Published in final edited form as:
J Polit Econ. 2015 April ; 123(2): 413–443. doi:10.1086/679498.

The Generalized Roy Model and the Cost-Benefit Analysis of Social Programs*

Philipp Eisenhauer,
The University of Chicago

James J. Heckman, and
The University of Chicago, University College Dublin, American Bar Foundation

Edward Vytlacil
New York University

Abstract

The literature on treatment effects focuses on gross benefits from program participation. We extend this literature by developing conditions under which it is possible to identify parameters measuring the cost and net surplus from program participation. Using the generalized Roy model, we nonparametrically identify the cost, benefit, and net surplus of selection into treatment without requiring the analyst to have direct information on the cost. We apply our methodology to estimate the gross benefit and net surplus of attending college.

Keywords

Cost-Benefit Analysis; Treatment Effects; Returns and Costs to Education

1 Introduction

The traditional approach to the evaluation of public policy compares the benefits and costs of policies. Measures of net surplus are used to determine whether policies should be undertaken (see Hotelling, 1938; Tinbergen, 1956; Harberger and Jenkins, 2002; and Chetty, 2009). The recent literature on program evaluation, or "treatment effects", focuses on gross benefits of policies and considers neither the marginal costs nor the perceived surplus associated with the programs being evaluated.[1]

*This research was supported by NIH R01-HD32058, NSF SES-024158, NSF SES-05-51089, NICHD R37HD065072, NIH R01-HD54702, The Pritzker Children's Initiative, the American Bar Foundation, the Human Capital and Economic Opportunity Working Group—an initiative of the Becker Friedman Institute for Research and Economics—funded by the Institute for New Economic Thinking (INET), and a European Research Council grant hosted by the University College Dublin, DEVHEALTH 269874. The website for this paper is https://heckman.uchicago.edu/generalized-roy-model. The views expressed in this paper are those of the authors and not necessarily those of the funders or commentators mentioned here. We have greatly benefited from comments received from Ismael Mourifié. We thank Luke Schmerold, Edward Sung, and Jake Torcasso for their outstanding research assistance.

Philipp Eisenhauer, Department of Economics, The University of Chicago, 1126 East 59th Street, Chicago, IL 60637, peisenha@uchicago.edu
James J. Heckman, Department of Economics, The University of Chicago, 1126 East 59th Street, Chicago, IL 60637, Phone: (773) 702-0634, jjh@uchicago.edu
Edward Vytlacil, Department of Economics, New York University, 19 W. 4th Street, 6FL, New York, NY 10012, Phone: (212) 992-8682, vytlacil@nyu.edu

We extend this literature using the generalized Roy model. In it, agents choose treatment if their expected surplus from doing so is positive, so the benefit outweighs the subjective cost. We present conditions under which we can use the economics of the model to identify cost and surplus parameters even without direct information on the costs of treatment. Information on revealed choices creates a simple relationship between the cost and benefit parameters: for individuals who are indifferent towards treatment participation, the benefit equals the cost and the surplus is zero. Building on existing identification results for benefit parameters, we show how to identify surplus and cost parameters by varying the margin of indifference. Our identification analysis applies traditional exclusion restrictions that separately shift costs and benefits from treatment. We use cost shifters to identify the benefit of treatment, and benefit shifters to vary the margin of indifference and thus to identify the cost of the treatment.

Our analysis complements and extends the work by Björklund and Moffitt (1987) who first noted the duality between cost and benefit parameters in the generalized Roy model. They estimate marginal gains and surpluses for policies within a parametric normal generalized Roy model. They use structural econometric methods to identify the components of the cost and benefit functions. This paper extends their analysis to a more general setting. It develops and applies a nonparametric identification analysis of benefits, costs, and surpluses without the need to identify all of the ingredients of a fully specified structural model. This approach implements *Marschak's Maxim* (Heckman, 2010) by directly estimating the cost, benefit, and surplus parameters rather then constructing them from the estimates of a full structural model.

We present *ex ante* and *ex post* analyses of costs and benefits. Applying our methods to the data on *ex post* gross benefits analyzed by Carneiro et al. (2011), we find that heterogeneity in benefits, and not costs, is the main driver of the variability in the decision to attend college.

Our analysis is reminiscent of the Heckman (1974) model of female labor supply. In that analysis, the econometrician observes the offered wage only for the agents who choose to work. The economist does not observe the reservation wage of any agent. Yet, his analysis identifies the parameters of the offered wage equation and the reservation wage equation by using the implication of the underlying economic model that agents decide to work if the offered wage exceeds the reservation wage.[2] In our analysis, we observe program outcomes for agents who select into treatment, and we observe the no treatment outcome for the agents who do not select into treatment. We do not observe the cost of treatment for any agent. Yet, using the economics of the model, we are able to identify the average benefit and average cost of treatment parameters by exploiting the agent's decision rule of selecting into treatment if the benefit exceeds the cost.

Our analysis is very different from analyses using randomized experiments to infer treatment effects. In commonly implemented randomizations, it is not possible to identify

[1]See the discussion in Heckman and Vytlacil (2007) and Heckman (2010).
[2]The same methodology applies to search theory, see Flinn and Heckman (1982).

the choice probability (Heckman, 1992; Heckman and Smith, 1995). Instead of using randomization to bypass problems of self-selection, we exploit the information that agents self-select into treatment and infer information on the cost of the treatment that cannot be recovered by standard randomized experiments.

The paper unfolds in the following way. Section 2 introduces the generalized Roy model. Section 3 reviews the average benefit of treatment parameters from Heckman and Vytlacil (1999, 2005, 2007), and develops and analyzes the dual cost parameters that match the benefit parameters. Section 4 presents our identification analysis of the cost and surplus parameters. Section 5 extends our analysis to allow agents to have imperfect foresight about future outcomes. We apply our analysis to study the decision to attend college in Section 6. Section 7 concludes.

2 The Generalized Roy Model

Suppose there are two potential outcomes (Y_0, Y_1), and a choice indicator D with $D = 1$ if the agent selects into treatment so that Y_1 is observed and $D = 0$ if the agent does not select into treatment so that Y_0 is observed. Anticipating our empirical analysis, Y_1 is the annualized flow of income from college, and Y_0 is the annualized flow of income from high school. The observed outcome Y can be written in switching regression form (Quandt, 1958, 1972):

$$Y = DY_1 + (1-D)Y_0, \quad (2.1)$$

where $E(Y_j \mid X) = \mu_j$ and

$$Y_j = \mu_j(X) + U_j \quad (2.2)$$

for $j = 0, 1$. X is a vector of regressors observed by the economist while (U_0, U_1) are not. Combining Equations (2.1) and (2.2),

$$Y = \mu_0(X) + \{[\mu_1(X) - \mu_0(X)] + U_1 - U_0\}D + U_0.$$

The individual gross benefit of treatment associated with moving an otherwise identical person from state “0” to “1” is $B = Y_1 - Y_0$ and is defined as the causal effect on Y of a *ceteris paribus* move from “0” to “1”. Defining $E(C \mid Z) = \mu_C(Z)$, the subjective cost of choosing treatment as perceived by the agent is

$$C = \mu_C(Z) + U_C, \quad (2.3)$$

where Z is an observed random vector of cost shifters and U_C is a random variable unobserved by the econometrician. Individuals choose treatment if the perceived benefit from treatment is greater than the subjective cost:

$$D = 1 \quad \text{if} \quad S \geq 0; \quad D = 0 \quad \text{otherwise}, \quad (2.4)$$

where S is the surplus, i.e. the net benefit, from treatment:

$$\begin{aligned}S&=(Y_1-Y_0)-C\\&=\{[\mu_1(X)-\mu_0(X)]-\mu_C(Z)\}-[U_C-(U_1-U_0)]\\&=\mu_S(X,Z)-V,\end{aligned}$$

with $\mu_S(X, Z) = [\mu_1(X) - \mu_0(X)] - \mu_C(Z)$ and $V = U_C - (U_1 - U_0)$. Our identification analysis of cost and surplus parameters does not assume particular functional forms for μ_0, μ_1, and μ_C, nor does it assume that the distributions of U_0, U_1, and U_C are of a known parametric form.

The original Roy (1951) model assumes that there are no observed regressors, X, that the cost of treatment is identically zero (i.e. $\mu_C = 0$, $U_C = 0$), and that $(U_0, U_1) \sim N(0, \Sigma)$. Heckman and Honoré (1990) present an identification analysis for a nonparametric version of the Roy model using variation in regressors and making no parametric assumption on the distribution of (U_0, U_1). Their version of the Roy model also imposes the condition that the cost of treatment is identically zero. In contrast, we allow for nonzero cost of treatment. In fact, for our identification analysis we require nondegenerate cost of treatment and observed cost shifters.[3] From the point of view of the observing economist, (X, Z) is observed and (U_1, U_0, U_C) is unobserved. This model assumes that agents know the gross benefit, $B = Y_1 - Y_0$, of treatment. We show in Section 5 that our results extend to a broader class of models, where agents only have imperfect foresight about the benefits of treatment. This model also supposes that there is no other aspect of the benefit of treatment than $Y_1 - Y_0$. Implicitly, any subjective benefits of the program are incorporated into the costs of treatment, i.e. the cost function includes the subjective benefits of the treatment. For example, if job training allows the individual to work in a job with preferred amenities, this is modeled as a (negative) contribution to the subjective cost of treatment. The classification of effects in either positive benefits or negative subjective cost (or vice versa) does not affect the definition of the surplus. To simplify the exposition, we suppose that Z and X do not contain any common elements. Thus, all of our analysis can be seen as implicitly conditioning on all common elements of X and Z.

