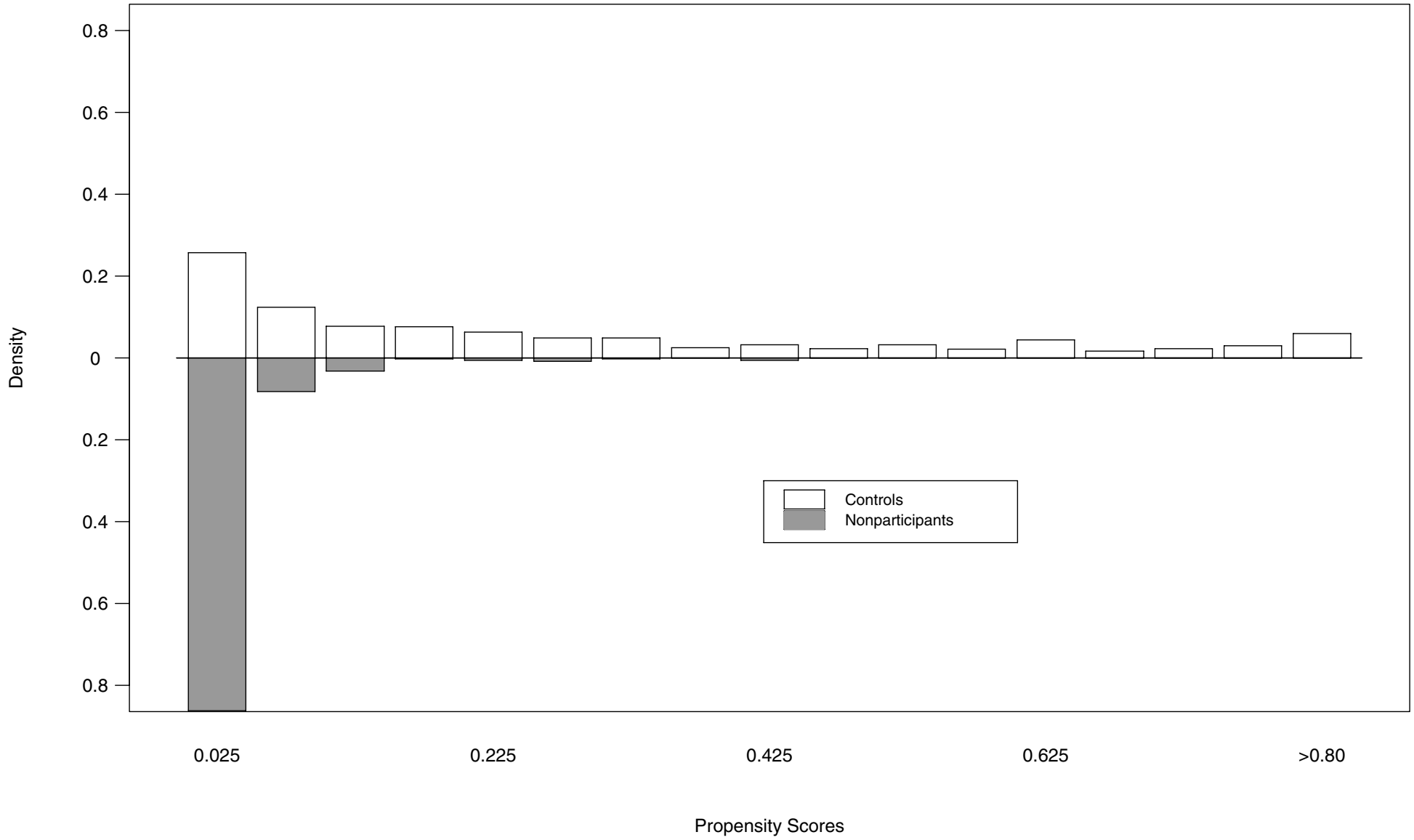


FIGURE BIAS - 7

Density of Estimated Probability of Program Participation

For Adult Female Controls and Eligible Nonparticipants in the National JTPA Study

