NBER WORKING PAPER SERIES

REMOVING THE VEIL OF IGNORANCE IN ASSESSING
THE DISTRIBUTIONAL IMPACTS OF SOCIAL POLICIES

Pedro Carneiro
Karsten T. Hansen
James J. Heckman

Removing the Veil of Ignorance in Assessing the Distributional Impacts of Social Policies
Pedro Carneiro, Karsten T. Hansen and James J. Heckman
NBER Working Paper No. 8840
March 2002
JEL No. I28, D33, H43

ABSTRACT

This paper summarizes our recent research on evaluating the distributional consequences of social programs. This research advances the economic policy evaluation literature beyond estimating assorted mean impacts to estimate distributions of outcomes generated by different policies and determine how those policies shift persons across the distributions of potential outcomes produced by them. Our approach enables analysts to evaluate the distributional effects of social programs without invoking the "Veil of Ignorance" assumption often used in the literature in applied welfare economics. Our methods determine which persons are affected by a given policy, where they come from in the ex-ante outcome distribution and what their gains are. We apply our methods to analyze two proposed policy reforms in American education. These reforms benefit the middle class and not the poor.

Pedro Carneiro
Department of Economics
University of Chicago
1126 E. 59th Street
Chicago, IL 60637

Karsten T. Hansen
Department of Economics
University of Chicago
1126 E. 59th Street
Chicago, IL 60637

James J. Heckman
Department of Economics
University of Chicago
1126 E. 59th Street
Chicago, IL 60637
and NBER
jjh@uchicago.edu

1　Introduction

Evaluating public policy is a central task of economics. Welfare economics presents different criteria. Research on program evaluation develops and applies a variety of different econometric estimators. Traditional empirical methods focus on mean impacts. Yet modern welfare economics emphasizes the importance of accounting for the impact of public policy on *distributions* of outcomes (Sen, 1997, 2000). A large body of empirical evidence indicates that people differ in their responses to the same policy and act on those differences, and that the representative agent paradigm is a poor approximation to reality because the marginal entrant into a social program is often different from the average participant. (Heckman, 2001a). This evidence highlights the importance of going beyond the representative agent framework when evaluating public policies.

This paper summarizes our recent research on evaluating the distributional consequences of public policy.[1] Our research advances the economic policy evaluation literature beyond estimating assorted mean impacts to estimate the distributions of outcomes generated by different policies and to determine how those policies shift persons across the distributions of potential outcomes produced by them. We distinguish the average participant in a program from the marginal entrant.

Our research advances the existing literature on evaluating the distributional consequences of alternative policies beyond the "Veil of Ignorance" assumption used in modern welfare economics (See Atkinson 1970, Sen 1997, 2000). Approaches based on that assumption compare two social states by assuming that the position of any particular individual in one distribution should be treated as irrelevant. In this approach the overall distribution

---

[1]Carneiro, Hansen and Heckman (2000, revised 2001).

of outcomes is all that matters. This is a consequence of the *anonymity postulate* that is fundamental to that literature. Anonymity is the property that only the distribution of outcomes matters and that reversing the positions of any two persons in the overall distribution does not affect the evaluation placed on the policy (or state of affairs) that produces the distribution.

There are normative arguments that support this criterion. (See Harsanyi, 1955, Vickery, 1960, and Roemer, 1996). As a positive description of actual social choice processes, the "Veil of Ignorance" seems implausible. Participants in the political process are likely to forecast their outcomes under alternative economic policies, and assess policies in this light. (Heckman, 2001b). This paper extends current practice by developing and applying methods that forecast how people fare under different policies. We link the literature in modern welfare economics to the treatment effect literature.

This paper proceeds as follows. We briefly present the evaluation problem for an economy with two sectors (e.g. schooled and unschooled) where agents select or are selected into "treatment" (one of the two sectors). We consider policies that affect choices of treatment (e.g. schooling) but not potential outcomes (the outcomes they experience under different treatments). We compare outcomes across two policy regimes that affect treatment choices. This task is much easier when individuals respond in the same way to treatment than when they differ in their response to treatment, and act on those differences in making treatment choice decisions. In the latter case, the marginal entrant into schooling is not the same as the average participant in treatment and the representative agent paradigm breaks down. In an appendix, we show how to generate the counterfactual distributions of outcomes produced by alternative policies.

We apply our analysis to estimate the distributional consequences of two proposed policy

reforms in American education. Even though the two policies barely affect the overall distribution of outcomes, and so would be judged to be equivalent to the pre-policy origin state under the Veil of Ignorance criterion, they have substantial effects on a small group of people concentrated in the middle to the high end of the pre-policy wage distribution. Marginal entrants attracted into college get smaller gains than average college students suggesting diminishing returns to programs that encourage college enrollment. Marginal entrants into junior college are about the same as average entrants, suggesting constant returns for that schooling level. Since most of the people affected by the policies come from the middle to the high end of the original wage distribution, there is little impact of these policies on the poor.

2 The Evaluation Problem for Means and Distributions

In order to place our work in the context of the current literature on social program evaluation, and to link it to the economics of education, it is helpful to consider a simple generalized Roy (1951) economy with two sectors. Let $S = 1$ denote college and $S = 0$ be high school. Persons (or their agents, such as their parents) can choose to be in either sector. There are two potential outcomes for each person (Y_0, Y_1), only one of which is observed, since it is assumed that only one option can be pursued at any time. For simplicity,

we assume that the decision rule governing sectoral choices is

$$S = \begin{cases} 1 & \text{if } I = Y_1 - Y_0 - C \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Here C is the cost of choosing $S = 1$. In the context of a schooling model, C is tuition or monetized psychic cost, while $Y_1 - Y_0$ is the net gain from schooling expressed, say, in present value terms.

We decompose Y_1 and Y_0 in terms of their means μ_1 and μ_0 and mean zero idiosyncratic deviations (U_1, U_0) or residuals:

$$\begin{aligned} Y_1 &= \mu_1 + U_1 \\ Y_0 &= \mu_0 + U_0. \end{aligned}$$

We condition on X variables, but for notational simplicity we keep this dependence implicit. Decomposing C in a similar fashion, we may write:

$$C = \mu_C + U_C,$$

so that

$$I = \mu_1 + \mu_0 - \mu_C + (U_1 - U_0 - U_C).$$

It is fruitful to distinguish two kinds of policies: (a) those that affect potential outcomes (Y_0, Y_1) through price and quality effects and (b) those that affect sectoral choices (through C) but do not affect potential outcomes. Tuition and access policies that do

4

not have general equilibrium effects fall into the second category of policy. Policies with general equilibrium effects and policies that directly affect rewards to potential outcomes and quality are examples of the first kind of policy. It is the second kind of policy that receives the most attention in empirical work on estimating economic returns to schooling (see e.g., the survey by Card (1999)) or in evaluating schooling policies (see e.g., Kane (1994)).