We make the following technical assumptions:

(A-1) *(U_0, U_1, U_C) is independent of (X, Z).*

(A-2) *The distribution of $\mu_C(Z)$ conditional on X is absolutely continuous with respect to Lebesgue measure.*

[3]Because Heckman and Honoré (1990) impose a Roy model with zero cost of treatment, they are able to identify the joint distribution of (U_0, U_1). In contrast, because we allow for nonzero cost of treatment (and, in particular, for unobserved costs of treatment), we are unable to identify the dependence between U_0 and U_1 which precludes the identification of some potentially interesting economic parameters. See Heckman (1990), Heckman and Smith (1998) and Heckman et al. (1997b) for related analysis. With additional information, the joint distribution of (U_1, U_0, U_C) can be identified. See, e.g., Carneiro et al. (2003), Aakvik et al. (2005), and Abbring and Heckman (2007). D'Haultfoeuille and Maurel (2013) identify the cost of treatment in a related Roy model in which the cost of treatment is a deterministic function of observed covariates. Their identification strategy is fundamentally different from ours, and critically relies on the restriction that the cost of treatment is constant conditional on covariates.

(A-3) *The distribution of* $V = U_C - (U_1 - U_0)$ *is absolutely continuous with respect to Lebesgue measure and has a cumulative distribution function that is strictly increasing.*

(A-4) *The population means* $E|Y_1|$, $E|Y_0|$ *and* $E|C|$ *are finite.*

Assumption (A-1) assumes that (U_0, U_1, U_C) is independent of (X, Z). Thus, D is endogenous but all other regressors in both the treatment equation and the outcome equation are exogenous. We implicitly condition on any regressors that enter both the outcome equations and the cost equation. Thus, this condition should be interpreted as an independence assumption for the error terms with regard to the unique elements of X and Z conditional on the regressors that enter both equations. No independence condition is required for the common elements. We also do not impose any restrictions on the dependence among the unobservables. (A-2) requires that there exists at least one continuous component of Z conditional on X. This assumption will only be required for our identification analysis, and is not needed for our definition or analysis of the cost and surplus parameters. (A-3) is a regularity condition. It allows for the possibility that U_C is degenerate (costs do not vary conditional on Z) or that $U_1 - U_0$ is degenerate (benefits do not vary conditional on X), though not both. Assumption (A-4) is required for the mean benefit and cost parameters to be well defined. An implication of our model with Assumptions (A-1) and (A-3) is that $0 < \Pr(D = 1 \mid X, Z) < 1$ w.p.1, so that there is a treated group and a control group for almost all (X, Z). Note that this restriction still allows the support of the distribution of $\Pr(D = 1|X, Z)$ to be the full unit interval.

Let $P(X, Z)$ denote the probability of selecting into treatment given (X, Z). Statisticians call this the "propensity score" $P(X, Z) \equiv \Pr(D = 1 \mid X, Z) = F_V(\mu_S(X, Z))$, where $F_V(\cdot)$ denotes the distribution of V.[4] We sometimes denote $P(X, Z)$ by P, suppressing the (X, Z) argument. We also work with U_S, a uniform random variable ($U_S \sim \text{Unif}[0, 1]$) defined by $U_S = F_V(V)$. Different values of U_S denote different quantiles of V. Given our previous assumptions, F_V is strictly increasing, and $P(X, Z)$ is a continuous random variable conditional on X.

The generalized Roy model presented in this paper is a special case of the model of Heckman and Vytlacil (1999, 2005). Under Assumptions (A-1)–(A-4), the model of Equations (2.1)–(2.4) implies the model and assumptions of Heckman and Vytlacil (1999, 2005). From the analysis of Vytlacil (2002), the more general model is equivalent to the conditions that justify the Local Average Treatment Effect (LATE) model of Imbens and Angrist (1994). We impose more restrictions here. In particular, we impose the generalized Roy model and the corresponding assumptions that will allow us to exploit its structure for identification of subjective cost parameters. As in the conventional Roy model (Heckman and Sedlacek, 1985), we assume additive separability in the outcome equations. Additive separability is not required in Heckman and Vytlacil (1999, 2005), but is required by our analysis in order to obtain additive separability in the latent index equation consistent with the generalized Roy model.[5] Thus our assumptions are most appropriate for continuous

[4]We will refer to the cumulative distribution function of a random vector A by $F_A(\cdot)$ and to the cumulative distribution function of a random vector A conditional on random vector B by $F_{A|B}(\cdot)$. We write the cumulative distribution function of A conditional on $B = b$ by $F_{A|B}(\cdot \mid b)$.

outcome variables, and we exclude discrete outcomes from our analysis. We also assume conditions on X that are not required in Heckman and Vytlacil (1999, 2005) to identify the gross benefit parameters. Their analysis conditions on X, and thus does not need to assume that X is independent of the error vector. In contrast, in order to use the generalized Roy model to recover subjective cost parameters, we require that the unique elements X are independent of the error vector.[6]

3 Benefit, Cost, and Surplus Parameters

This section defines and analyzes the benefit, cost, and surplus parameters. We maintain the model of Equations (2.1)–(2.4), and invoke Assumptions (A-1) and (A-3)–(A-4). We do not require Assumption (A-2) for the definition of the parameters, but do require it for our identification analysis.

Standard treatment effect analyses identify averaged parameters of the gross benefit of treatment, $B = Y_1 - Y_0$. The most commonly studied treatment effect parameter is the average benefit of treatment $B^{ATE}(x) \equiv E(Y_1 - Y_0 \mid X = x) = \mu_1(x) - \mu_0(x)$. This is the effect of assigning treatment randomly to everyone of type $X = x$ assuming full compliance, and ignoring any general equilibrium effects. Another commonly used parameter is the average benefit of treatment on persons who actually take the treatment, referred to as the benefit of treatment on the treated: $B^{TT}(x) \equiv E(Y_1 - Y_0 \mid X = x, D = 1) = \mu_1(x) - \mu_0(x) + E(U_1 - U_0|X = x, D = 1)$. Heckman and Vytlacil (1999, 2005) unify a broad class of treatment effect parameters including the $B^{ATE}(x)$ and $B^{TT}(x)$ through the marginal benefit of treatment, defined as $B^{MTE}(x, u_S) \equiv E(Y_1 - Y_0|X = x, U_S = u_S) = \mu_1(x) - \mu_0(x) + E(U_1 - U_0|U_S = u_S)$. $B^{MTE}(x, u_S)$ is the treatment effect parameter that conditions on the unobserved desire to select into treatment.

The conventional analysis of treatment effects does not define, identify, or estimate any aspect of the cost of the treatment. We define a set of cost parameters parallel to the benefit parameters, where cost is the subjective cost as perceived by the agent. Thus, we define the average cost of treatment, the average cost of treatment on the treated, and the marginal cost of treatment as follows:

$$\begin{array}{rcccl} C^{ATE}(z) & = & E(C|Z=z) & = & \mu_C(z) \\ C^{TT}(z) & = & E(C|Z=z, D=1) & = & \mu_C(z)+E(U_C|Z=z, D=1) \\ C^{MTE}(z, u_S) & = & E(C|Z=z, U_S=u_S) & = & \mu_C(z)+E(U_C|U_S=u_S). \end{array}$$

Recalling that $S = B - C = \mu_S(X, Z) - V$, where $\mu_S(X, Z) = [\mu_1(X) - \mu_0(X)] - \mu_C(Z)$ and $V = U_C - (U_1 - U_0)$, we can define the corresponding surplus parameters:

[5]Recall again that we are implicitly conditioning on all common elements of (X, Z), so that these need not be additively separable from the error term.

[6]In this respect, our analysis is broadly analogous to the identification strategies and conditions of Vytlacil and Yildiz (2007) and Shaikh and Vytlacil (2011), who also require that there be exogenous regressors in the outcome equation that is excluded from the treatment choice equation, and they exploit variation in such regressors for identification.

$$\begin{aligned} S^{ATE}(x,z) &= E(S|X{=}x,Z{=}z) &= \mu_S(x,z) \\ S^{MTE}(x,z,u_S) &= E(S|X{=}x,Z{=}z,U_S{=}u_S) &= \mu_S(x,z) - E(V|U_S{=}u_S) \end{aligned}$$

and

$$S^{TT}(x,z) = E(S|X{=}x,Z{=}z,D{=}1) = \mu_S(x,z) - E(V|X{=}x,Z{=}z,D{=}1).$$

With these parameters, we can answer questions not only about the outcome change from treatment, but also about the subjective cost of treatment and the net surplus as well. As the surplus from treatment participation $S^{TT}(x, z)$ is always positive among the treated, it follows immediately that $B^{TT}(x) > C^{TT}(z)$ holds as well. Following Heckman and Vytlacil (1999, 2005), we can represent the average treatment effects and treatment on the treated as averaged versions of the marginal effects of treatment:

$$\begin{aligned} B^{ATE}(x) &= \int_0^1 B^{MTE}(x,u_S)du_S \\ B^{TT}(x) &= \int_0^1 B^{MTE}(x,u_S)\frac{1-F_{P|X}(u_S|x)}{\int_0^1(1-F_{P|X}(t|x))dt}du_S. \end{aligned} \quad (3.1)$$

Following the same line of argument as used by Heckman and Vytlacil (1999, 2005),

$$\begin{aligned} C^{ATE}(z) &= \int_0^1 C^{MTE}(z,u_S)du_S \\ C^{TT}(z) &= \int_0^1 C^{MTE}(z,u_S)\frac{1-F_{P|Z}(u_S|z)}{\int_0^1(1-F_{P|Z}(t|z))dt}du_S, \end{aligned} \quad (3.2)$$

and

$$\begin{aligned} S^{ATE}(x,z) &= \int_0^1 S^{MTE}(x,z,u_S)du_S \\ S^{TT}(x,z) &= \frac{1}{P(x,z)}\int_0^{P(x,z)} S^{MTE}(x,z,u_S)du_S. \end{aligned} \quad (3.3)$$