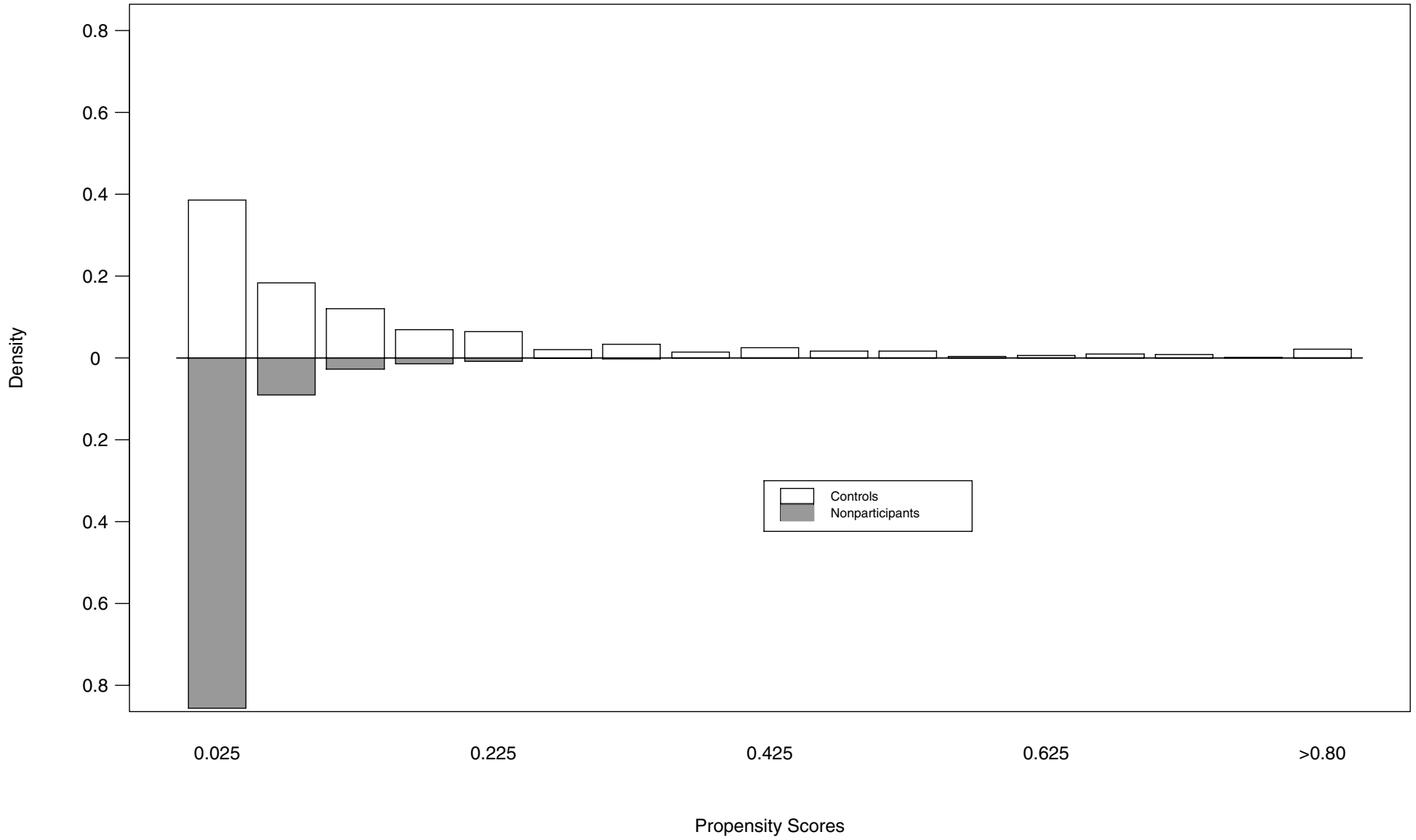


FIGURE BIAS - 8

Density of Estimated Probability of Program Participation

For Adult Male Controls and No-Shows in the National JTPA Study

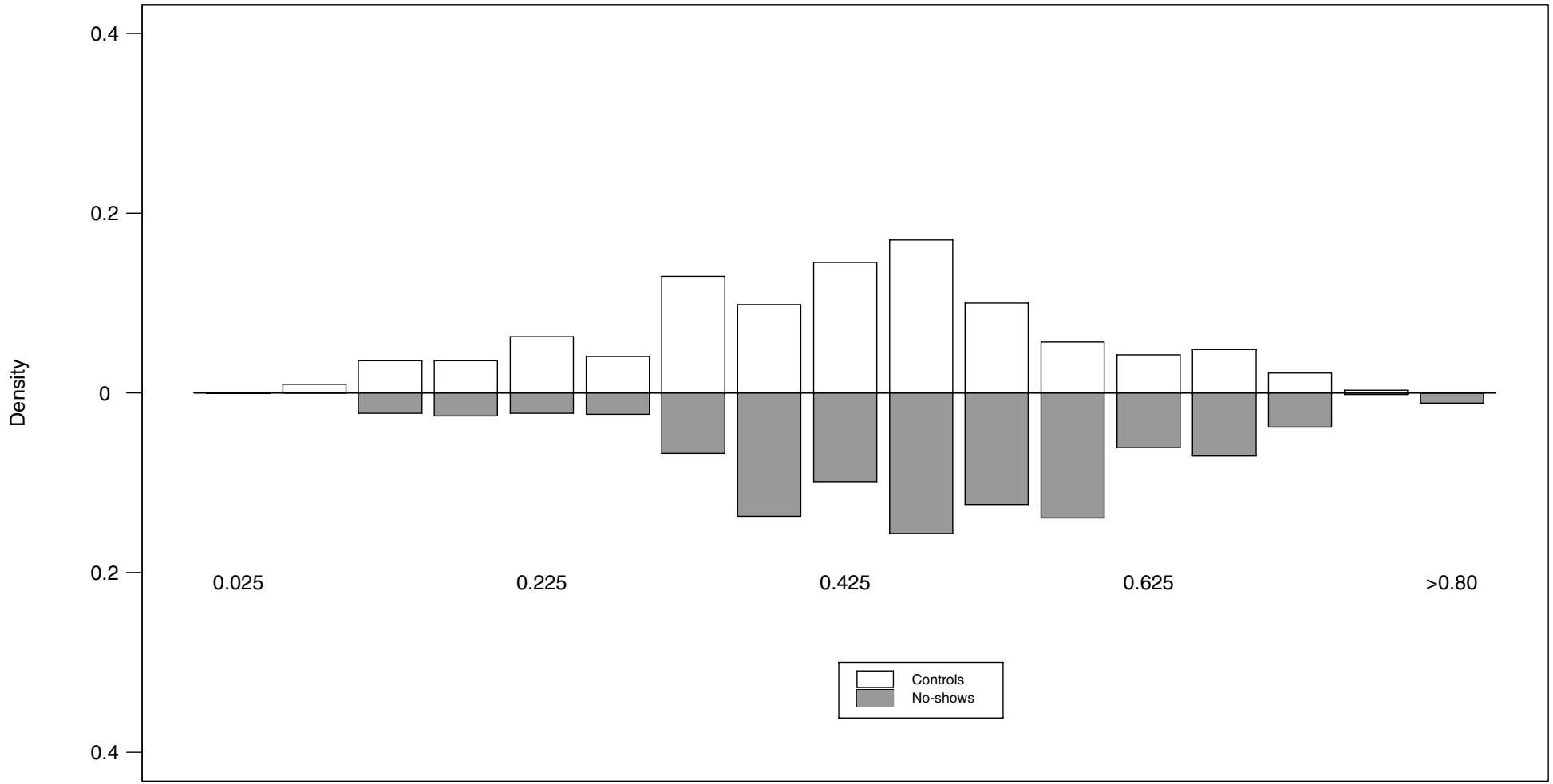


Table 6
Bias Associated with Alternative Matching Estimators
Comparison Group: CPS Adult Male Sample, Dependent Variable: Real Earnings in 1978
(NN = nearest neighbor, CS = common support)

Sample and Propensity Score Model	Mean Diff.	1 NN w/o CS	10 NN w/o CS	1 NN with CS	10 NN with CS	Local Linear (bw = 1.0)	Local Linear (bw = 4.0)
Lalonde Sample	-9757	-555	-270	-838	-1299	-1380	-1431
DW Prop. Score	(255)	(596)	(493)	(628)	(529)	(437)	(441)
as % of $886 impact	-1101%	-63%	-30%	-95%	-147%	-156%	-162%
DW Sample with	-10291	407	-5	-27	-261	-88	-67
-96DW Prop. Score	(306)	(698)	(672)	(723)	(593)	(630)	(611)
as % of $1794 impact	-574%	23%	-0.3%	-1.5%	-15%	-5%	-4%
Early RA sample with	-11101	-7781	-3632	-5417	-2396	-3427	-2191
DW Prop. Score	(461)	(1245)	(1354)	(1407)	(1152)	(1927)	(1069)
as % of $2748 impact	-404%	-283%	-132%	-197%	-87%	-125%	-80%
Lalonde Sample with	-10227	-3602	-2122	-3586	-2342	-3562	-2708
Lalonde Prop. Score	(296)	(1459)	(1299)	(1407)	(1165)	(3969)	(1174)
as % of $886 impact	-1154%	-406%	-240%	405%	264%	402%	306%

Source: Smith, Jeffrey and Petra Todd. (2001). "Reconciling Conflicting Evidence on the Performance of Propensity Score Matching Methods." American Economic Review. 91(2): 112-118

TABLE 7
Descriptive Statistics for Adult Male Experimental and Comparison Group Samples

	NSW Experimental Samples			Comparison Groups	
Variable	**Lalonde**	**Dehejia-Wahba**	**Early Random Assignment**	**CPS sample**	**PSID sample**
Age	24.52	25.37	25.74	33.23	34.85
	(6.63)	(7.10)	(6.75)	(11.05)	(10.44)
Education	10.27	10.2	10.37	12.03	12.12
	(1.70)	(1.79)	(1.6)	(2.87)	(3.08)
Black	0.80	0.84	0.82	0.07	0.25
	(.40)	(0.37)	(0.38)	(0.26)	(0.43)
Hispanic	0.11	0.09	0.10	0.07	0.03
	(0.31)	(0.28)	(.30)	(0.26)	(0.18)
Married	0.16	0.17	0.20	0.71	0.87
	(0.37)	(0.37)	(0.40)	(0.45)	(0.34)
No H.S. Degree	0.78	0.78	0.76	0.30	0.31
	(0.41)	(0.41)	(0.43)	(0.46)	(0.46)
"Real Earnings	3631	2102	3742	14017	19429
in 1974"	(6221)	(5364)	(6718)	(9570)	(13407)
Real Earnings	3043	1377	2415	13651	19063
in 1975	(5066)	(3151)	(3894)	(9270)	(13597)
Real Earnings	5455	5301	5796	14847	21554
in 1978	(6253)	(6632)	(7582)	(9647)	(15555)
Real Earnings	…	…	…	14730	…
in 1979				(11028)	
"Zero Earnings	0.45	0.73	0.524	0.12	0.09
in 1974"	(0.50)	(0.44)	(0.50)	(0.32)	(0.28)
Zero Earnings	0.40	0.65	0.41	0.11	0.10
in 1975	(0.49)	(0.48)	(0.49)	(0.31)	(0.30)
Experimental	886	1794	2748	…	…
Impact (1978	(488)	(670)	(1005)		
earnings)					
Sample Size	297 Treatments	185 Treatments	108 Treatments	15992	2490
	425 Controls	260 Controls	142 Controls		

Notes: Estimated standard deviations in parentheses. Robust standard errors are reported for experimental impact estimates.

Source: Smith, Jeffrey and Petra Todd. "Is Matching the Answer to LaLonde's Critique of Nonexperimental Methods?" Forthcoming, Journal of Econometrics

Do a correct cost benefit analysis

Table 8

E;ects of Accounting for Discounting, Expected Horizon and Welfare Costs of Taxes: Benefit Minus Cost Estimates for JTPA under Alternative Assumptions Regarding Benefit Persistence, Discounting, and Welfare Costs of Taxation (National JTPA Study, 30-Month Impact Sample)

Benefit Duration	Direct Costs Included?	6-Month Interest Rate	Welfare Cost of Taxes	Adult Males	Adult Females	Male Youth	Female Youth
30 Months	No	0.000	0.00	1,345	1,703	-967	136
30 Months	Yes	0.000	0.00	523	532	-2,922	-1,180
30 Months	Yes	0.000	0.50	108	-54	-3,900	-1,838
30 Months	Yes	0.025	0.00	433	432	-2,859	-1,195
30 Months	Yes	0.025	0.50	17	-154	-3,836	-1,853
7 Years	No	0.000	0.00	5,206	5,515	-3843	865
7 Years	Yes	0.000	0.00	4,375	4,344	-5,798	-451
7 Years	Yes	0.000	0.50	3,960	3,758	-6,775	1,109
7 Years	Yes	0.025	0.00	3,523	3,490	-5,166	-610
7 Years	Yes	0.025	0.50	3,108	2,905	-6,143	-1,268

Notes: (1) "Benefit Duration" indicates how long the estimated benefits from JTPA are assumed to persist. Actual estimates are used for the first 30 months. For the 7-year duration case, the average of the benefits in months 18-24 and 25-30 is used for the benefits in each future period.

(2) "Welfare Cost of Taxes" indicates the additional cost in terms of lost output due to each additional dollar of taxes raised. The value 0.50 lies in the range suggested by Browning (1987).

(3) Estimates are constructed by breaking up the time after random assignment into 6-month periods. All costs are assumed to be paid in the first 6-month period, while benefits are received in each 6-month period and discounted by the amount indicated for each row of the table.