Consider two policy environments denoted A and B. These produce two social states for outcomes that we wish to compare. In the general case, we may distinguish an economy operating under policy A with associated cost and outcome vector (Y_0^A, Y_1^A, C^A) for each person, from an economy operating under policy B with associated cost and outcome vector (Y_0^B, Y_1^B, C^B). Policy interventions with no effect on potential outcomes can be described as producing two choice sets (Y_0, Y_1, C^A) and (Y_0, Y_1, C^B) for each person. In this paper we focus on evaluating the second kind of policy that keeps invariant the distribution of potential outcomes across policy states, but affects the cost of choosing sector 1 within each state.

Our framework differs in its emphasis from the standard model of modern welfare economics. Analysts writing in that tradition focus on the distribution of outcomes produced by each policy without inquiring how those outcomes are produced. All policies that produce the same aggregate outcome distributions are judged to be equally good. The details of how the observed distribution is produced are deemed irrelevant. The distinctions we make between policies that affect potential outcomes and policies that affect which potential outcomes are selected are also ignored in that literature. There is no explicit discussion of sectoral choice within policy states. The literature starts, and stops, with an analysis of distributions of the observed outcomes for each person in each policy state (Y^A, Y^B)

defined as

$$Y^A = S_A Y_1^A + (1 - S_A)Y_0^A, \quad \text{and} \quad Y^B = S_B Y_1^B + (1 - S_B)Y_0^B,$$

where S_A and S_B are schooling choice indicators under policies A and B respectively, without inquiring more deeply into the sources of the differences in the distributions of outcomes.

The modern treatment effect literature focuses on these details and distinguishes choice of treatments from the treatment outcomes. However, it only inquires about certain mean treatment effects. The operating assumption in the literature is that policies do not affect potential outcomes (so $(Y_0^A, Y_1^A) = (Y_0^B, Y_1^B)$), but do affect choices of sectors.

This literature distinguishes three cases. Case I arises when everyone (with the same X) gets the same effect from treatment ($Y_1 - Y_0$ is the same for everyone). Case II occurs when $Y_1 - Y_0$ differs among people of the same X but decisions to enroll in the program are not affected by these differences:

$$\Pr(S = 1|Y_1 - Y_0) = \Pr(S = 1)$$

Case III occurs when $Y_1 - Y_0$ differs among people and people act on these differences. In cases I and II, the marginal entrant into a program is the same as the average entrant. In case III, this is not so. People select in part on gains. If they select solely on gains, then the marginal entrant gets a lower return than those participants (in 1) who are inframarginal; that is, the marginal treatment effect (MTE)

$$E(Y_1 - Y_0|I = 0) < E(Y_1 - Y_0|S = 1)$$

6

See Heckman 2001a for more discussion of the various treatment effects.[2]

3 Comparing Two Policy States

Consider two policies, A and B, that affect sectoral choices without affecting the distributions of potential outcomes. For concreteness, we can think of these as policies that affect C (e.g., tuition or access) by shifting its mean, changing its variance or changing the covariance between C and (Y_0, Y_1). Each policy produces a distribution of outcomes. For concreteness, think of the outcome as wages associated with different schooling levels.

In the literature on evaluating inequality, comparisons of policies are made in terms of comparisons of distributions. If policy B produces an aggregate distribution of wages that stochastically dominates that produced from policy A, B is preferred.[3] The details of who benefits or loses from the policy are considered to be irrelevant as a consequence of the anonymity postulate.

The literature on evaluating inequality in modern welfare economics compares two aggregate outcome distributions. If policy A has been implemented, but policy B has not, evaluation of B entails construction of a counterfactual aggregate outcome distribution. Under the assumptions used in the treatment effect literature, all that is required is determination of how policy B sorts persons into sectors "0" and "1", and how such sorting affects *observed* outcome distributions in sectors "0" and "1". In our example, what is required is a schooling choice equation and a selection model to identify the invariant potential outcome distributions. The selection model enables analysts to go from observed

---

[2]Bjorklund and Moffitt (1987) introduced the marginal treatment effect into the evaluation literature. See Heckman (2001a) for a summary of extensions of this literature.

[3]See Sen (1997).

(selected) distributions of Y_0 and Y_1 to the population potential distributions. With sufficient individual variation in C within an economy governed by policy A, it is possible to accurately forecast the effect of policy B on the overall distribution without previously observing it, as we demonstrate in this paper.

Our approach to the evaluation of public policy is more ambitious in some respects than the recent literature in welfare economics and is more in line with the objectives of modern political economy. (Persson and Tabellini, 2000). We relax the anonymity postulate and determine how individuals at different positions within the initial overall distribution respond to policies in terms of their treatment choices and gains. We estimate the number of people directly affected by the policy, where they start, and where they end up in the overall distribution.

In the context of the treatment effect framework, this task is broken down into two sub-tasks. The first sub-task is to determine who shifts treatment state in response to the policy and where they are located in the initial overall distribution. The second sub-task is to determine where they end up in the overall distribution after taking the treatment, and how much they gain. Since this approach assumes that potential outcome distributions are not affected by the policies, it is less ambitious, in this respect, than the approach advocated in modern welfare economics which entertains that possibility.

Under case I, this task is greatly simplified. Everyone who shifts from "0" to "1" gets the same gain Δ. The only problem is to find where in the initial overall distribution the switchers are located. Under case II, Δ varies among observationally identical people. The gain is not necessarily the same for persons with different initial Y_0 values. However, *on average*, across all movers, the gain is the same as the mean difference between the two potential outcome distributions *within* policy regime A. Hence the marginal entrant has

8

the same mean as the average person and the average participant:

$$E(Y_1 - Y_0 | I = 0) = E(Y_1 - Y_0) = E(Y_1 - Y_0 | S = 1).$$

Case III differs from case II in that in general the gains to the average switcher are not the same as the gains to the previous participants. If $(Y_1 - Y_0)$ is positively correlated with $I = (Y_1 - Y_0 - C)$, the marginal entrant receives lower gains on average than does the average participant. The details of constructing the transition densities for the switchers are presented in our companion paper.

4 Identifying Counterfactual Distributions Under Treatment Effect Assumptions

Identifying the joint distribution of potential outcomes under treatment effect assumptions is more difficult than identifying the various mean treatment effects.[4] The fundamental problem is that we never observe both components of (Y_0, Y_1) for anyone.[5] Thus we cannot directly form the joint distribution of potential outcomes (Y_0, Y_1).

In the Appendix, we review various approaches to estimating, or bounding, counterfactual distributions that have appeared in the literature. In our source paper, we develop a new method for identifying these distributions. It is based on an idea common in factor

---

[4]A large econometric literature identifies the mean impacts under a variety of assumptions. See Heckman, LaLonde and Smith (1999) for one survey. Heckman and Vytlacil (2000, 2001) consider identification of marginal treatment effects and unify the treatment effect literature.

[5]Panel data estimators sometimes enable analysts to observe both components. See Heckman, LaLonde and Smith (1999).

analysis but applied to model counterfactual distributions. If potential outcomes are generated by a low dimensional set of factors, then it is possible to estimate the distributions of factors and generate distributions of the counterfactuals. Here, low dimensional refers to the number of factors relative to the number of measured outcomes. See the Appendix for the intuitive idea that motivates the analysis in our source paper. We next turn to an application of our analysis to American data.