We now establish relationships among these parameters. First, consider the marginal surplus parameter. Recall that $U_S = F_V(V)$ with F_V strictly increasing. Thus $U_S = u_S$ is equivalent to $V{=}F_V^{-1}(u_S)$, and

$$S^{MTE}(x,z,u_S) = \mu_S(x,z) - E\left(V|V{=}F_V^{-1}(u_S)\right) = \mu_S(x,z) - F_V^{-1}(u_S).$$

F_V^{-1} is strictly increasing, and thus $S^{MTE}(x, z, u_S)$ is strictly decreasing in u_S. Individuals with low u_S want to enter the program the most and are those with the highest surplus from the program, while individuals with high u_S want to enter the program the least and have the smallest surplus from the program. Using the fact that F_V is strictly increasing and that $P(X,$

$Z) = F_V(\mu_S(X, Z))$, conditioning on $u_s = P(x, z)$ is equivalent to conditioning on $V = \mu_S(x, z)$. Thus

$$S^{MTE}(x,z,P(x,z)){=}\mu_{S}(x,z){-}E\left(V|V{=}\mu_{S}(x,z)\right){=}0.$$

An individual with $u_S = P(x, z)$ is an individual who is indifferent between being treated and untreated if assigned $X = x$ and $Z = z$. Since $S^{MTE}(x, z, u_S)$ is strictly decreasing in u_S, $S^{MTE}(x, z, u_S)$ is positive for $u_S < P(x, z)$, is equal to zero at $u_S = P(x, z)$, and is negative if $u_S > P(x, z)$. If we instead fix evaluation point u_S and consider how $S^{MTE}(x, z, u_S)$ varies with (x, z), $S^{MTE}(x, z, u_S)$ will be positive for all (x, z) such that $P(x, z) > u_S$ and will be negative for all (x, z) such that $P(x, z) < u_S$.

We have thus far discussed only the marginal surplus function. Using the relationship $S^{MTE}(x, z, u_S) = B^{MTE}(x, u_S) - C^{MTE}(z, u_S)$, we can translate statements about $S^{MTE}(x, z, u_S)$ into inequalities about the marginal benefit and marginal cost functions:

$$\begin{array}{lllll} B^{MTE}(x,u_S) & > & C^{MTE}(z,u_S) & \forall & (x,z,u_S)\,\text{s. t.}\;\; P(x,z){>}u_S \\ B^{MTE}(x,u_S) & = & C^{MTE}(z,u_S) & \forall & (x,z,u_S)\,\text{s. t.}\;\; P(x,z){=}u_S \\ B^{MTE}(x,u_S) & < & C^{MTE}(z,u_S) & \forall & (x,z,u_S)\,\text{s. t.}\;\; P(x,z){<}u_S. \end{array}$$

The benefit and cost parameters coincide when evaluated at $u_S = P(x, z)$, because at this point the marginal cost equals the marginal benefit. We exploit this equality at the margin of indifference in the next section to achieve identification of the cost parameters.

To fix ideas, in Figure 1 we display the full set of marginal effects for a numerical example. We plot the marginal effect functions for fixed values of (x, z), where it happens that $P(x, z) = 0.50$. Individuals at that margin, $u_S = 0.50$, have their benefit of treatment just offset by their subjective cost and are thus indifferent between participation in treatment and nonparticipation. The benefits are positive, but so are the costs. Overall, the surplus is zero. For $u_S < 0.50$, the marginal benefit function lies above the marginal cost function and thus the marginal surplus is strictly positive. The reverse is true for $u_S > 0.50$.

This example is constructed to have intuitive properties, with the marginal benefit of treatment $B^{MTE}(x, u_S)$ decreasing in u_S and the marginal cost of treatment $C^{MTE}(z, u_S)$ increasing in u_S. Agents with the greatest unobserved desire to select into treatment not only have higher benefits, but also have lower costs. These properties, while intuitive, need not hold in general—individuals with lower values of u_S (and thus a greater unobserved desire to take treatment) must necessarily have higher net surplus than those with higher values of u_S, but they need not have higher benefits and lower costs. It is possible, for example, that benefits and costs are so strongly positively correlated that those with the greatest unobserved desire to participate have either the smallest benefits and the lowest costs or the largest benefits and the highest costs. In Appendix A, we establish sufficient conditions for intuitive properties on $B^{MTE}(x, u_S)$ and $C^{MTE}(z, u_S)$ to hold, as well as testable implications of those conditions.

4 Identifying the Surplus and Benefit Functions Nonparametrically

Heckman and Vytlacil (1999, 2005) show that local instrumental variables (LIV) identify the marginal benefit of treatment:

$$\frac{\partial E(Y|X=x, P=p)}{\partial p} = B^{MTE}(x, p). \quad (4.1)$$

We can identify $E(Y|X = x, P = p)$ and its derivative for all $(x, p) \in \text{Supp}(X, P)$, where $\text{Supp}(X, P)$ denotes the support of $(X, P(X, Z))$.[7] We can thus identify $B^{MTE}(x, u_S)$ for all values of $(x, u_S) \in \text{Supp}(X, P)$. For a fixed x, we can identify $B^{MTE}(x, u_S)$ for $u_S \in \text{Supp}(P|X = x)$. The more variation in propensity scores P conditional on $X = x$, the larger the set of evaluation points u_S for which we identify $B^{MTE}(x, u_S)$. Variation in propensity scores conditional on X is driven by variation in Z, the cost shifters. Thus, if we observe regressors that produce large variations in costs, we will be able to identify $B^{MTE}(x, u_S)$ on a larger set.

We can identify $B^{ATE}(x)$ and $B^{TT}(x)$ by identifying $B^{MTE}(x, u_S)$ over the appropriate support and then integrating the latter with the appropriate weights, which are known given data on X and Z. By Equation (3.1), we identify $B^{ATE}(x)$ if $\text{Supp}(P|X = x) = [0, 1]$. For fixed $X = x$, this requires that there be enough variation in the cost shifters Z to drive the probabilities $P(x, Z)$ all the way to zero and to one. In other words, holding fixed the regressors that enter the outcome equation, we must observe cost shifters such that conditional on some values of those cost shifters, the cost to the agent is so low that the agent will select into treatment with probability arbitrarily close to one, and, conditional on other values of the cost shifters, the cost to the agent is so high that the agent will select into treatment with probability arbitrarily close to zero. Likewise, we identify $B^{TT}(x)$ if $\text{Supp}(P|X=x)=[0, p_x^{max}]$ where p_x^{max} is the supremum of $\text{Supp}(P|X = x)$. This support requirement in turn requires that, for fixed $X = x$, there be enough variation in the cost shifters Z to drive the selection probability arbitrarily close to zero.[8]

Using Equation (4.1) and the relationship for people on the margin of choice that $B^{MTE}(x, P(x, z)) = C^{MTE}(z, P(x, z))$, we have

$$\frac{\partial E(Y|X=x, P=p)}{\partial p}\Big|_{p=P(x,z)} = C^{MTE}(z, P(x, z)). \quad (4.2)$$

Using this relationship, we identify $C^{MTE}(z, u_S)$ for all values of $(z, u_S) \in \text{Supp}(Z, P)$. We thus identify the marginal cost parameter without direct information on the cost of treatment by using the structure of the Roy model and by identifying the marginal benefit of treatment for individuals at the margin of participation. For a fixed z, we identify $C^{MTE}(z, u_S)$ for $u_S \in$

[7]For any random vectors A and B, we will write the support of the distribution of A as $\text{Supp}(A)$, and the support of distribution of A conditional on $B = b$ as $\text{Supp}(A|B = b)$.

[8]Heckman and Vytlacil (2001a) show that one can identify $B^{ATE}(x)$ and $B^{TT}(x)$ under slightly weaker conditions than those required to follow this strategy of first identifying $B^{MTE}(x, u)$ over the appropriate support. In particular, they show that the necessary and sufficient condition for identification of $B^{ATE}(x)$ is that $\{0, 1\} \in \text{Supp}(P|X = x)$, and for $B^{TT}(x)$ that $\{0\} \in \text{Supp}(P|X = x)$.

Supp($P|Z = z$). The greater the variation in propensity scores conditional on $Z = z$, the larger the set of evaluation points for which we identify $C^{MTE}(z, u_S)$. Variation in propensity scores conditional on $Z = z$ is driven by variation in X, the regressors that affect the potential outcomes and thus that drive the benefit of treatment. If we observe X regressors that cause large variations in benefits, we will be able to identify $C^{MTE}(z, u_S)$ at a larger set of u_S evaluation points. In contrast, if there are no X regressors, then P only depends on Z and we can only identify $C^{MTE}(z, u_S)$ for $u_S = P(z)$.

From Equation (3.2), we can identify $C^{ATE}(z)$ if Supp($P|Z = z$) = [0, 1]. This requires, for fixed $Z = z$, for there to be enough variation in the outcome shifters X to drive the probabilities $P(X, Z)$ all the way to zero and to one. In other words, holding fixed the regressors that enter the cost equation, we must observe outcome shifters such that conditional on some values of those outcome shifters, the benefit to the agent is so high that the agent will select into treatment with probability arbitrarily close to one; conditional on other values of the outcome shifters, the benefit to the agent is so low that the agent will select into treatment with probability arbitrarily close zero. Likewise, we identify $C^{TT}(x)$ if $\text{Supp}(P|Z=z)=[0, p_z^{max}]$ where p_z^{max} is the supremum of Supp($P|Z = z$). This support requirement in turn requires that, for fixed $Z = z$, there is sufficient variation in the outcome shifters X to drive the probabilities arbitrarily close to zero.