Source: Heckman and Smith (1998)

TABLE 9

Accounting of Estimated Social Benefits and Costs per Treatment in Selected Social Experiments Evaluating Employment and Training Services for Female Welfare Applicants and Recipients

[1997 U.S. Dollars]

	Program/Main Services Provided					
	NSW	San Diego CWEP/	San Diego CWEP/	San Diego SWIM/	Florida PI/	NJS
	WE	JSA/WE	JSA Only	JSA/CT	JSA	All
Benefits:						
Increased Output from Employment (Includes earnings and fringe benefit impacts)	24,486	3,571	2,457	2,913	757	2,066
From Projected Period Only[1]	19,084	2,101	1,161	0	298	0
Value of In-Program Output	12,039	3,280	-5	262	NA	NA
Reduced Cost of Using Transfer Programs	2,160	131	82	53	130	NA
Reduced Cost of Using Other Programs (e.g., other education and training programs)	1,619	85	74	NA	NA	NA
Costs:						
Program Operating Costs, Including JSA or WE	-13,850	-968	-857	-866	-417	-1,421[3]
Education and Training Costs	0	0	0	-360	-846	NA[3]
Forgone Earnings and Fringe Benefits[2]	-2,341	NA	NA	NA	NA	NA

Participant Out-of-Pocket Expenses (e.g., Transportation, child care, clothing costs)	-431	-24	-	-	-	-
Value of Reduced Non-market Time	-	-	-	-	-	-
Displacement of Other Workers	-	-	-	-	-	-
Deadweight Loss from Taxes to Pay for Programs	-	-	-	-	-	-
Net Present Value of Benefits Minus Costs	21,708	3,123	1,753	2,003	-377	645
Observation Period in Years	2.25	1.00	1.16	5.00	2.00	2.50

Sources: Kemper, et al. (1981) Table IV.1, p. 100, Table IV.2, p. 106, Table IV.6, p. 121; Goldman, et al. (1986), p. 139, Table 5.4, p. 153, Table 5.8, p. 166; Friedlander and Hamilton (1993), Table 5.1, p. 57; Kemple, et al. (1995), Table 7.3, p. 174, Table 7.5, p. 177; Orr, et al. (1994), Exhibit 6.2, p. 162.

Notes:

1. Projected earnings are based on earnings impacts during the last four quarters of the observation period and are discounted at a rate of five percent per year. Subsequent earnings impacts were assumed to depreciate at a rate of 25 percent per year.

2. For all the first column, the social costs associated with forgone earnings are embodied in the estimates of the increased output from employment. In these cases this measure is net of the forgone earnings costs of the program.

3. The -1421 figure for the NJS includes both program operating costs and education and training costs.

4. A "-" denotes costs that were not estimated in the indicated study.

Even If Effect Identified, Does Not Produce Costs (Show Costs)

Experiments

(a) Disrupt Environments

(Heckman, 1992; Hotz, 1992)
Randomization BIAS

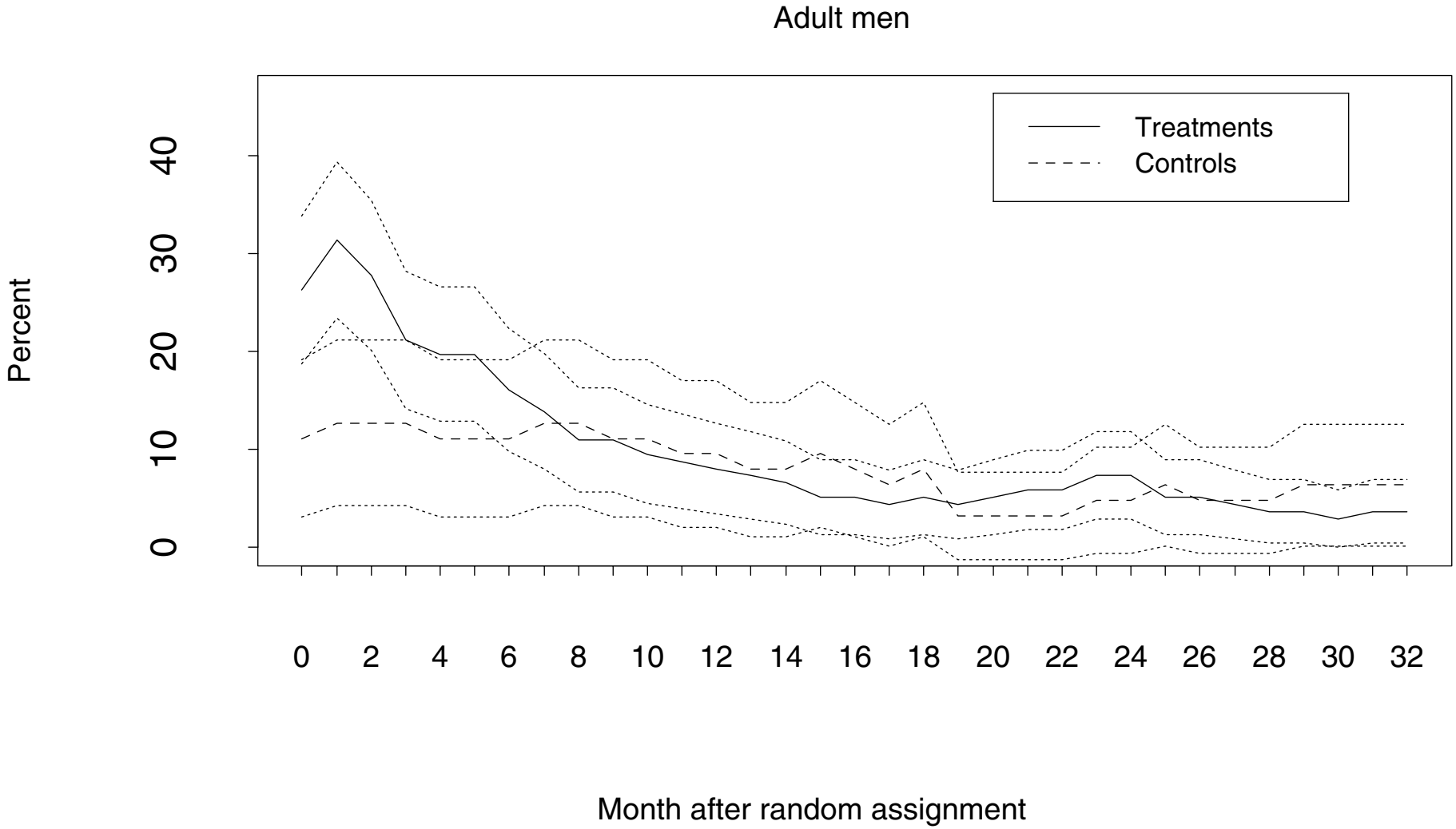
(b) Do not capture entry effects

(Heckman 1992; Moffitt 1992)

(c) Substitution BIAS

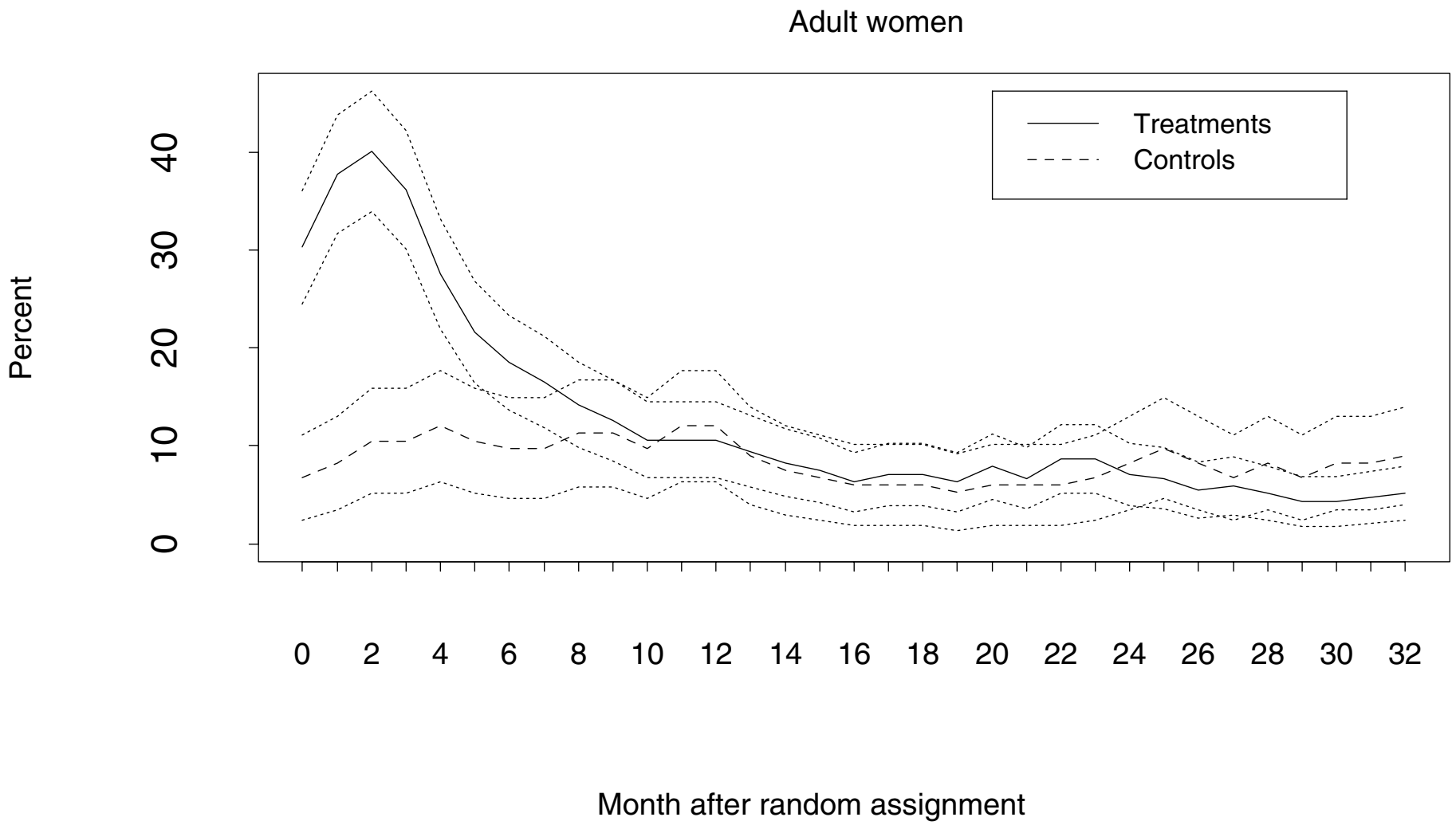
(Heckman, Hohmann and Khoo)

Figure 9
Percentage Receiving Classroom Training



The percentages are the proportion of persons among the sample who report the receipt of classroom training in each month following random assignment. The sample includes only those persons who responded for the entire 32 months of the survey. Month 0 is the month of random assignment. Standard error bars indicate +/- 2 standard errors about the mean.