5 Some Evidence From America on Two Educational Reforms

Our companion paper uses data on wages, schooling choices and covariates for white males from the National Longitudinal Study of Youth (NLSY) to estimate a three factor version of the model described in the Appendix using a Bayesian semiparametric mixture of normals econometric framework. We consider four schooling levels: dropout, high school graduate, junior college and four year college. We use local labor market variables, tuition and family background information to identify the model. The estimated model fits the data well. Observed wage distributions are closely approximated. There is no need for more than three factors to fit our data which includes panel data measurements on wages as well as indicators of ability and motivation.[6]

Our paper estimates models for a variety of schooling groups. Here, for the sake of brevity, we focus only on certain key empirical results. We report the wages returns to college and high school, and selection on levels and gains into those schooling categories.

---

[6]The factor model is strongly overidentified so that it would have been possible to estimate many more than three factors.

We analyze two policies: (a) a full tuition subsidy for junior colleges and (b) a policy promoting access to four year colleges which places an institution in the immediate vicinity (the county of residence) of each American. We consider only partial equilibrium treatment effects and do not consider the full cost of financing the reforms.

Our evidence shows considerable dispersion in terms of levels and returns (gains) to various schooling categories. Indeed, *ex post* returns are negative for a substantial fraction of people. There is little evidence of selection either on levels or gains for high school graduates. There is a lot of evidence of selection on levels and gains for college graduates. The marginal entrants into four year colleges induced by the access policy we consider have wage outcomes below the average college participant both in terms of levels and gains. This is not true for the junior college tuition subsidy policy we also analyze. For that case, there is little impact on overall quality of junior college graduates.

Figure 1 shows the potential high school wages for all four schooling groups– what people who actually attend various schooling levels would have earned had they gone to high school. The four densities are nearly the same suggesting that there is little evidence of selection on levels into high school. Three of these four densities are counterfactual. The density for high school graduates is factual. For college (Figure 2), there is strong evidence of selection on levels. Persons who attend college do better in college than dropouts, high school graduates or junior college graduates would do. This result contrasts sharply with the corresponding result for the factual and counterfactual wage densities for high school graduates.

There is also little evidence of selection on gains $(Y_1 - Y_0)$ to high school (high school vs. dropout). See Figure 3 which plots the counterfactual returns to high school for all four schooling groups. The returns (high school vs. four year college) are greater for persons

11

who become college graduates than for the other schooling groups, although there is a lot of overlap in the distributions. See Figure 4. *Ex post* many persons who actually stop their schooling at the high school level would make fine college graduates. Many college graduates experience negative returns. The marginal treatment effect comparing high school to college (Figure 6) suggests that as the unobservables that lead to a higher likelihood of attending college increase, (so $P(S=1)$ increases) the return to college increases. People most likely to attend college have the highest marginal returns. The corresponding figure for the return to high school is flat, suggesting that the marginal participant has the same return as the average participant.

Using the estimated model, we compare two policies: a full subsidy to community college tuition and a policy that places a four year college in each county in America. Table 1 shows the average log wages of participants before the policy change and their average return. It compares these levels and returns with what the marginal participant attracted into the indicated schooling by the policy would earn. Marginal and average log wages and returns are about the same for the community college policy. There is little decline in quality among the entrants. For the access policy, there is a sharp difference. Average participants in four year colleges earn more and have higher returns than marginal entrants. There is a sharp decline in the average quality of college graduates.

Despite the substantial sizes of the policy changes we consider, the induced effects on participation are small. The four year access policy only raises four year graduation rates by 1.3 percent. The junior college subsidy raises attendance at those institutions by 3.8 percent.

The policies operate unevenly over the deciles of the initial outcome distribution. Mobility is greatest at the center of the distribution for the community college policy. See

12

Table 2 and Figure 6. Mobility is from the top of the initial wage distribution for the four year college policy. See Table 3 and Figure 7. Neither policy benefits the poor.

Our approach to the evaluation of social policy is much richer, and more informative, than an analysis of aggregate outcomes of the sort contemplated in modern welfare theory. The overall Gini coefficient does not change (to two decimal points) when we implement the two policies. By the standards of that literature, the pre- and post-policy distributions are the same. A focus on the aggregate outcome distribution masks important details which our approach reveals. Only a small group of persons are directly affected by the policy. The vast majority of persons would be unaffected by these policies, and presumably, would be indifferent to the policy.[7] Our approach to policy evaluation lifts the Veil of Ignorance and provides a more complete interpretation of who benefits from the policy and where beneficiaries come from in the overall distribution of outcomes.

6 Summary and Conclusions

This paper summarizes our recent research on evaluating the distributional consequences of social programs. We move beyond the mean treatment effects that dominate discussion in the recent applied evaluation literature to analyze the impacts of policy on distributions of outcomes. We develop and apply methods for determining which persons are affected by the policy, where they come from in the initial distribution, and what their gains are.

We contrast the outcomes of participants in schooling before the policy change with the outcomes of marginal entrants induced into the treatment state by the policy. We compare

---

[7]Counting their tax burden, they might even be hostile to these policies.

our approach to the approach advocated in modern welfare economics. That approach focuses attention solely on the aggregate distribution and does not identify gainers and losers from a policy. Our approach identifies where gainers and losers are located in the overall distribution. The output produced from our approach generates the information required in positive political economy.

Our analysis has been conducted for a partial equilibrium treatment effect model that assumes that policies do not affect the distribution of potential outcomes, just the choice probabilities of particular treatments. It would be desirable to extend our framework to analyze the effects of more general policies that affect both outcome distributions and choices using the general equilibrium framework described in Heckman (2001b). We leave that task for another occasion.

7   References

1. Atkinson, Anthony (1970). "On The Measurement of Inequality," *Journal of Economic Theory.*

2. Björklund, A. and Moffitt, R. (1987). "The Estimation of Wage Gains and Welfare Gains in Self-Selection," *Review of Economics and Statistics*, 69(1):42-9.

3. Card, David (1999). "Estimating the Causal Effect of Education,", in O. Ashenfelter and D. Card, editors, *Handbook of Labor Econometrics*, Chapter 10, 1801-1863.

4. Carneiro, Pedro; Hansen, Karsten T.; and Heckman, James J. (2001). "Educational Attainment and Labor Market Outcomes: Estimating Distributions of The Returns to Educational Interventions," unpublished paper presented as Klein Lecture, University of Pennsylvania, September 2001; First Draft presented at Midwest Econometrics Group (MEG), October, 2000, forthcoming, *International Economic Review*, 2002.

5. Harsanyi, J. (1955). "Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63(4) (August), 309-21.

6. Heckman, James (1990). "Varieties of Selection Bias," *American Economic Review*, 80(2):313-318.

7. Heckman, James J. (2001a). "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture (2001a)," *Journal of Political Economy* 109(4):673-748.