Finally, consider identification of the surplus parameters. Using the fact that

$$S^{MTE}(x, z, u_S)=B^{MTE}(x, u_S)-C^{MTE}(z, u_S),$$

we can identify the marginal surplus parameter at (x, z, u_S) such that $(x, u_S) \in$ Supp(X, P) and $(z, u_S) \in$ Supp(Z, P). By Equation (3.3), we can integrate $S^{MTE}(x, z, u_S)$ using the appropriate weights (which are identified from the data on X and Z) to identify $S^{ATE}(x, z)$ and $S^{TT}(x, z)$ under the appropriate support conditions. For example, we identify $S^{ATE}(x, z)$ if Supp($P|X = x$) = [0, 1] and Supp($P|Z = z$) = [0, 1].

Thus, for identification of the treatment parameters we need sufficient variation in cost shifters conditional on the outcome shifters. For identification of the cost parameters, we need sufficient variation in the outcome shifters conditional on the cost shifters. For identification of the surplus parameters we need sufficient variation in both sets of regressors. We can thus identify the marginal cost, the average cost, and the cost of treatment without direct information on the cost. Consequently, we can also identify the corresponding surplus parameters as well. Our ability to do so is directly related to the extent of variation in observed regressors that shift the benefit of the treatment.

We summarize our discussion in the form of a theorem:

Theorem 1

Assume that Equations (2.1)–(2.4) and our Assumptions (A-1)–(A-4) hold.

1. *$B^{MTE}(x, u_S)$ is identified for $(x, u_S) \in Supp(X, P)$; $C^{MTE}(z, u_S)$ is identified for $(z, u_S) \in Supp(Z, P)$; and $S^{MTE}(x, z, u_S)$ is identified for (x, z, u_S) such that $(x, u_S) \in Supp(X, P)$ and $(z, u_S) \in Supp(Z, P)$.*

2. *$B^{ATE}(x)$ is identified if $Supp(P|X = x) = [0, 1]$; $C^{ATE}(z)$ is identified if $Supp(P|Z = z) = [0, 1]$; $S^{ATE}(x, z)$ is identified if $Supp(P|X = x) = [0, 1]$ and $Supp(P|Z = z) = [0, 1]$.*

3. *$B^{TT}(x)$ is identified if $Supp(P|X=x)=[0, p_x^{max}]$; $C^{TT}(z)$ is identified if $Supp(P|Z=z)=[0, p_z^{max}]$; $S^{TT}(x, z)$ is identified if $Supp(P|X=x)=[0, p_x^{max}]$ and $Supp(P|Z=z)=[0, p_z^{max}]$.*

Our results allow for unobserved heterogeneity in costs and benefits conditional on the observed regressors. If there is no unobserved (by the economist) heterogeneity in the costs of treatment, $U_C = 0$, then $C^{MTE}(z, u_S) = C^{TT}(z) = C^{ATE}(z)$ and thus we can identify the cost of treatment on the treated and average cost parameters without the additional support conditions. Likewise, if we impose that there is no unobserved heterogeneity in the benefits of treatment, $U_1 - U_0 = 0$, we have $B^{MTE}(z, u_S) = B^{TT}(z) = B^{ATE}(z)$ and can thus identify all of the benefit parameters without additional support conditions.

We establish identification of the marginal effect parameters within the conditional support of P. However, exploiting additive separability, we are able to extend the margin of identification to the unconditional support of P by a chaining argument. We illustrate the reasoning behind this for the $B^{MTE}(x, u_S)$, but the analogous result applies to the marginal cost and surplus functions as well.

Recall that $B^{MTE}(x, u_S) = \mu_1(x) - \mu_0(x) + E(U_1 - U_0|U_S = u_S)$ is identified for all $(x, u_S) \in$ Supp(X, P). How $B^{MTE}(x, u_S)$ varies with x does not depend on the point of evaluation of u_S, and how $B^{MTE}(x, u_S)$ varies with u_S does not depend on the point of evaluation of x. This insight is helpful in securing identification of $B^{MTE}(x, u_S)$ for other (x, u_S) pairs.

For example, consider two potential values of X, x_0 and x_1, and suppose that there exists some p^* such that $p^* \in$ Supp$(P|X = x_0) \cap$ Supp$(P|X = x_1)$ so that $B^{MTE}(x_0, p^*)$ and $B^{MTE}(x_1, p^*)$ are both identified by Theorem 1. $B^{MTE}(x, u_S)$ is additively separable in x and u_S. As a consequence of additive separability, it follows directly that

$$B^{MTE}(x_0, u_S) - B^{MTE}(x_0, p^*) = B^{MTE}(x_1, u_S) - B^{MTE}(x_1, p^*). \quad (4.3)$$

If $u_S \in$ Supp$(P|X = x_1)$, we identify $B^{MTE}(x_1, u_S)$ by Theorem 1. We can solve Equation (4.3) to identify $B^{MTE}(x_0, u_S)$ even if $u_S \notin$ Supp$(P|X = x_0)$. Alternatively, if $u_S \in$ Supp$(P|X = x_0)$, we identify $B^{MTE}(x_0, u_S)$ by Theorem 1 and can now solve Equation (4.3) to identify $B^{MTE}(x_1, u_S)$ even if $u_S \notin$ Supp$(P|X = x_1)$. Thus, if there exists some p^* such that $p^* \in$ Supp$(P|X = x_0) \cap$ Supp$(P|X = x_1)$, we can chain together identification of $B^{MTE}(x_0, u_S)$ for $u_S \in$ Supp$(P|X = x_0)$ and identification of $B^{MTE}(x_1, u_S)$ for $u_S \in$ Supp$(P|X = x_1)$ to obtain identification of $B^{MTE}(x_0, u_S)$ and $B^{MTE}(x_1, u_S)$ for all $u_S \in$ Supp$(P|X = x_0) \cup$ Supp$(P|X = x_1)$. One can iterate to further increase the range of values for which $B^{MTE}(x, u_S)$ is identified. Under an additional rank condition, we can use this strategy to identify $B^{MTE}(x, u_S)$ for all $(x, u_S) \in$ Supp$(X) \times$ Supp(P). In particular, we consider the following assumption:

(A-5) *X and $P(X, Z)$ are measurably separated; i.e., any function of X that almost surely equals a function of $P(X, Z)$ must be almost surely equal to a constant.*

Measurable separability between X and P is a rank condition. A necessary condition for measurable separability between X and $P(X, Z)$ is for $P(X, Z)$ to be nondegenerate conditional on X, as implied by $P(X, Z) = F_V(\mu_S(X, Z))$ along with Assumptions (A-2) and (A-3). In Theorem 5 in Appendix A, we build on Theorem 2 of Florens et al. (2008) to provide sufficient conditions on our model for measurable separability between X and $P(X, Z)$. As shown by that theorem, strengthened versions of Assumptions (A-2) and (A-3), along with an additional support condition, are sufficient for measurable separability between X and $P(X, Z)$.

Using Assumption (A-5), we obtain the following identification result:

Theorem 2

Assume that Equations (2.1)–(2.4) and our Assumptions (A-1)–(A-5) hold. Then, for $x \in \text{Supp}(X)$ and $z \in \text{Supp}(Z)$,

1. $B^{MTE}(x, u_S)$, $C^{MTE}(z, u_S)$ *and* $S^{MTE}(x, z, u_S)$ *are identified for* $u_S \in Supp(P)$.
2. $B^{ATE}(x)$, $C^{ATE}(z)$ *and* $S^{ATE}(x, z)$ *are identified if* $Supp(P) = [0, 1]$, *and*
3. $B^{TT}(x)$, $C^{TT}(z)$ *and* $S^{TT}(x, z)$ *are identified if* $Supp(P) = [0, p^{max}]$.

The proof of Theorem 2 is in Appendix B. The theorem shows that, under our maintained assumptions and condition (A-5), identification of the treatment parameters depends on the marginal support of P, not on the support of P conditional on X or Z.

5 Extension to the Case of Limited Information by the Agent

Thus far, our analysis has assumed choice Equation (2.4), i.e., that $D = \mathbf{1}[S \geq 0]$ where $S = (Y_1 - Y_0) - C$. This implicitly assumes that agents have perfect foresight about their net benefit. In this section, we extend the choice model of Equation (2.4) to allow for limited information on the part of the agents, while maintaining the model for latent outcomes (Y_0, Y_1) and cost C of Equations (2.2) and (2.3). We assume that agents form valid expectations about their outcomes and costs given the information that they have at the time of their treatment choice and that they select into treatment if the expected surplus is positive. We allow agents to know only some elements of (X, Z), and possibly to have incomplete knowledge of (U_0, U_1, U_C) and hence their own idiosyncratic benefit and cost of treatment. We now show that the preceding analysis goes through with minor modifications, though it is now important to distinguish conditioning sets: what is known to the agent at the time of treatment choice (which might include some information not known to the econometrician), what is known to the econometrician (which might include some information not known to the agent at the time of treatment choice), and what is realized *ex post*. The essential change in our procedure in the case of incomplete information is that the marginal benefit of treatment identified by LIV must be projected onto the agent's information set when selecting treatment to form the expected marginal benefit of treatment conditional on the information available to the agent. This coarsened version of B^{MTE} is used to identify the marginal cost parameter. In addition, only components of X that are known to the agent at the time of treatment choice can aid in identification of the cost parameters. The exclusion restrictions for identification of the cost parameter are variables in X that are not in Z and that are known to the agent at the time of choosing treatment.

Let (X_I, Z) denote components of (X, Z) that are observed by the agent when choosing whether to select into treatment.[9] Suppose that the agent's information set is (X_I, Z, U_I).[10] U_I is the private information of the agent relevant to his or her own benefits and cost of treatment, and is not observed by the econometrician.