Figure 10
Percentage Receiving Classroom Training



The percentages are the proportion of persons among the sample who report the receipt of classroom training in each month following random assignment. The sample includes only those persons who responded for the entire 32 months of the survey. Month 0 is the month of random assignment. Standard error bars indicate +/- 2 standard errors about the mean.

TABLE 10
Treatment Group Dropout and Control Group Substitution
in Experimental Evaluations of Active Labor Market Policies
[Fraction of Experimental Treatment and Control Groups Receiving Services]

Study	Authors/Time Period	Target Group(s)	Fraction of Treatments Receiving Services	Fraction of Controls Receiving Services
1. NSW*	Hollister, et al. (1984)	Long Term AFDC Women	0.95 ~	0.11
	(9 months after RA)	Ex-addicts	NA	0.03
		17 - 20 year old H.S. dropouts	NA	0.04
2. SWIM	Friedlander and Hamilton (1993)	AFDC Women: Applicants and Recipients		
	(Time period not reported)	a. Job Search Assistance	0.54	0.01
		b. Work Experience	0.21	0.01
		c. Classroom Training/OJT	0.39	0.21
		d. Any activity	0.69	0.30
		AFDC-U Unemployed Fathers		
		a. Job Search Assistance	0.60	0.01
		b. Work Experience	0.21	0.01
		c. Classroom Training/OJT	0.34	0.22
		d. Any activity	0.70	0.23
3. JOBSTART	Cave, et al. (1993)	Youth High School Dropouts		
	(12 months after RA)	Classroom Training/OJT	0.90	0.26
4. Project Independence	Kemple, et al. (1995)	AFDC Women: Applicants and Recipients		
	(24 months after RA)			

		a. Job Search Assistance	0.43	0.19
		b. Classroom Training/OJT	0.42	0.31
		c. Any activity	0.64	0.40
5. New Chance	Quint, et al. (1994) (18 months after RA)	Teenage Single Mothers		
		Any education services	0.82	0.48
		Any training services	0.26	0.15
		Any education or training	0.87	0.55
6. NJS	Heckman and Smith (1998c) (18 months after RA)	Self-reported from Survey Data		
		Adult Males	0.38	0.24
		Adult females	0.51	0.33
		Male youth	0.50	0.32
		Female youth	0.58	0.41
		Combined Administrative and Survey Data		
		Adult males	0.74	0.25
		Adult females	0.78	0.34
		Male youth	0.81	0.34
		Female youth	0.81	0.42

Notes: RA = random assignment. H.S. = high school. Service receipt includes any employment and training services. The services received by the controls in the NSW study are CETA and WIN jobs. For the Long Term AFDC Women, this measure also includes regular public sector employment during the period.

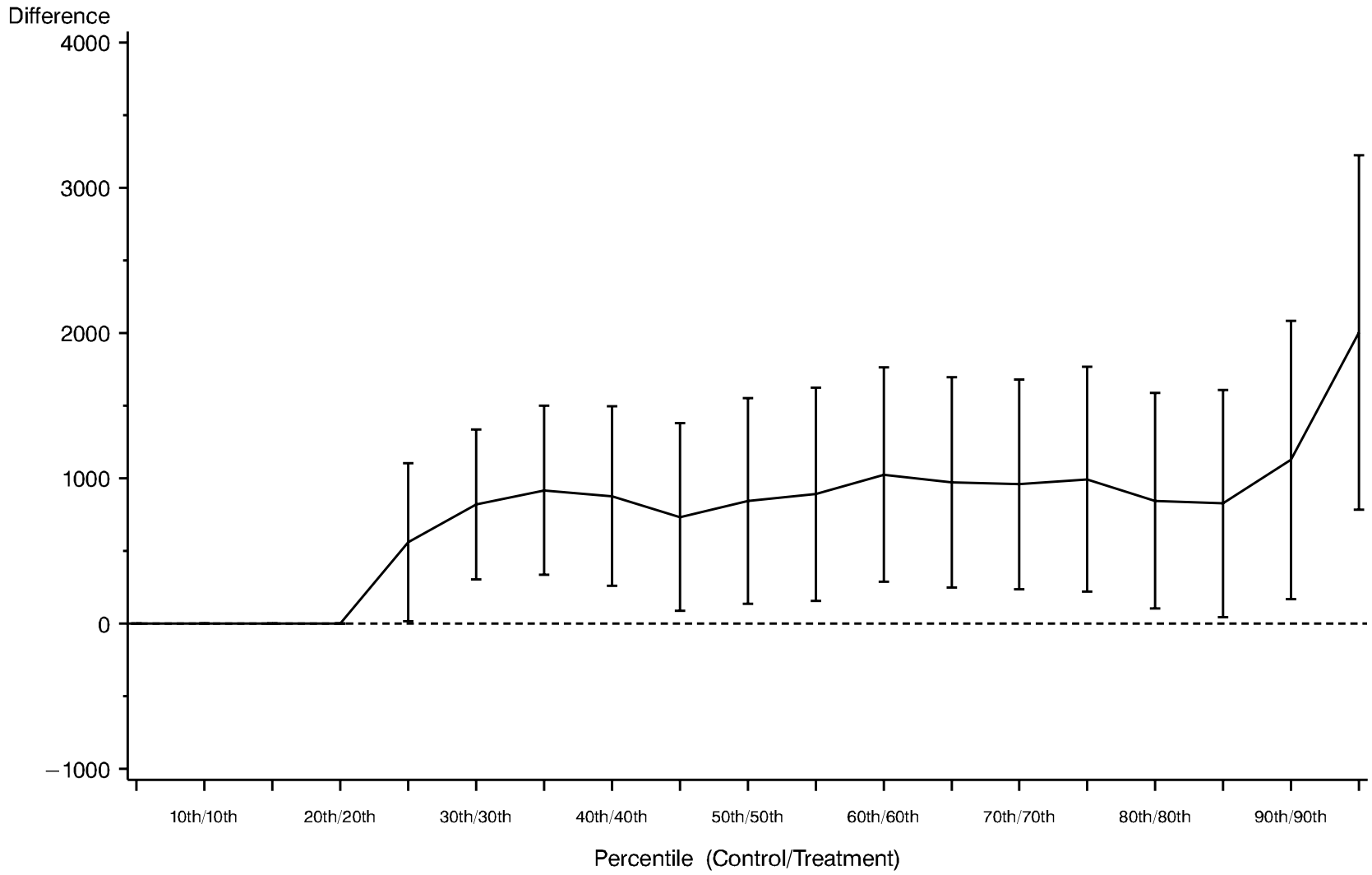
Sources: Masters and Maynard (1981), p. 148, Table A.15; Maynard (1980), p. 169, Table A14. Friedlander and Hamilton (1993), p. 22, Table 3.1; Cave, et al. (1993), p. 95, Table 4-1; Kemple, et al. (1995), p. 58, Table 3.5; Quint, et al. (1994), p. 110, Table 4.9; Heckman and Smith (1998c) and calculations by the authors.

Does Not Produce Distribution of Benefits (Only Determines Marginals—Show Bounds)

Treatment – Control Differences at Percentiles of the 18 Month Earnings Distribution

Perfect Positive Dependence Case

Adult Females



1. Sample consists of ABT's experimental 18-month study samples
2. ABT Imputed values were used in place of outlying values.
3. Standard errors for the quantiles are obtained using methods described in Csorgo (1993).

Source: Heckman and Smith (1993) and Heckman, Smith and Clements (1997)

ESTIMATED PARAMETERS OF THE IMPACT DISTRIBUTION

PERFECT POSITIVE DEPENDENCE, POSITIVE DEPENDENCE WITH $\tau = 0.95$, INDEPENDENCE AND PERFECT NEGATIVE DEPENDENCE CASES
National JTPA Study 18 Month Impact Sample
Adult Females

Statistic	Perfect Positive Dependence ($\tau = 1.0$)	Positive Dependence with $\tau = 0.95$	Independence of Y^1 and Y^0 ($\tau = 0.0$)	Perfect Negative Dependence ($\tau = -1.0$)
5th Percentile	0.00	0.00	-18098.50	-22350.00
	(47.50)	(360.18)	(630.73)	(547.17)
25th Percentile	572.00	125.50	-6043.00	-11755.00
	(232.90)	(124.60)	(300.47)	(411.83)
50th Percentile	864.00	616.00	0.00	580.00
	(269.26)	(280.19)	(163.17)	(389.51)
75th Percentile	966.00	867.00	7388.50	12791.00
	(305.74)	(272.60)	(263.25)	(253.18)
95th Percentile	2003.00	1415.50	19413.25	23351.00
	(543.03)	(391.51)	(423.63)	(341.41)
Percent Positive	100.00	96.00	54.00	52.00
	(1.60)	(3.88)	(1.11)	(0.81)
Impact Std Dev	1857.75	6005.96	12879.21	16432.43
	(480.17)	(776.14)	(259.24)	(265.88)
Outcome Correlation	0.9903	0.7885	-0.0147	-0.6592
	(0.0048)	(0.0402)	(0.0106)	(0.0184)

*τ is the rank correlation coefficient.