8. Heckman, James J. (2001b). "Accounting For Heterogeneity, Diversity and General Equilibrium In Evaluating Social Programs," *Economic Journal*, November, 2001

9. Heckman, J. and B. Honoré (1990). "The Empirical Content of the Roy Model," *Econometrica*, 58(5):1121-1149.

10. Heckman, J., R. LaLonde and J. Smith (1999). "The Economics and Econometrics of Active Labor Market Programs," in O. Ashenfelter and D. Cards, eds., *Handbook of Labor Economics*, Chapter 31, pp.

11. Heckman, J. and J. Smith (1993). "Assessing The Case for Randomized Evaluations of Social Programs," in K. Hansen and P. Madsen, eds, *Measuring Labor Market Outcomes*, (Copenhagen: Ministry of Labor).

12. Heckman, J. and J. Smith (1998). "Evaluating The Welfare State," in S. Strom (ed), *Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, Econometic Society, Monograph Series*, 16, (Cambridge University Press, 1998), Chapter 8, 241-318.

13. Heckman, J. and E. Vytlacil (2000). "Local Instrumental Variables," in C. Hsiao, K. Morimune, and J. Powells, (eds.), *Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya*, (Cambridge: Cambridge University Press, 2000), 1-46.

14. Heckman, J. and Vytlacil, E. (2001). "Structural Equations, Treatment Effects and Econometric Policy Evaluation," Fisher Schultz Lecture, World Meetings of the Econometric Society, Seattle, forthcoming in *Econometrica.*

15. Joreskog, K. and D. Sorbom (1979). *Advances in Factor Analysis and Structural Equations Models.* (Cambridge: Abt Associates).

16. Kane, T. (1994). "College Entry by Blacks since 1970: The Role of College Costs, Family Background, and the Returns to Education," *Journal of Political Economy*, 102 (October 1994): 878-911.

17. Persson, T. and G. Tabellini (2000). *Political Economics -Explaining Economic Policy.* (MIT Press: Cambridge, Massachusetts).

18. Rawls, J. (1971). *A Theory of Justice.* (Cambridge, MA: Harvard University Press).

19. Roemer, J. (1996). *Theories of Distributive Justice.* (Cambridge: Harvard University Press).

20. Roy, A. (1951). "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3:135-146.

21. Sen, A. K. (1997). *On Economic Inequality.* Second Edition; Oxford: Clarendon Press.

22. Sen, A. K. (2000). "Social Justice and the Distribution of Income," in *Handbook of Income Distribution* (eds. A. B. Atkinson and F. Bourguignon). (Amsterdam: North Holland).

23. Vickery, W. (1960). "Utility, Strategy and Social Decision Rules." *Quarterly Journal of Economics*, 74(4):507-35.

Appendix

Identifying Counterfactual Distributions Under Treatment Effect Assumptions

Heckman and Honoré (1990) show that in the context of the original Roy (1951) model under normality or exclusion restrictions, it is possible to identify the joint density of potential outcomes. The original Roy model sets $C = 0$. Sectoral choices are then determined solely by potential outcomes. This extra information identifies the full model and lets analysts identify the joint distributions of outcomes across policy states. If there is variation in C across persons, this method breaks down and it is only possible to identify $g(Y_0|S = 0)$ and $g(Y_1|S = 1)$, the conditional densities of the potential outcomes, as well as $\Pr(S = 1)$, but not the joint density, $g(Y_0, Y_1)$ (Heckman (1990)). Another special case that is discussed in Heckman (1992), is case I where $Y_1 \equiv Y_0 + \Delta$, and Δ is a constant. Then from the marginal distribution of Y_0 or Y_1 it is possible to form the joint distribution (Y_0, Y_1) which is degenerate. Heckman and Smith (1993) and Heckman, Smith and Clements (1997) generalize this case to assume that the persons at the qth percentile in the density of Y_0 are at the qth percentile of Y_1. Even without imposing this information, from the marginals it is also possible to bound the joint densities using classical results in probability theory. In practice these bounds turn out to be rather wide (Heckman and Smith (1993); Heckman, Smith and Clements (1997)).

In our source paper (Carneiro, Hansen and Heckman (2001)), we generate the distributions of potential outcomes using a panel data factor structure model. For the details of our method we refer the reader to our source paper. Here we present the intuitive idea that underlies our method and in the text we report its application. We discuss the most

elementary case, leaving a complete discussion of the more general case for our companion paper.

Suppose that the mean of C depends on shifter variables Z that do not affect (are independent of) potential outcomes (Y_0, Y_1). These are instruments. Suppose that for some values of Z within available samples we observe

$$\Pr(S = 1|Z) \doteq 1 \qquad Z \in \mathcal{Z}_1$$

while for other values of Z

$$\Pr(S = 1|Z) \doteq 0 \qquad Z \in \mathcal{Z}_0.$$

Thus if Z is tuition, people who face a low tuition cost (possibly even a large subsidy) are almost surely likely to go to college while those who face a very high tuition cost are almost certainly likely not to go to school.[8] We assume that the distribution of potential outcomes is the same in these subsets as they are in the overall distribution. Thus we can identify the marginal distribution of Y_1 from the first sample and the marginal distribution of Y_0 from the second sample.

Within these samples, we observe post schooling outcomes

$$Y_{0t}, \qquad t = 1, \ldots, T, \qquad \text{for } Z \in \mathcal{Z}_0,$$

$$Y_{1t}, \qquad t = 1, \ldots, T, \qquad \text{for } Z \in \mathcal{Z}_1.$$

From these data we can form the joint densities of each outcome over time on $f(y_{01}, \ldots, y_{0T})$

---

[8]This is the version of identification at infinity discussed in Heckman (1990).

and $f(y_{11}, ..., y_{1T})$, but not the joint densities over time over both outcomes.

Now suppose that Y_{0t} and Y_{1t} are both generated by a common factor f (e.g., ability, motivation) so that

$$Y_{0t} = \mu_{0t} + \alpha_{0t}f + \varepsilon_{0t}, \qquad t = 1, \ldots, T,$$

$$Y_{1t} = \mu_{1t} + \alpha_{1t}f + \varepsilon_{1t}, \qquad t = 1, \ldots, T,$$

where the ε_{0t} and ε_{1t} are mutually independent of each other, f, and all other $\varepsilon_{0t'}, \varepsilon_{0t''}, t \neq t', t''$.[9] All of these error components are assumed to have mean zero. A common factor generates both potential outcomes. If we can get our hands on the distribution of the common factor, we can compute the joint distribution of counterfactuals.