We revise assumption (A-1) in the following way:

(A-1′) (U_0, U_1, U_C, U_I) *is independent of* (X, Z), *and* X *is independent of* Z *conditional on* X_I.

Assumption (A-1′) imposes the requirement that the private information of the agent is independent of the observed regressors. Note that, under this independence assumption, $(U_0, U_1, U_C, U_I) \perp\!\!\!\perp (X_I, Z)$, and

$$E(V|X, Z, U_I) = E(V|X_I, Z, U_I) = E(V|U_I),$$

using the definition $V = U_C - (U_1 - U_0)$.

Assumption (A-1′) implies that $(X, Z) \perp\!\!\!\perp U_I \mid (X_I, Z)$, so that U_I does not help the agent predict elements of (X, Z) that are not contained in (X_I, Z). Thus, we allow the agents to have private information about their own idiosyncratic benefits $(U_1 - U_0)$ and costs U_C, though we impose the restriction that the only information known by the agent that is useful for predicting X is (X_I, Z). Furthermore, Assumption (A-1′) requires that, conditional on the components of X known to the agent at the time of selecting into treatment, Z does not help to predict those elements of X not known at the time of treatment selection. This restriction is only imposed for notational convenience and can be easily relaxed.

We restate Assumption (A-3) as:

(A-3′) *The distribution of* $\tilde{V} = E(V|U_I)$ *is absolutely continuous with respect to Lebesgue measure, and the cumulative distribution function of* $\tilde{V}$ *is strictly increasing.*

An implication of (A-3′) is that $E(V|U_I)$ is a nondegenerate random variable, and thus that agents have some nontrivial information about their own idiosyncratic cost or benefit from treatment when deciding whether to select into treatment. We maintain Assumptions (A-2) and (A-4) as before.

---

[9]We assume that agents know all components of Z, while we allow agents to be ignorant of some components of X. This assumption simplifies our notation and conforms to our empirical analysis of Section 6. The analysis can be extended (at the cost of somewhat more cumbersome notation) to allow agents to know only a subvector of Z as well as only a subvector of X at the time of selection into treatment.

[10]In other words, the information set of the agent equals $\sigma(X, Z, U_I)$, the sigma-algebra generated by (X, Z, U_I).

Define $\mu_j^I(X_I)=E(Y_j|X_I)$ for j = 0, 1, and $\mu_C^I(Z)=E(C|Z)$, and note that given our independence assumptions and the law of iterated expectations,

$\mu_j^I(X_I)=E(\mu_j(X)|X_I), \mu_C^I(Z)=E(\mu_C(Z)|Z)$. Define $\mu_S^I(X_I,Z)=E(S|X_I,Z)$. Under our assumptions,

$$E(S|X_I,Z,U_I)=\mu_S^I(X_I,Z)-\tilde{V}=\mu_1^I(X_I)-\mu_0^I(X_I)-\mu_C^I(Z)-\tilde{V}.$$

The decision rule becomes

$$D=1 \quad \text{if} \quad E(S|X_I,Z,U_I) \geq 0; \quad D=0 \quad \text{otherwise}, \tag{5.1}$$

where $E(S|X_I, Z, U_I)$ is the expected surplus from treatment, with the expectation conditional on the agents information set. We thus have

$$D=\mathbf{1}\,[\mu_S^I(X_I,Z)-\tilde{V} \geq 0],$$

where our independence assumptions imply $\tilde{V} \perp\!\!\!\perp (X_I, Z)$, and thus the selection model is of the same form as that used by Heckman and Vytlacil (1999), which allows us to use LIV to identify $B^{MTE}(x, u_S)$. Redefining $U_S = F_{\tilde{V}}(\tilde{V})$ and

$P(X_I,Z)=\Pr[D=1|X_I,Z]=F_{\tilde{V}}(\mu_S^I(X_I,Z))$, we have

$$D=\mathbf{1}\,[P(X_I,Z)-U_S \geq 0],$$

with U_S distributed unit uniform and independent of (X, Z) and thus independent of (X_I, Z).

Define $B_I^{MTE}(x_I,u_S) \equiv E(Y_1-Y_0|X_I=x_I,U_S=u_S), C_I^{MTE}(z,u_S) \equiv E(C|Z=z,U_S=u_S)$, and $S_I^{MTE}(x_I,z,u_S) \equiv B_I^{MTE}(x_I,u_S)-C_I^{MTE}(z,u_S)$, the marginal benefit, cost, and net surplus of treatment conditional on the agent's information set, where again by the law of iterated expectations and our independence assumptions

$$\begin{array}{rcccc} B_I^{MTE}(x_I,u_S) & = & E(B^{MTE}(X,u_S)|X_I=x_I,U_S=u_S) & = & E(B^{MTE}(X,u_S)|X_I=x_I) \\ C_I^{MTE}(z,u_S) & = & E(C^{MTE}(Z,u_S)|Z=z,U_S=u_S) & = & E(C^{MTE}(Z,u_S)|Z=z). \end{array}$$

Evaluating $S_I^{MTE}(x_I,z,u_S)$ at $u_S = P(x_I, z)$, we obtain

$$\begin{aligned} S_I^{MTE}(x_I, z, P(x_I, z)) &= \mu_S^I(x_I, z) - E(V|U_S = P(x_I, z)) \\ &= \mu_S^I(x_I, z) - E(V|\tilde{V} = \mu_S^I(x_I, z)) \\ &= \mu_S^I(x_I, z) - E(V|E(V|U_I) = \mu_S^I(x_I, z)) \\ &= \mu_S^I(x_I, z) - E(E(V|U_I)|E(V|U_I) = \mu_S^I(x_I, z)) \\ &= \mu_S^I(x_I, z) - \mu_S^I(x_I, z) \\ &= 0, \end{aligned}$$

where the second equality is obtained by plugging in the definition of U_S, the third equality is obtained by plugging in the definition of $\tilde{V}$, and the fourth equality is obtained using the law of iterated expectations and the fact that $E(V|U_I)$ is degenerate given U_I. Since $S_I^{MTE}(x_I, z, u_S) = B_I^{MTE}(x_I, u_S) - C_I^{MTE}(z, u_S)$, we have

$$B_I^{MTE}(x_I, u_S) = C_I^{MTE}(z, u_S) \text{ for } u_S \text{ such that } u_S = P(x_I, z).$$

Thus, identification of $B_I^{MTE}(x_I, P(x_I, z))$ provides identification of $C_I^{MTE}(z, P(x_I, z))$.

Since our model is a special case of Heckman and Vytlacil (1999), we can follow them in using LIV to identify $B^{MTE}(x, u_S)$ for (x, u_S) in the support of $(X, P(X_I, Z))$. It is important to note that LIV does not identify the $B^{MTE}(x, u_S)$ that is relevant to the agent's decision problem. LIV identifies $B^{MTE}(x, u_S) = E(Y_1 - Y_0|X = x, U_S = u_S)$, not $B_I^{MTE}(x_I, u_S) = E(Y_1 - Y_0|X_t = x_I, U_S = u_S)$. However, we can project the $B^{MTE}(x, u_S)$ identified by LIV on the information known to the agent at the time of selecting into treatment and coarsen the set used to define and identify $B^{MTE}(x, u_S)$, to identify the $B_I^{MTE}(x_I, u_S)$ relevant to the agent's decision problem. It is the latter that is relevant for identifying the cost functions. By the law of iterated expectations, we obtain

$$B_I^{MTE}(x_I, u_S) = E(B^{MTE}(X, u_S)|X_I = x_I) = \int B^{MTE}(x, u_S)dF_x(x|X_I = x_I), \quad (5.2)$$

where $F_X(\cdot|X_I = x_I)$ is the cumulative distribution function of X conditional on $X_I = x_I$. We directly identify $F_X(\cdot|X_I = x_I)$, and thus, for given u_S, obtain identification of $B^{MTE}(x, u_S)$ for all $x \in \text{Supp}(X|X_I = x_I)$ implies identification of $B_I^{MTE}(x_I, u_S)$. Since, for a given x, we identify $B^{MTE}(x, u_S)$ if $u_S \in \text{Supp}(P(X_I, Z)|X = x)$, we thus identify $B_I^{MTE}(x_I, u_S)$ if

$$u_S \in \bigcap_{x \in \text{Supp}(X|X_I = x_I)} \text{Supp}(P(X_I, Z)|X = x).$$

In other words, to identify *ex ante* $B_I^{MTE}(x_I, u_S)$, we need to identify *ex post* $B^{MTE}(x, u_S)$ for every value x that X can take given $X_I = x_I$, and thus we need for u_S to be an element of $\text{Supp}(P(X_I, Z)|X = x)$ for each value x that X can take given $X_I = x_I$. However, using the fact that X_I is a subvector of X and independence assumption (A-1′), it follows that Supp($P(X_I$,

$Z)|X)$ = Supp($P(X_I, Z)|X_I$), and thus using Equation (5.2) we identify $B_I^{MTE}(x_I, u_S)$ for (x_I, u_S) in the support of (X_I, $P(X_I, Z)$). Using the fact that $B_I^{MTE}(x_I, P(x_I, z)) = C_I^{MTE}(z, P(x_I, z))$, we identify $C_I^{MTE}(z, u_S)$ for (z, u_S) in the support of (Z, $P(X_I, Z)$). We have thus identified the marginal cost parameter, and can integrate it to obtain other cost parameters. We can also combine it with the benefit parameters to identify net surplus parameters as before. The only elements of X that are useful for identifying the cost parameters are those elements that are in X, but not in Z, and which are known to the agent at the time of selection into treatment (i.e., are contained in X_I).