Source: Heckman, Smith and Clements (1997)

Other Difficulties With Randomization (Lessons Learned)

Does Not Produce All Treatment Effects (Depends on What is Randomized.)

Does Not Extrapolate

Does Not Account for Full General Equilibrium Effects.

Answers One Set of Questions Well, Subject to Standard Caveats.
(Attrition, Randomization Bias, Substitution Bias)

General Equilibrium

Account For Large Scale Effects of the Programs

Tuition Effects in General Equilibrium are 1/10 those estimated from microdata

Feedback

Tuition $\downarrow \Longrightarrow$

College Enrollment $\uparrow \Longrightarrow$

Return to schooling declines

If agents anticipate or partly anticipate this, then effects overstated.

III. Structural Approaches

1. Consider a Model of Schooling and Training Choice.

2. Use Economic Theory to Identify Joint Distributions (Heckman et al., 2003, *IER*; Cunha, Heckman and Navarro, 2005). (Choice Theory and Indicators of Effects)

3. Explicitly Consider a Variety of Counterfactuals.

4. Respond to Criticism Against Old Structural Models of Being Too Parametric. Adopt a Semiparametric Approach. (Robust Methods)

5. Theory Produces Counterfactuals from a Common Set of Behavioral Parameters.

Fit Good

Accounting For General Equilibrium and Heterogeneity in Evaluating Human Capital Policies

A Dynamic General Equilibrium Model of Human Capital Accumulation with Heterogeneous Agents

1. Treatment Effect and Micro Structural Partial Equilibrium Approaches Ignore Effect of Large Scale Programs on Outcomes. But Many Programs We Wish to Evaluate are Large Scale in Character and Change the Prices and Constraints of Agents Through Tax and Benefit Schedules.

2. Microbased General Equilibrium Models Enrich Microdata By Synchronizing Estimation at Microdata Level With Estimation of Aggregate Levels. Aggregate Models Determine Prices and Interest Rates. Micro Models Determine Quantities.

3. Allows $(Y_0(\omega), Y_1(\omega))$ To Be Affected By the Scale of the Activity, and its Distribution.

4. Estimates Distributions and Feeds Results Into Modern Political Economy Frameworks.

5. Ingredients of Model

6. Results

7. Future: Unite Both

A Dynamic General Equilibrium Model of Human Capital Accumulation with Heterogeneous Agents

Individuals Live for $\bar{a}$ Years and Retire After $a_R \leq \bar{a}$ Years.

- K_{at} : Stock of Physical Capital at Time t.
- H_{at}^{S} : Is the Stock of Human Capital at Time t of Schooling Type S at Age a.

- C_{at} : Consumption
- I_{at}^{S} : Investment in On-the-Job Training for Each Schooling Level, Proportion of Investing.
- θ : Ability

Individuals live for $\bar{a}$ years and retire after $a_R \leq \bar{a}$ years.

K_{at} : Stock of physical capital at time t.

H_{at}^S : Is the stock of human capital at time t of type S at age a.

C_{at} : Consumption

I_{at}^S : For each schooling level.

(III-1) $$V_{at}(H^S_{at}, K_{at}, S) =$$

$$\max_{C_{at}, I^S_{at}} \frac{C^\gamma_{at} - 1}{\gamma} + \delta V_{a+1,t+1}$$

$$(H^S_{a+1,t+1}, K_{a+1,t+1}, S)$$

(III-2) $$K_{a+1,t+1} \leq K_{a,t}(1 + (1 - \tau_k)r_t$$

$$+ R^S_t H^S_{at}(1 - I^S_{at})$$
$$-\tau_1(R^S_t H^S_{at}(1 - I^S_{at})) - C_{at}$$

(III-3) $H^S_{a+1,t+1} = A^S(\theta) I^{\alpha_S}_{at} H^{\beta_S}_{at} + (1-\sigma^S) H^S_{at}$

(III-4) $\hat{S} = \underset{S}{\text{Argmax}}\left[V^S(\theta) - D^S + \varepsilon^S\right]$

$$F(\bar{H}_t^1, \bar{H}_t^2, \bar{K}_t) =$$

$$a_3 \left(\begin{array}{r} a_2 \left(a_1 (\bar{H}_t^1)^{\rho_1} + (1 - a_1)(\bar{H}_t^2)^{\rho_1} \right)^{\rho_2/\rho_1} \\ \\ + (1 - a_2)\bar{K}_t{}^{\rho_2} \end{array} \right)^{1/\rho_2}.$$

HLT estimate that $\rho_2 = 0$ but $\rho_1 = .693$.

Tax Effects on Human Capital Accumulation
Analyzing Two Tax Reforms

Table 7

Closed Economy Effects of Alternative Tax Proposals
General Equilibrium (Steady State) and Partial Equilibrium Effects§
Percentage Difference from Progressive Case†

	Flat Tax‡		Flat Cons. Tax‡	
	PE	GE	PE	GE
After Tax Interest Rate	0.00	1.96	17.65	3.31
Skill Price College HC	0.00	-1.31	0.00	3.38
Skill Price HS HC	0.00	-0.01	0.00	4.65
Stock of Physical Capital	-15.07	-0.79	86.50	19.55
Stock of College HC	22.41	2.82	-15.77	1.85
Stock of HS HC	-9.94	0.90	1.88	0.08
Stock of College HC per College Graduate	3.04	2.55	-4.08	1.72
Stock of HS HC per HS Graduate	1.84	1.07	-5.23	0.16
Aggregate Output	-0.09	1.15	15.76	4.98
Aggregate Consumption	-0.08	0.16	7.60	3.66
Mean Wage College	3.39	2.60	0.12	6.96
Mean Wage HS	2.44	2.44	0.25	6.82
Standard Deviation Log Wage	4.09	1.56	-1.94	0.69
Gini Coefficient	7.61	3.05	-3.14	2.44
College/HS Wage Premium at 10 Yrs Exp*	1.92	-0.45	3.10	0.18
Fraction attending college	18.79	0.26	-12.18	-1.92
Type 1: Fraction Attending College	50.29	-1.25	-42.57	2.14
Type 2: Fraction Attending College	28.50	-5.89	-15.60	-7.88
Type 3: Fraction Attending College	14.13	-6.93	-5.20	-9.56
Type 4: Fraction Attending College	15.27	6.13	-11.77	7.50
Type 1: College HC Gain First 10 Years**	5.81	3.12	-7.53	1.51
Type 2: College HC Gain First 10 Years**	5.33	2.86	-6.84	1.38
Type 3: College HC Gain First 10 Years**	5.60	3.10	-6.70	1.61
Type 4: College HC Gain First 10 Years**	6.85	4.17	-6.41	2.56
Type 1: HS HC Gain First 10 Years**	3.42	1.06	-7.79	-0.34
Type 2: HS HC Gain First 10 Years**	4.49	1.97	-7.60	0.46
Type 3: HS HC Gain First 10 Years**	5.36	2.67	-7.62	1.06
Type 4: HS HC Gain First 10 Years**	5.29	2.55	-7.95	0.92

§ General equilibrium (GE) effects allow skill prices to change, while partial equilibrium (PE) effects hold prices constant.

† In the progressive case we allow for a progressive tax on labor earnings, but assume a flat tax on capital at 15% .

‡ In the flat tax regime we hold the tax on capital fixed to the same level as the progressive tax, but the tax on labor income is flat and is calculated to balance the budget in the new GE steady state. This yields a tax rate on labor income of 7.7% . In the consumption regime, we tax only consumption at a 10.0% rate, again balancing the budget in steady states.

∗ The college - high school wage premium measures the difference in log mean earnings between college graduates and high school graduates with ten years of experience.

∗∗ These rows present changes in the ratio of human capital at ten years of experience versus human capital upon entering the labor force.

Table 8A
Votes for Policy Reform in the Initial State and
Outcomes in Final Steady State

	Movement to Flat Income Tax	Movement to Flat Consumption Tax
% in Favor in Initial State	43%	52%
Final Steady State Utility Gain		
High School Ability 1:	-0.61	0.27
High School Ability 2:	-0.20	0.71
High School Ability 3:	0.09	0.93
High School Ability 4:	-0.13	0.78
College Ability 1:	-0.53	0.35
College Ability 2:	-0.32	0.58
College Ability 3:	-0.18	0.72
College Ability 4:	0.23	1.16
Ability 1	-0.57	0.30
Ability 2	-0.28	0.64
Ability 3	-0.03	0.85
Ability 4	0.11	1.05

Tax Effects on Human Capital Accumulation

General Equilibrium Treatment Effects: A Study of Tuition Policy

Conventional Models of Treatment Effects

The Treatment Effect for Person i is

$$\Delta_i = Y_i^1 - Y_i^0.$$

General Equilibrium Effect of Tuition on Enrollment: 10% of Microeconomic Effect.