Within each regime we can compute the following covariances:

$$\text{Cov}(Y_{0t}, Y_{0t'}) = \alpha_{0t}\alpha_{0t'}\sigma_f^2, \quad t \neq t', t, t' = 1, \ldots, T, \quad \text{for } Z \in \mathcal{Z}_0,$$

$$\text{Cov}(Y_{1t}, Y_{1t'}) = \alpha_{1t}\alpha_{1t'}\sigma_f^2, \quad t \neq t', t, t' = 1, \ldots, T, \quad \text{for } Z \in \mathcal{Z}_1$$

For concreteness suppose $T = 3$, so we have three panel wage observations. Then

$$\text{Cov}(Y_{01}, Y_{02}) = \alpha_{01}\alpha_{02}\sigma_f^2,$$

$$\text{Cov}(Y_{01}, Y_{03}) = \alpha_{01}\alpha_{03}\sigma_f^2, \qquad \text{for } Z \in \mathcal{Z}_0,$$

$$\text{Cov}(Y_{01}, Y_{02}) = \alpha_{02}\alpha_{03}\sigma_f^2,$$

and

---

[9] The means may depend on the covariates.

$$\text{Cov}(Y_{11}, Y_{12}) = \alpha_{11}\alpha_{12}\sigma_f^2,$$

$$\text{Cov}(Y_{11}, Y_{13}) = \alpha_{11}\alpha_{13}\sigma_f^2, \qquad \text{for } Z \in \mathcal{Z}_1,$$

$$\text{Cov}(Y_{12}, Y_{13}) = a_{12}a_{13}\sigma_f^2.$$

If we assume $\alpha_{01} = 1$ or $\sigma_f^2 = 1$, we can identify all of the rest of the factor loadings.[10]

With this information in hand, we can identify the variances of the uniquenesses, $\varepsilon_{0t}, \varepsilon_{1t}$ of the outcomes:

$$\text{Var}(\varepsilon_{0t}) = \text{Var}(Y_{0t}) - \alpha_{0t}^2\sigma_f^2 \qquad t = 1, \ldots, T$$

$$\text{Var}(\varepsilon_{1t}) = \text{Var}(Y_{1t}) - \alpha_{1t}\sigma_f^2 \qquad t = 1, \ldots, T$$

Suppose that $f, \varepsilon_{0t}, \varepsilon_{12t}, t = 1, \ldots, T$ are normally distributed. Then from the information just presented obtained from the subsamples associated with $\mathcal{Z}_0$ and $\mathcal{Z}_1$ we can identify the density of f and hence the joint density of $(Y_{01}, Y_{11}, \ldots, Y_{0T}, Y_{1T})$. Using the outcome data within schooling choices we can identify the distribution of f and hence estimate the

---

[10]Proof:

$$\frac{\text{Cov}(Y_{01}, Y_{02})}{\text{Cov}(Y_{01}, Y_{03})} = \frac{\alpha_{02}}{\alpha_{03}}.$$

Given $\sigma_f^2 = 1$, we can use

$$\text{Cov}(Y_{02}, Y_{03}) = \alpha_{02}\alpha_{03}.$$

to obtain

$$(\alpha_{03})^2 = \frac{\text{Cov}(Y_{02}, Y_{03})}{\text{Cov}(Y_{01}, Y_{02})\text{Cov}(Y_{02}, Y_{03})}$$

and we can identify α_{03} up to sign and hence can identify α_{02} and α_{01}. If we normalize $\alpha_{01} = 1$, we can identify, α_{02}, α_{03} up to sign and σ_f^2. Since the sign of f is unknown, the sign of the factor loadings is unknown. Using the data on Y_1, under either normalization, we can identify $\alpha_{11}, \alpha_{12}, \alpha_{13}$ since σ_f^2 is known.

joint distribution of schooling choices across potential outcomes.

In our companion paper we show that we can obtain this joint density without a normality assumption for f or $\varepsilon_{0t}, \varepsilon_{1t}, t = 1, ..., T$. We extend our analysis to allow for vector f so there may be many factors, not just one. We show that it is possible to nonparametrically identify the joint density of potential outcomes provided that the number of panel data wage measurements is large, in a sense we make precise in our companion paper, relative to the number of factors.[11] We do not need to invoke "identification at infinity" *i.e.* we can dispense with the requirement that there are subsets of $\mathcal{Z}$ where there is no selection. We also consider a model with multiple discrete choices (schooling levels) instead of just two. With these counterfactual distributions determined, we can identify the impact of social policy on the distributions of outcomes and returns.

---

[11]In our companion paper, we show how indicators of f can be used to supplement, or replace, panel data. This type of identification is familiar to users of LISREL (see Joreskog and Sorbom, 1979).

Figure 1
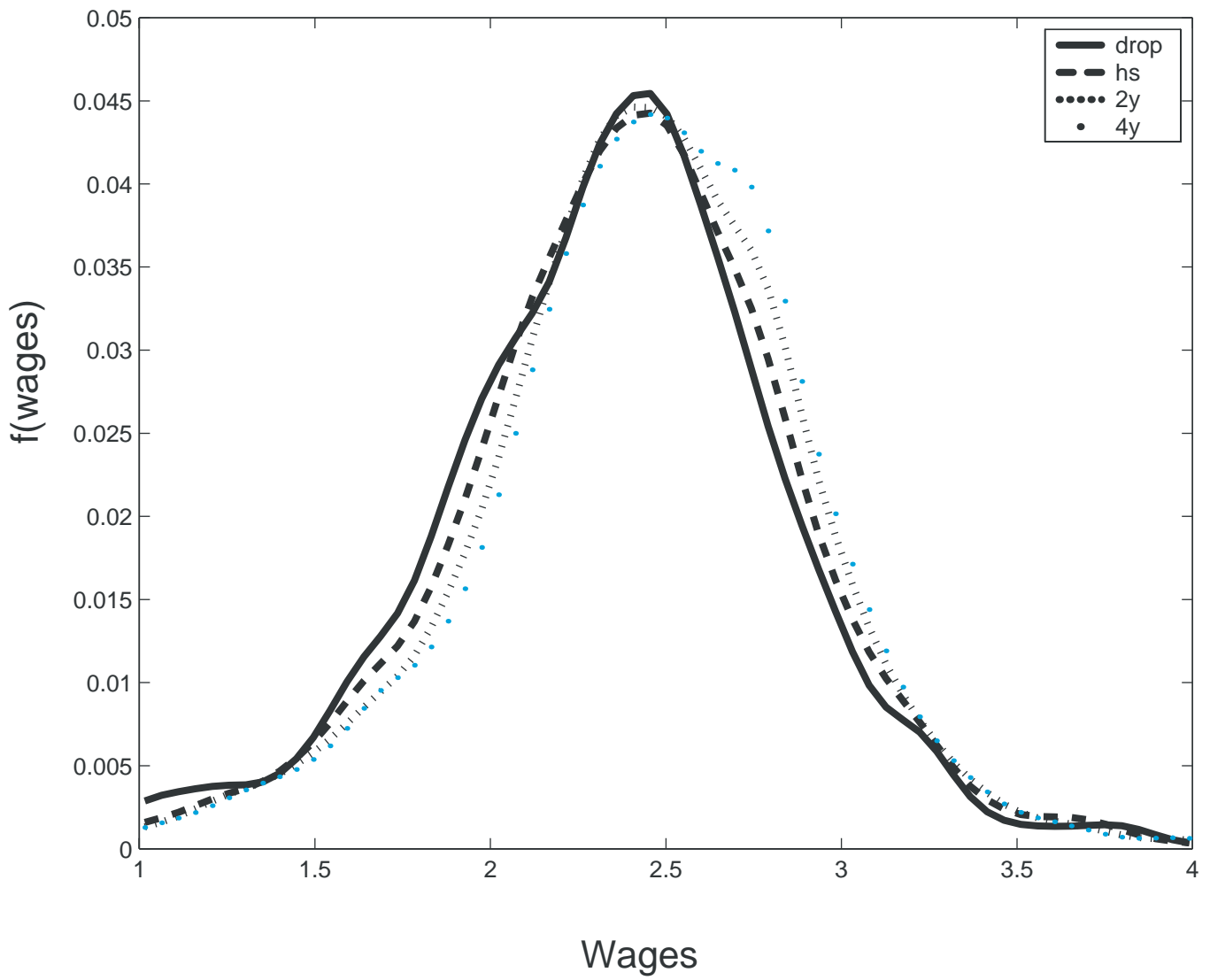## Distributions of Wages, High School Graduates
White Males, age 29 from NLSY