6 Estimating the Cost and Surplus from Educational Choices

We apply our methodology to an analysis of educational choice and estimate the marginal benefit, cost, and surplus from a college education. Carneiro et al. (2011) provide estimates of the marginal benefit of attending college. We extend their work by adding results for the subjective cost and surplus. Björklund and Moffitt (1987) provide fully parametric estimates of cost and surplus in the context of a manpower training program in Sweden. Application of their approach offers a useful benchmark to gauge our more flexible estimation strategy. Our nonparametric identification analysis follows Marschak (1953) who noted that for many policy analyses only combinations of structural parameters are required. We embrace *Marschak's Maxim* (Heckman, 2010) and implement an estimation strategy with minimal assumptions and transparent sources of identification for the marginal effects of treatment.

We analyze a sample of 1,747 white males from the National Longitudinal Survey of Youth of 1979 (NLSY79).[11] The outcome variable is the log of the mean non-missing values of the hourly wage between 1989 and 1993, which we interpret as an estimate of the log hourly wage in 1991, and an approximation to the long-run wage. Schooling is measured in 1991 when individuals are between 28 and 34 years of age. We separate individuals into two groups: persons with no college (D = 0) and persons with at least some college (D = 1). We present annualized returns to education, obtained by dividing all our estimates by four which is the average difference in years of schooling between those with $D = 1$ and those with $D = 0$.

To identify the $C_I^{MTE}(z, u_S)$, we require variables that do not affect the cost of attending college, but that change future wages and are known to the agent at college entry (benefit shifters). We measure long-run labor market conditions by permanent local wages and compute average earnings between 1973 and 2000 for each location of residence at 17 as a proxy. Since we will also condition on current labor market conditions at the time of potential enrollment, these regressors should only affect the schooling decision through their effect on agent's expected future wages and thus the expected benefit of treatment. We assume that the main benefits to a higher education are through earnings. Any other subjective benefits, such as allowing access to jobs with preferred amenities, are implicitly included (as a negative contribution) in costs. The validity of our exclusion restriction would

[11]See Bureau of Labor Statistics (2005) for a detailed description of the NLSY79 and Appendix C for details on the construction of the variables.

be threatened if our measure of permanent local wages affects the subjective benefit of education.

We identify $B^{MTE}(x, u_S)$ and $B_I^{MTE}(x_I, u_S)$ using variables that do not affect future wages, but only the cost of attending college (cost shifters). We use current fluctuations in local labor market conditions such as local wages at the time of the educational decision, which shift the opportunity cost of schooling. They should not help to predict the agent's expected future wages as we also control for permanent local labor market conditions. Effectively, we use only the innovations in local wages as cost shifters. We also include tuition cost, a dummy variable indicating urban residence at age 14, and distance to college as shifters that affect the direct cost of attending college.

Table 1 presents the covariates used in our empirical analysis. We highlight the two different types of exclusion restrictions. Variables that affect benefits as well as costs of treatment (common elements) include the Armed Forces Qualifying Test (AFQT) scores, mother's education, number of siblings, and cohort dummies. In what follows, we keep this set of observables in the background to ease notation. X and Z continue to denote the benefit and cost shifters respectively. X_I is the subvector of X which is known to the agent at the college entry decision. We include two variables in X not included in X_I : years of experience and wages in the county of residence. The excluded variables are measured approximately 12 years after the agent's college entry decision and thus not in the individual's information set at the time of the treatment decision. We follow the analysis of Section 5 and allow agents to have imperfect foresight about the realizations of these variables. They form expectations about their future wages, but do not have perfect information. In line with our exposition, we assume that Z does not help to predict the *ex post* realization of X conditional on X_I and denote the agent's information about their idiosyncratic cost and benefit from treatment as $\tilde{V} = E(V \mid U_I)$.

We specify a linear version of the generalized Roy model. Define potential outcomes:

$$Y_1 = X\beta_1 + U_1 \quad \text{and} \quad Y_0 = X\beta_0 + U_0.$$

The choice equation is:

$$D = \mathbf{1}\left[X_I(\alpha_1 - \alpha_0) - Z\gamma > \tilde{V}\right],$$

where we assume that agents form valid expectations about their own outcomes so that $E(X(\beta_1 - \beta_0) \mid X_I) = X_I(\alpha_1 - \alpha_0)$ holds. Note that X_I does not only affect the returns to education directly, but also helps to predict the *ex post* realization of those elements of X not contained in X_I.

We first implement the traditional structural approach and explicitly estimate all components of the generalized Roy model and combine them to form the marginal effect parameters (Björklund and Moffitt, 1987). We impose normality for the unobservables and fit the model by maximum-likelihood. As the participation decision is based on the net surplus and X does not affect the cost of treatment, this implies a cross-equation restriction between the coefficients on X in the outcome equations and X_I in the choice equation. We account for agents' imperfect foresight and set $(\alpha_1-\alpha_0)=(\overline{X}_I'\overline{X}_I)^{-1}\overline{X}_I'\overline{X}(\beta_1-\beta_0)$, where $(\bar{X}, \bar{X}_I)$ denote the matrices with the outcome shifters of the whole sample. We estimate the whole model in one step. In a standard Probit model, the coefficients can only be identified up to a factor of proportionality. However, as the wage gain $(\alpha_1 - \alpha_0)X_I$ appears with a coefficient of one in the choice equation, we do not need to normalize the variance of $\tilde{V}$ and estimate it instead. We can then construct the marginal effects of treatment based on the results:

$$
\begin{aligned}
B^{MTE}(x,u_S) &= x(\beta_1-\beta_0) + \left(\frac{\sigma_{U_1-U_0,\tilde{V}}}{\sigma^2_{\tilde{V}}}\right)\Phi^{-1}_{\sigma_{\tilde{V}}}(u_S)\\
B_I^{MTE}(x_I,u_S) &= x_I(\alpha_1-\alpha_0) + \left(\frac{\sigma_{U_1-U_0,\tilde{V}}}{\sigma^2_{\tilde{V}}}\right)\Phi^{-1}_{\sigma_{\tilde{V}}}(u_S)\\
C_I^{MTE}(z,u_S) &= z\gamma + \left(\frac{\sigma_{U_C,\tilde{V}}}{\sigma^2_{\tilde{V}}}\right)\Phi^{-1}_{\sigma_{\tilde{V}}}(u_S)\\
S_I^{MTE}(x_I,z,u_S) &= x_I(\alpha_1-\alpha_0)-z\gamma - \Phi^{-1}_{\sigma_{\tilde{V}}}(u_S),
\end{aligned}
$$

where $\sigma_{U_1-U_0,\tilde{V}}$ and $\sigma_{U_C\tilde{V}}$ denote the covariance between $(U_1 - U_0, \tilde{V})$ and $(U_C, \tilde{V})$ respectively. $\Phi^{-1}_{\sigma_{\tilde{V}}}$ indicates the inverse of a normal cumulative distribution function with standard deviation $\sigma_{\tilde{V}}$.

The sign of the slope of the marginal effect parameters is determined by $\sigma_{U_C\tilde{V}}$ and $\sigma_{U_1-U_0\tilde{V}}$ as $\sigma^2_{\tilde{V}}>0$. We present our results for these parameters in Table 2. The estimate for $\sigma_{U_1-U_0\tilde{V}}$ is negative and thus the marginal benefits of treatment decrease when moving along the margins of $\tilde{V}$. The opposite is true for $\sigma_{U_C\tilde{V}}$ and so the marginal cost increases in u_S. However, only $\sigma_{U_1-U_0\tilde{V}}$ is significantly different from zero at the 10% level.

Figure 2 presents our fully parametric results for the *ex post* marginal benefit and *ex ante* cost and surplus parameters. We plot them as a function of u_S and evaluate them at the sample mean of (X_I, Z). As agents are assumed to form valid expectations about their future benefits, the *ex ante* and *ex post* marginal benefits are identical. Individuals with a high unobserved desire for treatment (low u_S) have the highest benefit, strictly decreasing from +16% to −4%. The estimated surplus is positive for low values of u_S and decreases when moving along the margins of $\tilde{V}$. The opposite holds for the marginal cost, which is always positive and slightly increasing. The cost is lowest for individuals with low values of u_S and ranges from +3% to +10%. In summary, the benefit is highest and cost lowest for those most likely to pursue a higher education. However, the estimates are not precisely determined. The marginal benefit of treatment is significantly different from zero for roughly half of the individuals. Along all margins of $\tilde{V}$, the marginal cost of a college education does not significantly differ from zero. *By construction*, the marginal surplus is strictly positive for all those individuals who participate in the treatment and negative for those that do not. Conditional on the observables set to their sample mean, individuals are indifferent towards treatment when $u_S = 0.51$.

Figure 2 presents the marginal effect parameters over the full unit interval from the structural model. The distributional assumptions on $(U_1, U_0, \tilde{V})$ expand the margins for which we can identify the marginal effects of treatment. As we assume full independence between all observables and unobservables, we identify the marginal effects of treatment over the unconditional common support of $P(X_I, Z)$. In our sample, this support ranges between 0.03 and 0.98. Adding joint normality, we can extrapolate even further and cover the full unit interval.

However, our formal analysis demonstrates that in a fully nonparametric setting we are only able to identify the $B_I^{MTE}(x_I, u_S)$ over the support of $P(X_I, Z)$ conditional on $X_I = x_I$ and the $C_I^{MTE}(z, u_S)$ over the support of $P(X_I, Z)$ conditional on $Z = z$. We identify the $S_I^{MTE}(x_I, z, u_S)$ over the intersection of the two supports. In Figure 3 we plot the conditional densities of $P(X_I, Z)$ in our data. As X_I and Z are both multidimensional, we condition on the decile of the relevant index, i.e. on $X_I(\alpha_1 - \alpha_0)$ for the $B_I^{MTE}(x_I, u_S)$ and $Z\gamma$ for the $C_I^{MTE}(z, u_S)$.[12] The support is very limited and thus the results of a fully parametric implementation rely heavily on extrapolation based on the distributional assumptions.