Analyzing Two Tax Reforms

Closed Economy Effects of Alternative Tax Proposals
General Equilibrium [GE] (Steady State) and
Partial Equilibrium [PE] Effects[§]
Percentage Difference from Progressive Case[†]

	Flat Tax[‡]		Flat Tax[‡]	Cons.
	PE	GE	PE	GE
After Tax Interest Rate	0.00	1.96	17.65	3.31
Skill Price College HC	0.00	-1.31	0.00	3.38
Skill Price HS HC	0.00	-0.01	0.00	4.65
Stock of Physical Capital	-15.07	-0.79	86.50	19.55
Stock of College HC	22.41	2.82	-15.77	1.85
Stock of HS HC	-9.94	0.90	1.88	0.08
Stock of College HC per College Graduate	3.04	2.55	-4.08	1.72
Stock of HS HC per HS Graduate	1.84	1.07	-5.23	0.16
Aggregate Output	-0.09	1.15	15.76	4.98
Aggregate Consumption	-0.08	0.16	7.60	3.66
Mean Wage College	3.39	2.60	0.12	6.96
Mean Wage HS	2.44	2.44	0.25	6.82
Standard Deviation Log Wage	4.09	1.56	-1.94	0.69
Gini Coefficient	7.61	3.05	-3.14	2.44
College/HS Wage Premium at 10 Yrs Exp*	1.92	-0.45	3.10	0.18
Fraction attending college	18.79	0.26	-12.18	-1.92
Type 1[§§]: Fraction Attending College	50.29	-1.25	-42.57	2.14
Type 2[§§]: Fraction Attending College	28.50	-5.89	-15.60	-7.88
Type 3[§§]: Fraction Attending College	14.13	-6.93	-5.20	-9.56
Type 4[§§]: Fraction Attending College	15.27	6.13	-11.77	7.50
Type 1[§§]: College HC Gain First 10 Years**	5.81	3.12	-7.53	1.51
Type 2[§§]: College HC Gain First 10 Years**	5.33	2.86	-6.84	1.38
Type 3[§§]: College HC Gain First 10 Years**	5.60	3.10	-6.70	1.61
Type 4[§§]: College HC Gain First 10 Years**	6.85	4.17	-6.41	2.56
Type 1[§§]: HS HC Gain First 10 Years**	3.42	1.06	-7.79	-0.34
Type 2[§§]: HS HC Gain First 10 Years**	4.49	1.97	-7.60	0.46
Type 3[§§]: HS HC Gain First 10 Years**	5.36	2.67	-7.62	1.06
Type 4[§§]: HS HC Gain First 10 Years**	5.29	2.55	-7.95	0.92

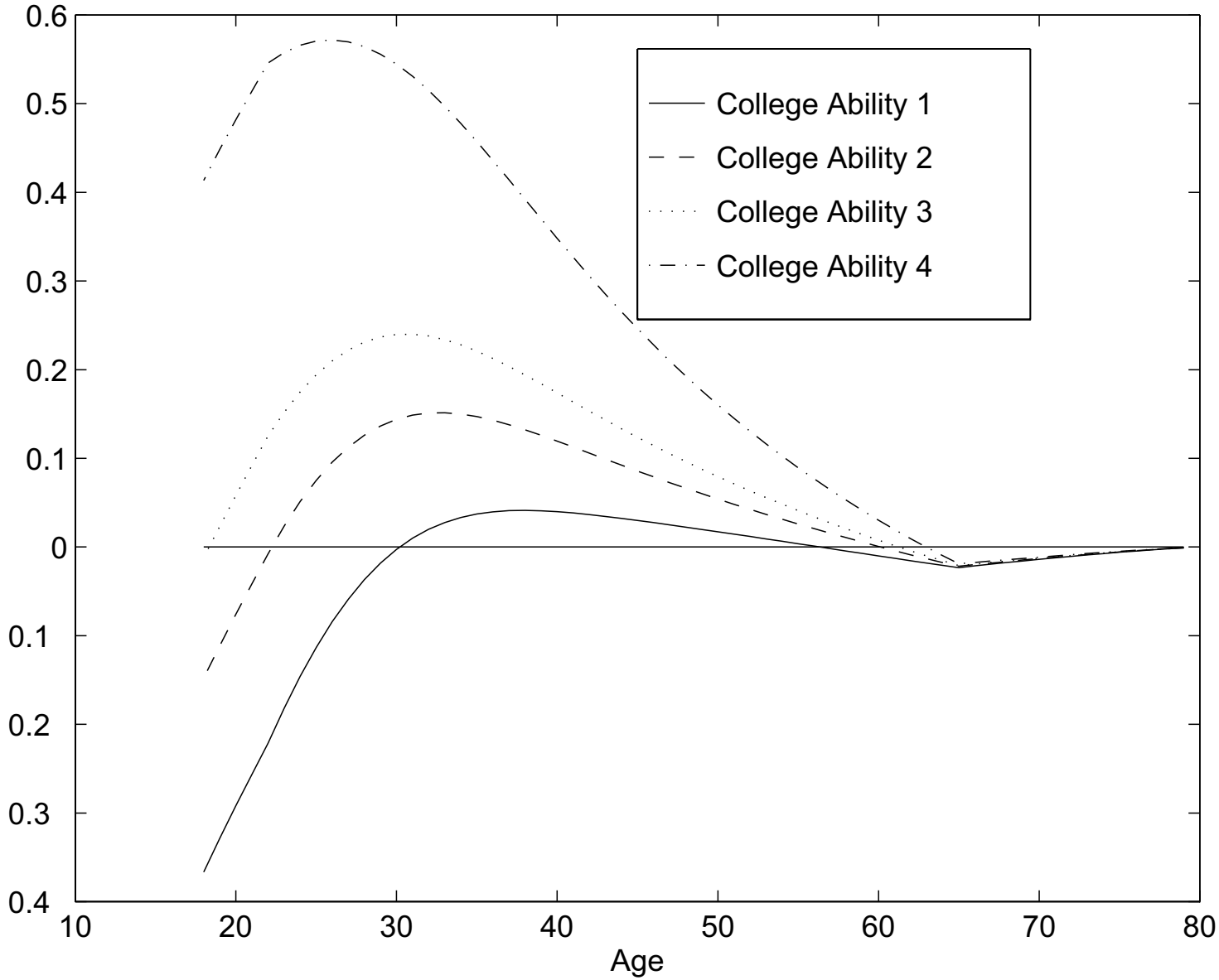
Votes for Policy Reform in the Initial State and Outcomes in Final Steady State

	Movement to Flat Income Tax	Movement to Flat Consumption Tax
% in Favor in Initial State	43%	52%
	Final Steady State Utility Gain	
High School Ability 1:	-0.61	0.27
High School Ability 2:	-0.20	0.71
High School Ability 3:	0.09	0.93
High School Ability 4:	-0.13	0.78
College Ability 1:	-0.53	0.35
College Ability 2:	-0.32	0.58
College Ability 3:	-0.18	0.72
College Ability 4:	0.23	1.16
Ability 1	-0.57	0.30
Ability 2	-0.28	0.64
Ability 3	-0.03	0.85
Ability 4	0.11	1.05

Ability 1 is the lowest ability and Ability 4 is the highest as measured by the AFQT distribution quartiles.

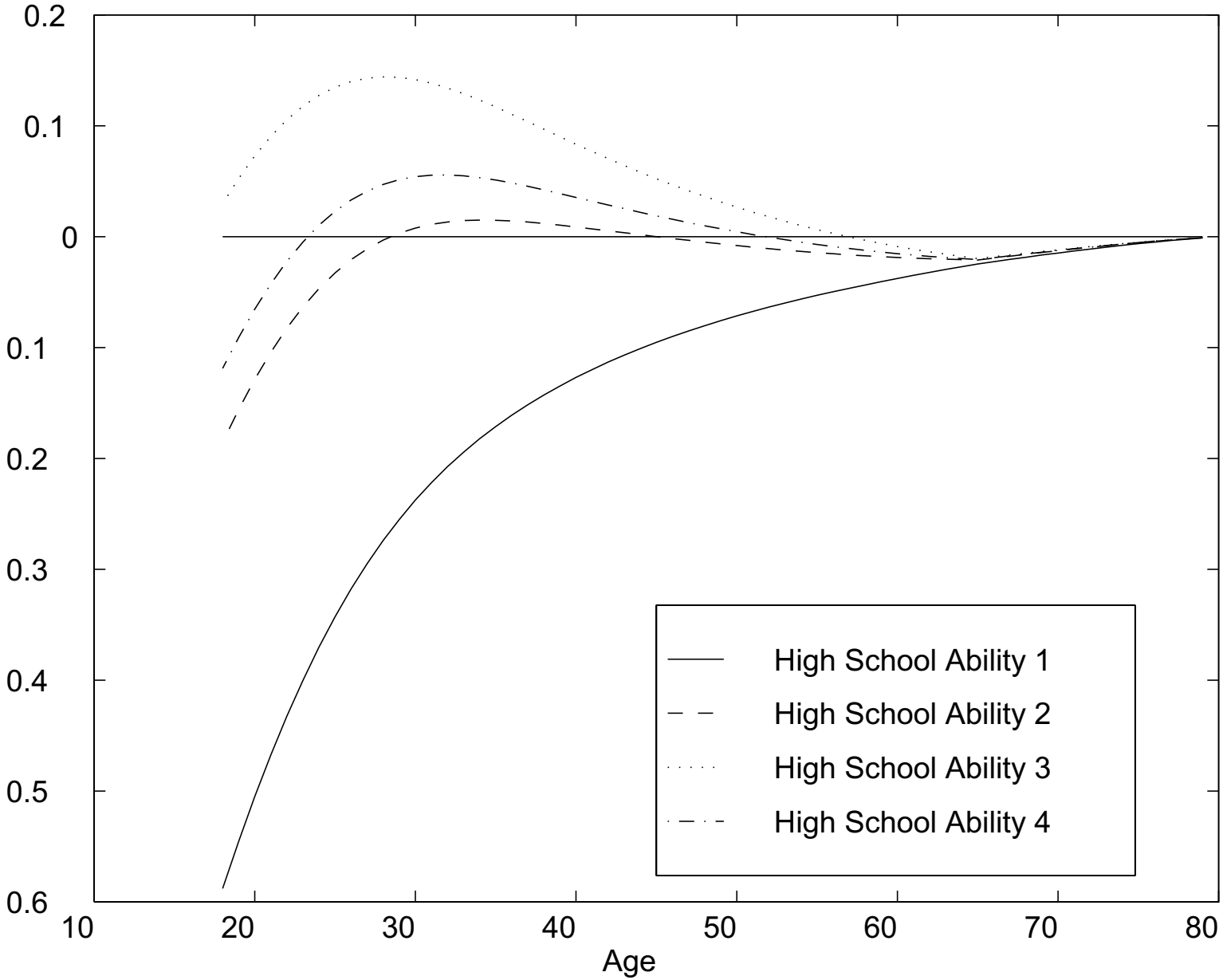
Simulations based on the closed economy estimates of Heckman, Lochner and Taber (1999).