Figure 2
Distributions of Wages, College
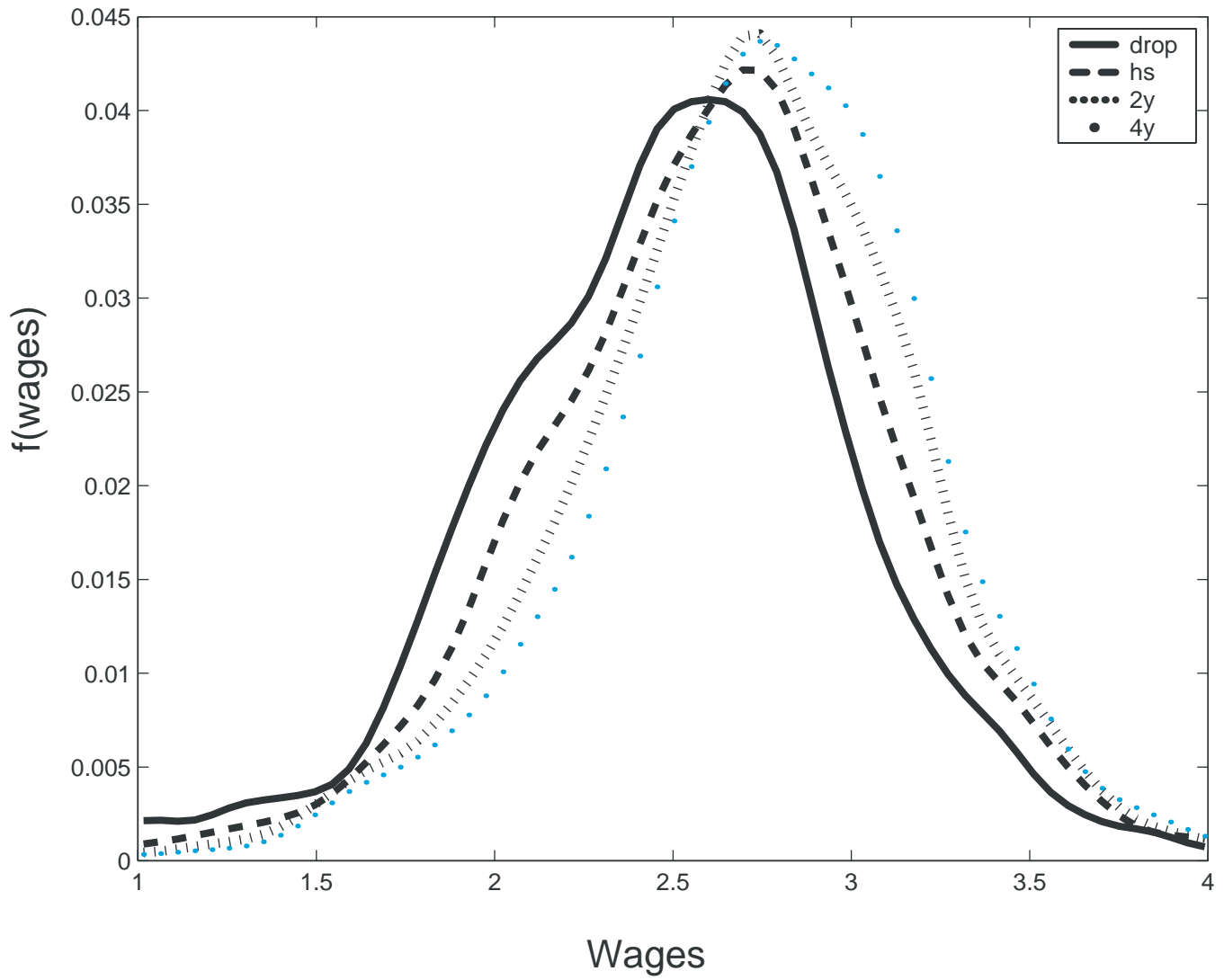
White Males, age 29 from NLSY