We now develop a semiparametric estimation strategy that relies on fewer assumptions and provides more transparent sources of identification. We apply *Marschak's Maxim*, estimating only those combinations of structural parameters needed for the marginal effect parameters. To fix ideas, consider the estimation of the $B^{MTE}(x, u_S)$, where the conditional expectation of $(U_1 - U_0)$ along the margins of $\tilde{V}$ is a key element. In the fully parametric normal-theory approach, it is directly constructed from estimates of $(\sigma_{U_1, \tilde{V}}, \sigma_{U_0, \tilde{V}})$ and $\sigma_{\tilde{V}}^2$:

$$E(U_1 - U_0 | U_S = u_S) = \left(\frac{\sigma_{U_1 - U_0, \tilde{V}}}{\sigma_{\tilde{V}}^2}\right) \Phi_{\sigma_{\tilde{V}}}^{-1}(u_S).$$

Instead, in what follows, we directly obtain $E(U_1 - U_0 \mid U_S = u_S)$ without having to estimate all structural components. We will also carefully recognize the relevant conditional support of P for each parameter and thus present a data-sensitive structural analysis (Heckman, 2010).

We determine the support of P by building on an estimator of the joint support of the distribution of (X, Z):

$$\hat{S}_{X,Z} = \{(x, z): \|(X_i, Z_i) - (x, z)\| \leq \varepsilon \quad \text{for some } i\},$$

[12]We trace out the remaining variation in $P(X_I, Z)$ by applying a two-dimensional kernel density estimation with a bivariate normal kernel.

where $\|\cdot\|$ corresponds to the Euclidean norm and i denotes a generic observation in our data.[13] Then, letting $x_I(x)$ indicate the appropriate subvector of x, our resulting estimator for the support of (X_I, Z) is:

$$\hat{S}_{X_I,Z}=\{(x_I^*,z^*):\exists(x,z)\in\hat{S}_{X,Z}\quad \text{such that}\quad (x_I(x),z)=(x_I^*,z^*)\}.$$

We can use these estimates to construct our desired support for the marginal cost and benefit parameters:

$$\begin{aligned}\hat{S}_{X,P}&=\{(x^*,p^*):\exists(x,z)\in\hat{S}_{X,Z}\quad \text{such that}\quad (x,P(x_I(x),z))=(x^*,p^*)\}\\ \hat{S}_{X_I,P}&=\{(x_I^*,p^*):\exists(x_I,z)\in\hat{S}_{X_I,Z}\quad \text{such that}\quad (x_I,P(x_I,z))=(x_I^*,p^*)\}\\ \hat{S}_{Z,P}&=\{(z^*,p^*):\exists(x_I,z)\in\hat{S}_{X_I,Z}\quad \text{such that}\quad (z,P(x_I,z))=(z^*,p^*)\}.\end{aligned}$$

Note, that the variation in p for a given x and $x_I(x)$ is the same in $\hat{S}_{X,P}$ and $\hat{S}_{X_I,P}$. Thus we can identify the $B^{MTE}(x, u_S)$ and $B_I^{MTE}(x_I(x),u_S)$ over the same margins. Finally, for the marginal surplus parameter, we collect in $\hat{S}_{X_I,Z,P}$ all (x_I, z, p) where the relevant subsets in $\hat{S}_{X_I,P}$ and $\hat{S}_{Z,P}$ overlap in p. We only report estimates for the margins within these sets and thus acknowledge the limitations of the data.

We estimate the $B^{MTE}(x, u_S)$ using the method of local instrumental variables (LIV) proposed in Heckman and Vytlacil (1999, 2001b, 2005). They show that under our conditions the $B^{MTE}(x, u_S)$ is identified by differentiating the conditional expectation of observed outcomes:

$$\frac{\partial E(Y|X=x,P=p)}{\partial p}\Big|_{p=u_S}=B^{MTE}(x,u_S). \quad (6.1)$$

Applied to sample data, this is the LIV estimator of Heckman and Vytlacil (1999).[14] As noted in Carneiro et al. (2011), it is empirically very difficult to apply the LIV estimator while conditioning on all variables in the outcome equations. Thus we proceed by invoking the stronger assumption that in addition to the variables in X, all elements common to outcome and choice equations are independent of $(U_1, U_0, \tilde{V})$ as well. Because our generalized Roy model is also linear, the conditional expectation of Y simplifies to:

$$\begin{aligned}E(Y|X=x,P=p)&=E(DY_1+(1-D)Y_0|X=x,P=p)\\ &=x\beta_0+px(\beta_1-\beta_0)+K(p),\end{aligned} \quad (6.2)$$

where $K(p) = E(U_1 - U_0 \mid D = 1, P = p)$ can be estimated nonparametrically. We determine the parameters of Equation (6.1) by a partially linear regression of Y on X and P. We proceed in two steps. The first step is the construction of P, and the second step is the estimation of β_1 and β_0 using the estimated P. We carry out the first step using a Probit

[13] In practice, we set ε such that at most 5% of the sample are within the support for a given pair of (X_i, Z_i).

[14] See the Web Appendix of Heckman et al. (2006) for a detailed description of the implementation of the LIV estimator.

regression of D on (X_I, Z). In the second step we use Robinson (1988)'s method for estimating partially linear models as extended in Heckman et al. (1997a).[15] Next, consider the estimation of $K(P)$. Equation (6.2) implies that $E(\tilde{Y}) = K(p)$, where $\tilde{Y} = Y - x\beta_0 - px(\beta_1 - \beta_0)$ is the residualized observed outcome. We thus use a local quadratic regression of $\tilde{Y}$ on P to estimate $K(P)$ and its partial derivative with respect to P.[16] We construct the *ex post* marginal benefit of treatment $B^{MTE}(x, u_S)$ based on these estimates:

$$B^{MTE}(x,p)=x(\beta_1-\beta_0)+\frac{\partial K(p)}{\partial p} \quad \forall \quad (x,p) \in \hat{S}_{X,P}.$$

For the *ex ante* marginal benefit of treatment, we account for the agents' imperfect foresight about the future realization of components of X. As agents form valid expectations, we calculate $(\alpha_1-\alpha_0)=(\overline{X}_I'\,\overline{X}_I)^{-1}\overline{X}_I'\,\overline{X}(\beta_1-\beta_0)$[17] and then construct the $B_I^{MTE}(x_I, u_S)$ as follows:

$$B_I^{MTE}(x_I,p)=x_I(\alpha_1-\alpha_0)+\frac{\partial K(p)}{\partial p} \quad \forall \quad (x_I,p) \in \hat{S}_{X_I,P}.$$

We can identify the $C_I^{MTE}(z, u_S)$ using the equality of the marginal cost and benefit parameter at the margin of indifference:

$$C_I^{MTE}(z,p)=B_I^{MTE}(x_I,p) \quad \forall \quad (z,p) \in \hat{S}_{Z,P}. \tag{6.3}$$

This step directly mirrors Equation (4.2) from our nonparametric identification analysis. We obtain an estimate for the marginal cost of treatment using only information on the marginal benefits. We do not exploit any additional distributional assumptions such as joint normality of the unobservables.

We finally determine the $S_I^{MTE}(x_I, z, u_S)$ by taking the difference between benefits and costs:

$$S_I^{MTE}(x_I,z,p)=B_I^{MTE}(x_I,p)-C_I^{MTE}(z,p) \quad \forall \quad (x_I,z,p) \in \hat{S}_{X_I,Z,P}. \tag{6.4}$$

Figure 4 presents our semiparametric results for the *ex ante* benefit, cost and surplus parameters as well as the *ex post* benefit. We calculate the marginal effects at the mean values in the sample $(\bar{x}, \bar{z})$ and at two additional points of evaluation (x^A, z^A) and (x^B, z^B).

[15]We run kernel regressions of each of the regressors on P using a bandwidth of $h = 0.05$. We compute the residuals of each of these regressions and then run a linear regression of Y on these residuals.

[16]We choose the bandwidth that minimizes the residual square criterion proposed in Fan and Gijbles (1996), which gives us a bandwidth of $h = 0.3$.

[17]The economics of the model imply a restriction on the coefficients $(\alpha_1 - \alpha_0)$ in the choice equation, which depend on the estimated values of $(\beta_1 - \beta_0)$. However, we only learn about the values of $(\beta_1 - \beta_0)$ using an initial estimate of P. We insure internal consistency of our estimation routine by iterating between the estimation of the $B^{MTE}(x, u_S)$ and P with restricted $(\alpha_1 - \alpha_0)$ until convergence.

We plot them as a function of u_S within the relevant conditional support and compute the 90% confidence bands using the bootstrap.[18]

Our estimates show that individuals with a high unobserved desire for treatment (low u_S) have high benefits as well as high costs from participation. When moving up the margins of u_S the benefits fall more quickly than the costs as the surplus decreases. The $B_I^{MTE}(x_I, x_S)$ ranges from +37% within the support of x^A to as low as −12% within the support of x^B. The $C_I^{MTE}(z, u_S)$ varies between +32% and −6% overall, but within each margin of support the variation is limited to about 4% in absolute value. We can calculate the $S_I^{MTE}(x_I, z, u_S)$ which ranges from +5% to −5% as the difference between *ex ante* benefits and costs within the overlap of the support. Note that the estimates for the marginal benefits at x^B are all negative. However, costs are as well and so the surplus is still positive at the lower end of the conditional support. After conditioning on observables, it is unobservable heterogeneity in benefits and not costs that is driving the college entry decision. However, all estimates are rather imprecise, precision is highest at the mean values in the sample.