Changes in Utility from the Reform in the Current Generation: Flat Tax



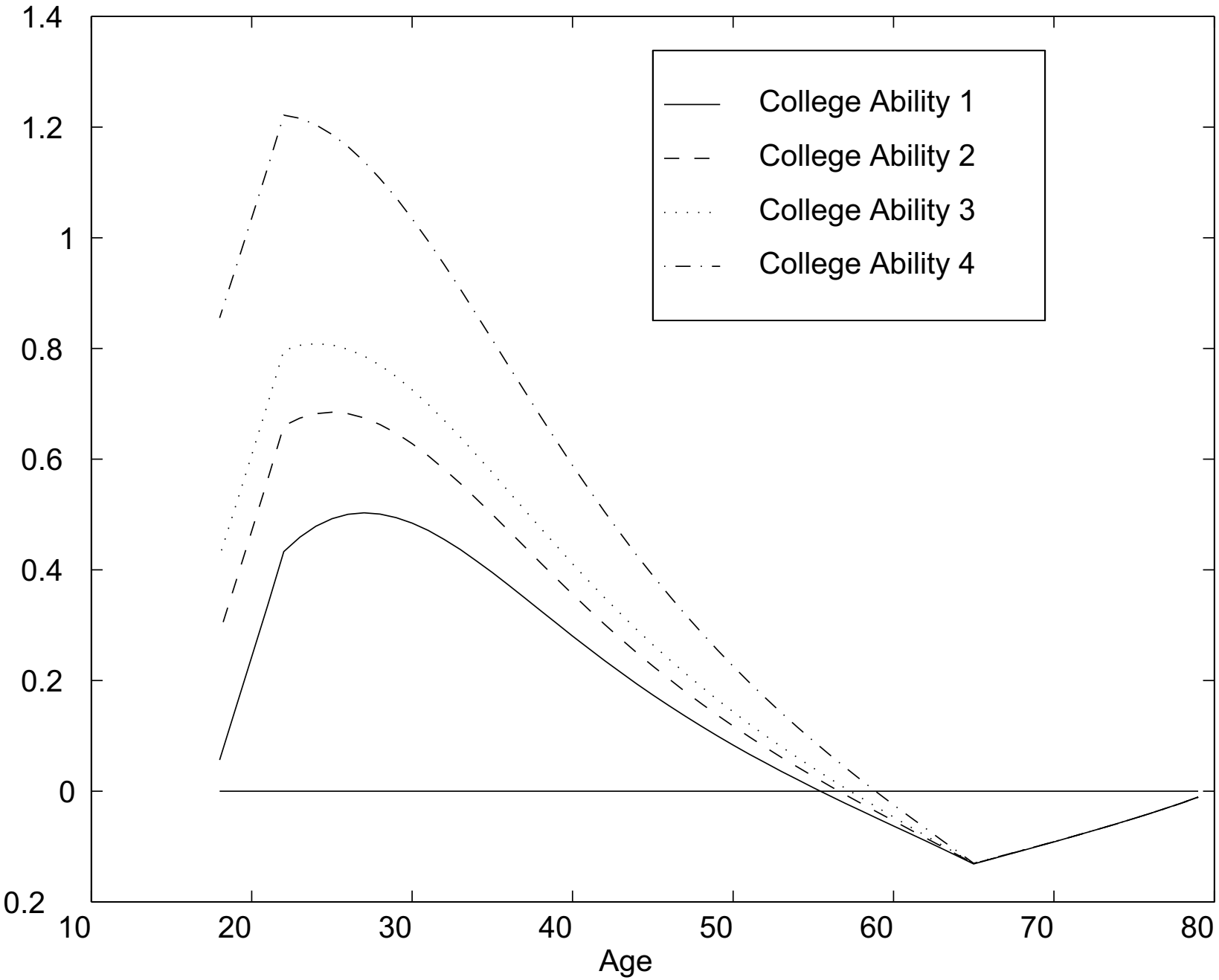
Source: Simulations based on Heckman, Lochner and Taber (1998, 1999)

Changes in Utility from the Reform in the Current Generation: Flat Tax



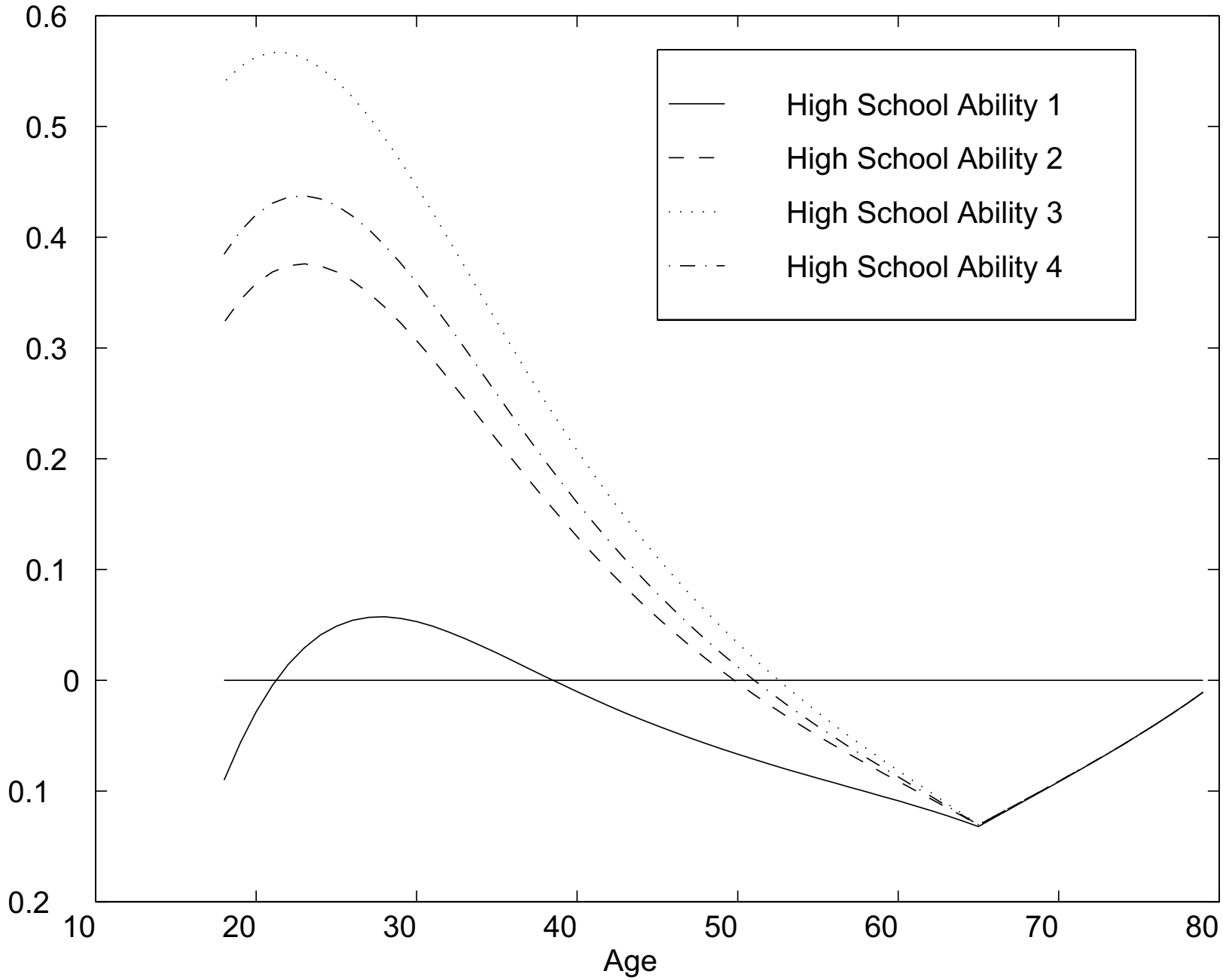
Source: Simulations based on Heckman, Lochner and Taber (1998, 1999)