Figure 3
Distributions of Returns to High School
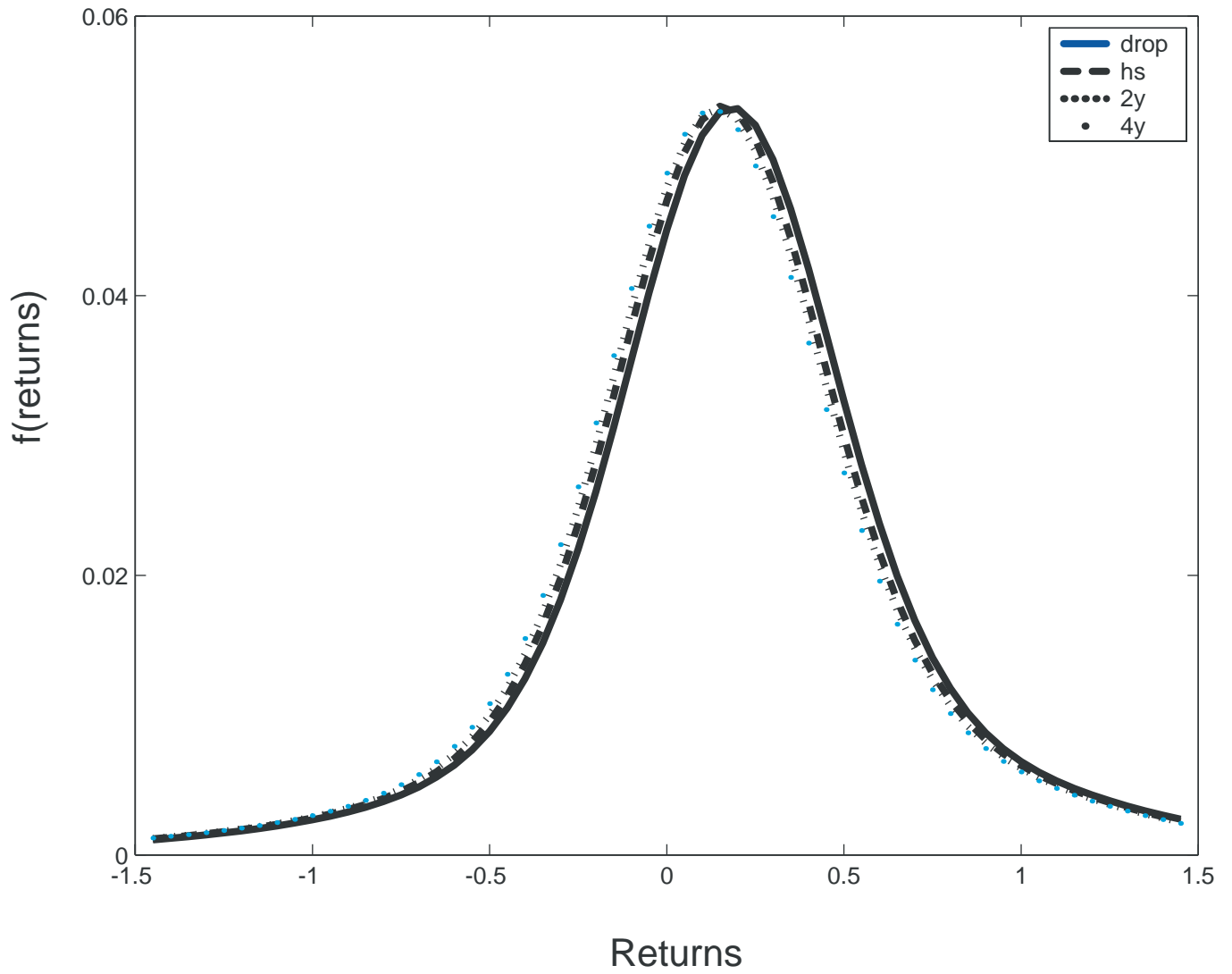### White Males, age 29 from NLSY

Figure 4
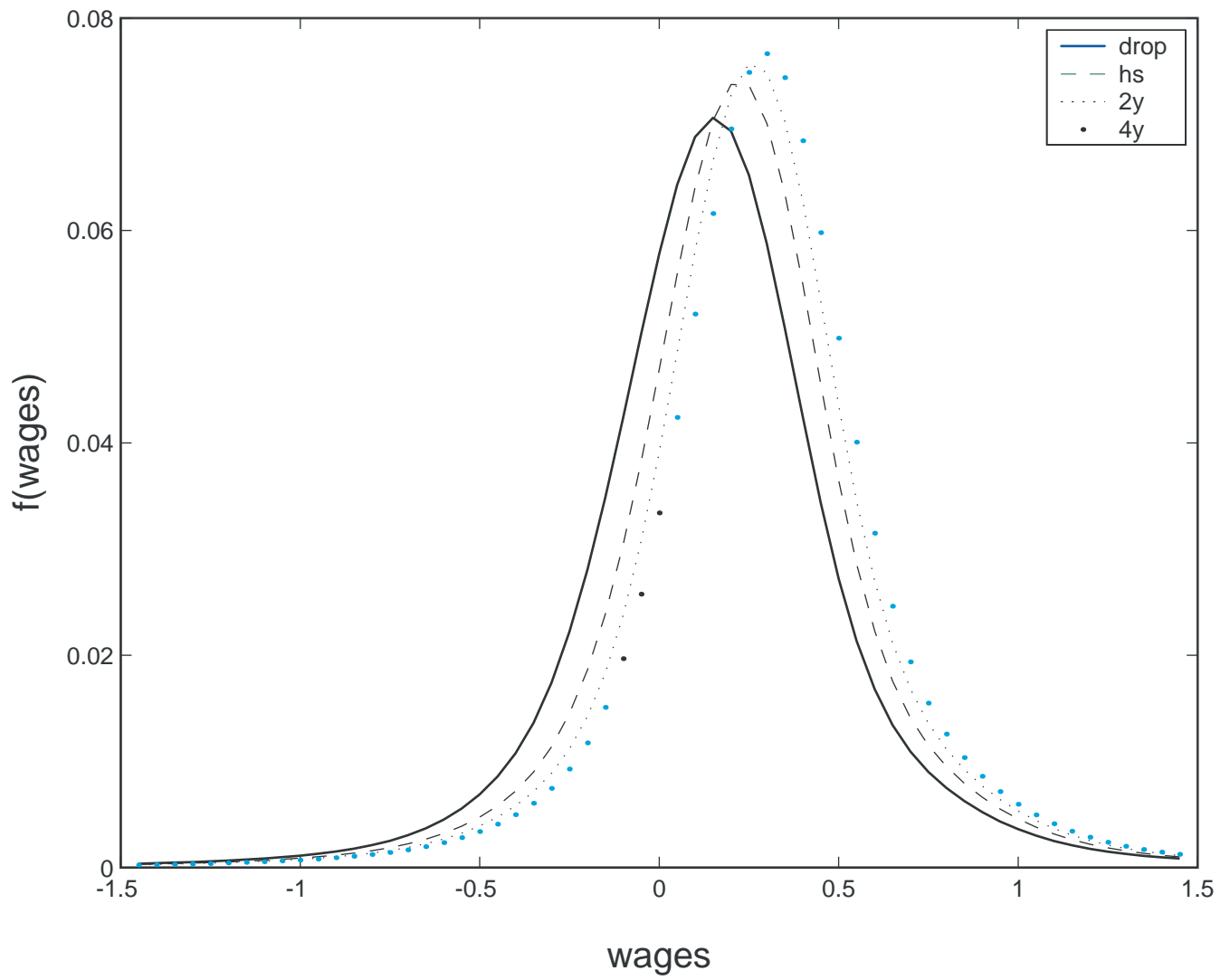Distribution of Returns to College vs. High School
White Males, Age 29 from NLSY

Figure 5
Marginal Treatment Effect
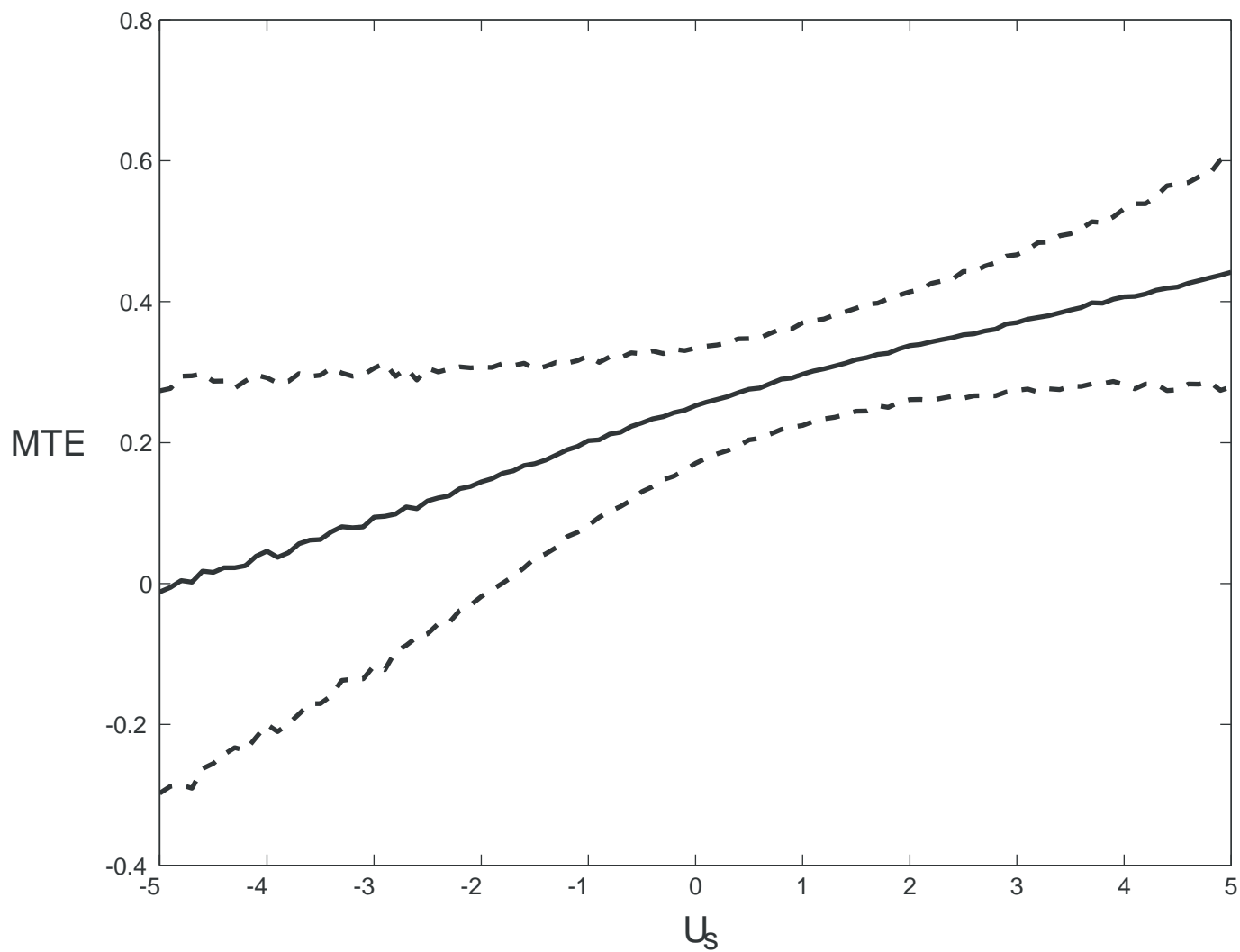## High School - College
NLSY, White Males

Figure 6

People affected by full subsidy to community college tuition by decile of initial overall wage distribution
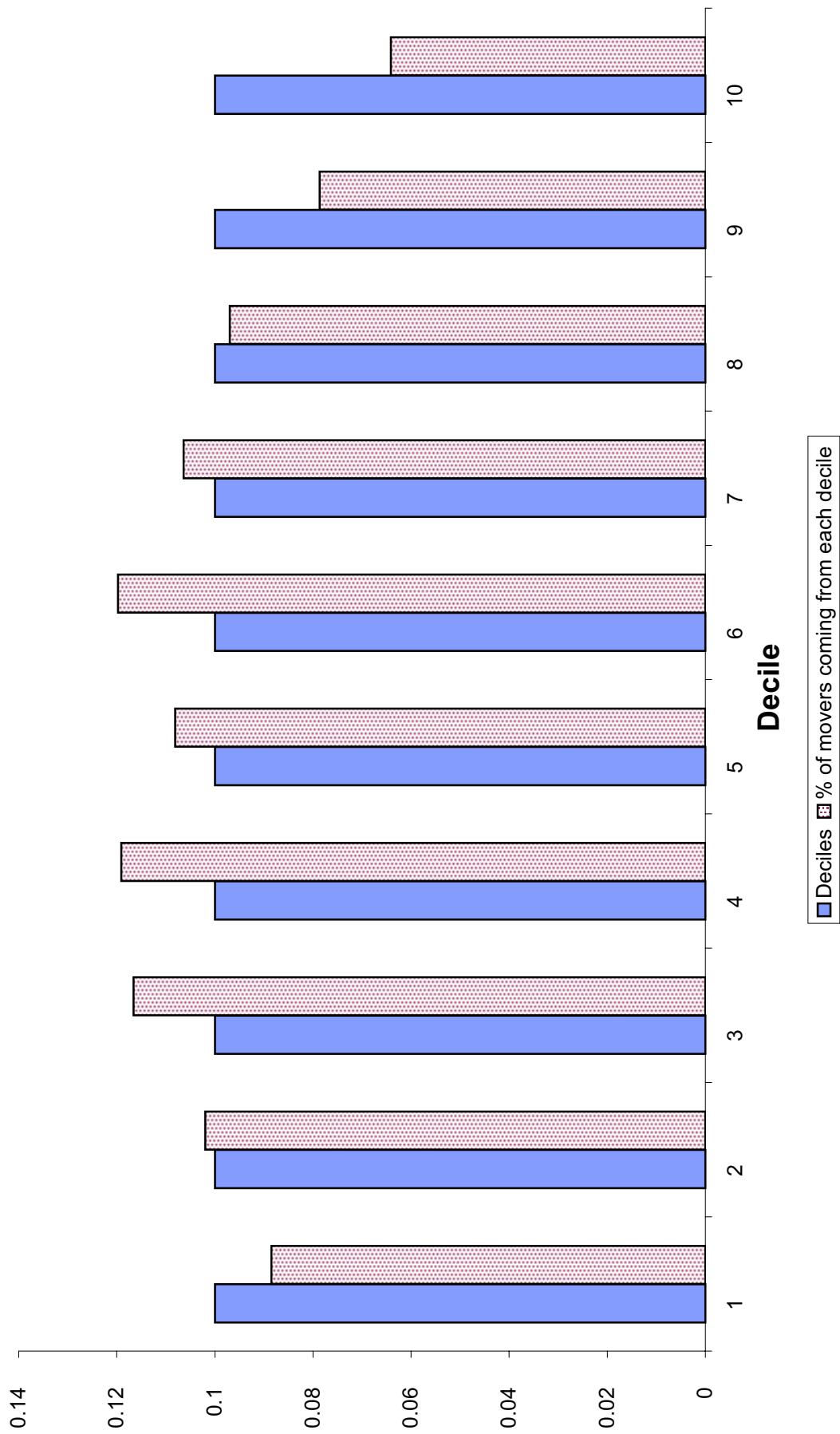
Figure 7

People affected by making distance to 4y college = 0 by decile of initial overall wage distribution