Changes in Utility from the Reform in the Current Generation: Consumption Tax



Source: Simulations based on Heckman, Lochner and Taber (1998, 1999)

Changes in Utility from the Reform in the Current Generation: Consumption Tax



Source: Simulations based on Heckman, Lochner and Taber (1998, 1999)

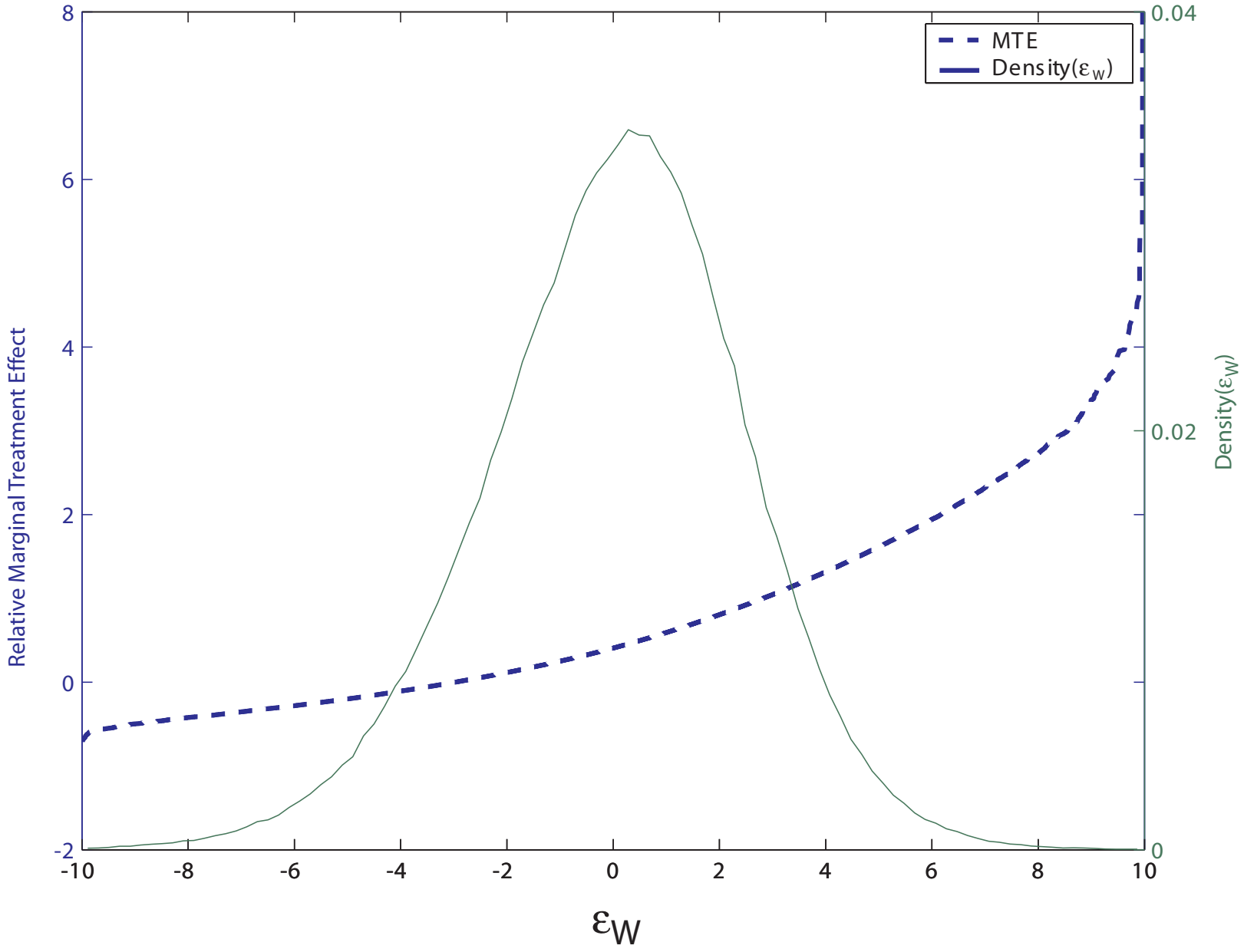
Summary

1. Evaluation using microdata improves government efficiency and raises level of discussion of policy
2. Quality of evaluations crucially depend on the quality of the data
3. Randomization is no panacea, serious problems with using it
4. Important to measure costs and benefit to fully assess social policy

5. General equilibrium effects of large scale programs are important
   Gives Us Marginal Effects (Figure 7)

6. ∴ We Can Input Into the Cost-Benefit the Marginal and Average Effects of a Policy.

Figure 7: Density of ε_W and Relative Marginal Ex Post Treatment Effect
For Present Value of Gross Earnings E((PVc/PVh)-1| ε_W)



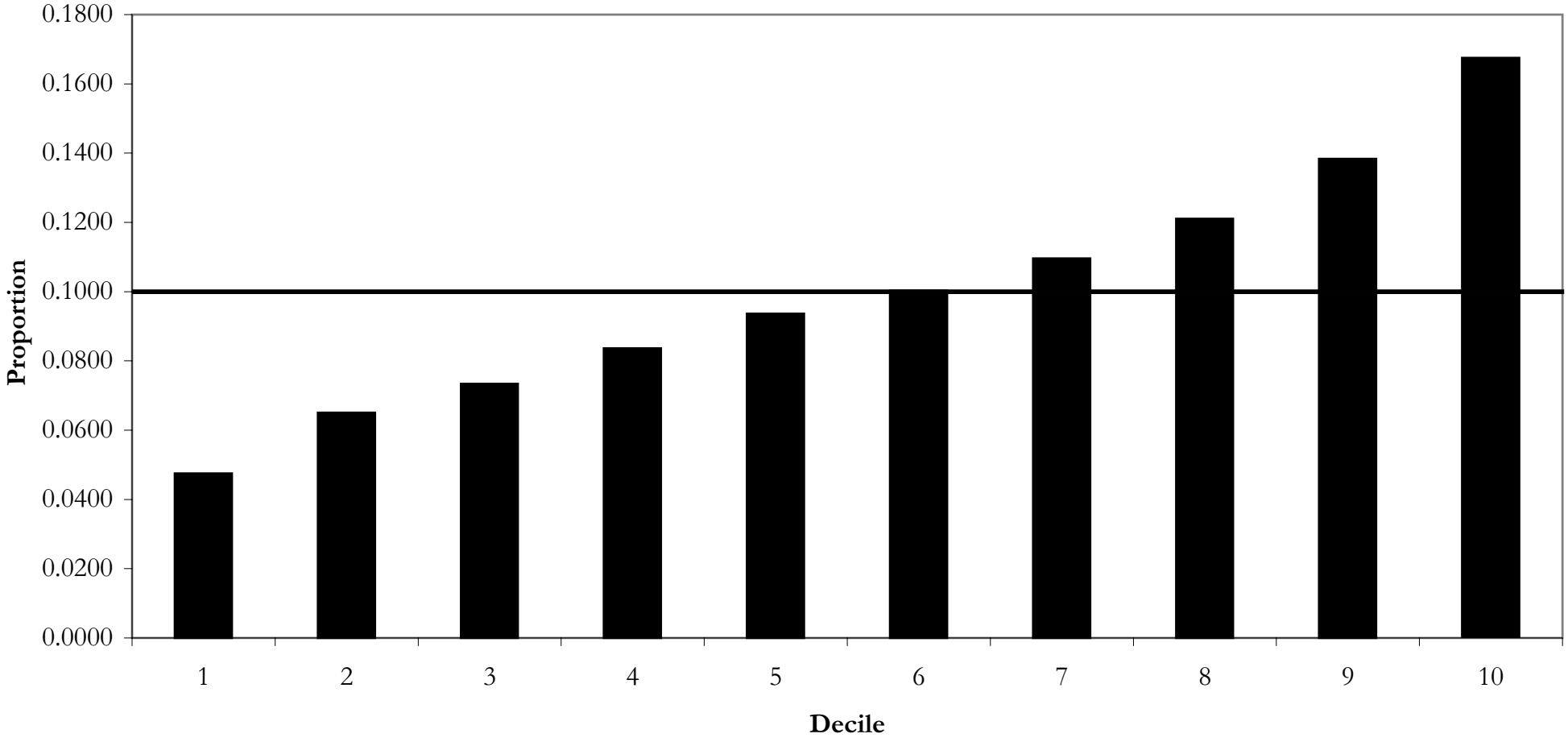
$$*\varepsilon_W = (\alpha'_{1,c} + \alpha'_{2,c} - \alpha'_{1,h} - \alpha'_{2,h} - \alpha'_p)\theta - \varepsilon_p$$

All densities are estimated using a 100 point grid over the domain and a Gaussian kernel with bandwidth of 0.12.

Can Explore Distributional Consequences of Subsidy to Tuition (Figure 8)

Figure 8

Proportion of People Induced Into College by Full Subsidy to College Tuition When Information Set = $\{\theta_1, \theta_2\}$ by Decile of Initial High School Earnings Distribution



One Can Evaluate a Full Range of Policies.

Still Partial Equilibrium in Character.