The conditional support is limited as shown in Figure 3. The location and range of the support depends on the point of evaluation. In general, we can identify $B^{MTE}(x, u_S)$ and $B_I^{MTE}(x_I, x_S)$ over longer stretches of u_S than the $C_I^{MTE}(z, u_S)$ function. In fact, for all x_I, z evaluation points considered, the values of u_S for which we identify $C_I^{MTE}(z, u_S)$ is a subset of the values of u_S for which we identify $B_I^{MTE}(x_I, x_S)$. Hence, for the x_I, z evaluation points considered, we can identify $S_I^{MTE}(x_I, z, u_S)$ only over the set of u_S values corresponding to the smaller set of u_S values for which we identify $C_I^{MTE}(z, u_S)$. The conditional variation in P is largest at $\bar{x}_I$ where we can identify the longest stretch for the $B_I^{MTE}(\overline{x}_I, x_S)$ with $u_S \in (0.42, 0.61)$, while it is smallest for $C_I^{MTE}(z^B, x_S)$ with $u_S \in (0.81, 0.89)$. Note that we identify all marginal effect parameters around the margin of indifference at $S_I^{MTE}(x_I, z, u_S)=0$.

We can also assess the magnitude of the expectation errors due to the agents' imperfect foresight about parts of their future benefits. Given our prediction model, the *ex post* and *ex ante* benefits coincide for the average individual $(\bar{x}, \bar{z})$. However, a comparison between realized and predicted benefits reveals that at x^A, *ex post* benefits are overestimated by about 9%, while at x^B the prediction is only off by 3%.

We can compare the results for the marginal effects of treatment between the two estimation approaches at $(\bar{x}, \bar{z})$ within the conditional support. The semiparametric approach indicates larger heterogeneity in benefits and costs due to the steeper slope of the marginal effect parameters. In both cases, benefits decrease considerably when moving along the margins of $\tilde{V}$ while variation in costs is limited. Thus, it is heterogeneity in benefits that drives the college attendance decision. This is in line with the results by Björklund and Moffitt (1987),

[18]We use 2,000 bootstrap replications. In each iteration of the bootstrap we re-estimate P so all standard errors account for the fact that P itself is an estimated object.

who also find that heterogeneity in rewards is more important than heterogeneity in costs for the participation decision in their context of a manpower training program in Sweden.

7 Summary and Conclusion

This paper extends the modern treatment effect literature by developing a framework for identifying both the marginal benefit and marginal cost of policies. The treatment effect literature focuses only on the benefit side, and does not address the question of the subjective cost of treatment as perceived by the agents attempting to take it. We build on the pioneering parametric analysis of Björklund and Moffitt (1987) by extending the nonparametric analysis of Heckman and Vytlacil (1999, 2005, 2007) to identify subjective cost and surplus functions. We provide identification results for the case of perfect foresight (as in the previous literature) as well as cases with imperfect foresight not previously considered. An analysis of college-going finds unobserved heterogeneity in the benefits as well as costs of attending college, with agents selecting into college based on both their idiosyncratic expected benefit and perceived cost of attending college. We find more heterogeneity in expected benefits than in perceived cost. Thus, the observed variability in college attendance is mainly driven by the variability in expected benefits.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Aakvik A, Heckman JJ, Vytlacil EJ. Treatment Effects for Discrete Outcomes When Responses to Treatment Vary Among Observationally Identical Persons: An Application to Norwegian Vocational Rehabilitation Programs. Journal of Econometrics. 2005; 125(1–2):15–51.

Abbring, J.; Heckman, JJ. Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation. In: Heckman, JJ.; Leamer, EE., editors. Handbook of Econometrics. Vol. 6B. Elsevier Science; Amsterdam, Netherlands: 2007. p. 5145-5303.

Björklund A, Moffitt R. The Estimation of Wage Gains and Welfare Gains in Self-Selection Models. The Review of Economics and Statistics. 1987; 69(1):42–49.

Bureau of Labor Statistics. NLS Hanbook 2005: The National Longitudinal Surveys. U.S. Department of Labor; Washington, DC: 2005.

Carneiro P, Hansen K, Heckman JJ. Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. International Economic Review. 2003; 44(2):361–422.

Carneiro P, Heckman JJ, Vytlacil EJ. Estimating Marginal Returns to Education. American Economic Review. 2011; 101(6):2754–2781. [PubMed: 25110355]

Chetty R. Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. Annual Review of Economics. 2009; 1:451–488.

D'Haultfoeuille X, Maurel A. Inference on an Extended Roy Model, With an Application to Schooling Decisions in France. Journal of Econometrics. 2013; 174(2):95–106.

Fan, J.; Gijbles, I. Local Polynomial Modelling and its Applications. Chapman and Hall; New York, NY: 1996.

Flinn C, Heckman JJ. New Methods for Analyzing Structural Models of Labor Force Dynamics. Journal of Econometrics. 1982; 18(1):115–168.

Florens JP, Heckman JJ, Meghir C, Vytlacil EJ. Identification of Treatment Effects Using Control Functions in Models with Continuous, Endogenous Treatment and Heterogeneous Effects. Econometrica. 2008; 76(5):1191–1206.

Harberger, AC.; Jenkins, GP., editors. Cost-Benefit Analysis. Edward Elgar Publishers; Northampton, MA: 2002. The International Library of Critical Writings in Economics

Heckman JJ. Shadow Prices, Market Wages, and Labor Supply. Econometrica. 1974; 42(4):679–694.

Heckman JJ. Varieties of Selection Bias. American Economic Review. 1990; 80(2):313–318.

Heckman, JJ. Randomization and Social Policy Evaluation. In: Manski, C.; Garfinkel, I., editors. Evaluating Welfare and Training Programs. Harvard University Press; Cambridge, MA: 1992. p. 201-230.

Heckman JJ. Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policies. Journal of Economic Literature. 2010; 48(2):356–398. [PubMed: 21743749]

Heckman JJ, Honoré BE. The Empirical Content of the Roy Model. Econometrica. 1990; 58(5):1121–1149.

Heckman JJ, Ichimura H, Todd PE. How Details Make a Difference: Semiparametric Estimation of the Partially Linear Regression Model. Unpublished Manuscript. 1997a

Heckman JJ, Sedlacek GL. Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market. Journal of Political Economy. 1985; 93(6):1077–1125.

Heckman, JJ.; Smith, J. Evaluating the Welfare State. In: Strom, S., editor. Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium. Cambridge University Press; New York, NY: 1998. p. 241-318.

Heckman JJ, Smith J, Clements N. Making the Most out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts. The Review of Economic Studies. 1997b; 64(4):487–535.

Heckman JJ, Smith JA. Assessing the Case for Social Experiments. Journal of Economic Perspectives. 1995; 9(2):85–110.

Heckman JJ, Urzua S, Vytlacil EJ. Understanding Instrumental Variables in Models with Essential Heterogeneity. The Review of Economics and Statistics. 2006; 88(3):389–432.

Heckman JJ, Vytlacil EJ. Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects. Proceedings of the National Academy of Sciences. 1999; 96(8):4730–4734.

Heckman, JJ.; Vytlacil, EJ. Instrumental Variables, Selection Models, and Tight Bounds on the Average Treatment Effect. In: Lechner, M.; Pfeiffer, F., editors. Econometric Evaluation of Labour Market Policies. Springer; New York, NY: 2001a. p. 1-15.

Heckman, JJ.; Vytlacil, EJ. Local Instrumental Variables. In: Hsiao, C.; Morimune, K.; Powell, JL., editors. Nonlinear Statistical Modeling: Proceedings of the Thirteenth International Symposium in Economic Theory and Econometrics: Essays in Honor of Takeshi Amemiya. Cambridge University Press; New York, NY: 2001b. p. 1-46.

Heckman JJ, Vytlacil EJ. Structural Equations, Treatment Effects, and Econometric Policy Evaluation. Econometrica. 2005; 73(3):669–738.

Heckman, JJ.; Vytlacil, EJ. Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Economic Estimators to Evaluate Social Programs and to Forecast their Effects in New Environments. In: Heckman, JJ.; Leamer, EE., editors. Handbook of Econometrics. Vol. 6B. Elsevier Science; Amsterdam, NL: 2007. p. 4875-5144.

Hotelling H. The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates. Econometrica. 1938; 6(3):242–269.

Imbens GW, Angrist JD. Identification and Estimation of Local Average Treatment Effects. Econometrica. 1994; 62(2):467–475.

Marschak, J. Economic Measurements for Policy and Prediction. In: Hood, W.; Koopman, T., editors. Studies in Econometric Method. Wiley; New York, NY: 1953. p. 1-26.

Quandt RE. The Estimation of the Parameters of a Linear Regression System Obeying two Separate Regimes. Journal of the American Statistical Association. 1958; 53(284):873–880.

Quandt RE. A New Approach to Estimating Switching Regressions. Journal of the American Statistical Association. 1972; 67(338):306–310.

Robinson PM. Root-N-Consistent Semiparametric Regression. Econometrica. 1988; 56(4):931–954.

Roy A. Some Thoughts on the Distribution of Earnings. Oxford Economic Papers. 1951; 3(2):135–146.

Shaikh AM, Vytlacil EJ. Partial Identification in Triangular Systems of Equations with Binary Dependent Variables. Econometrica. 2011; 79(3):949–955.

Tinbergen, J. Economic Policy: Principles and Design. North Holland Publishing Company; Amsterdam: 1956.

Vytlacil EJ. Independence, Monotonicity, and Latent Index Models: An Equivalence Result. Econometrica. 2002; 70(1):331–341.

Vytlacil EJ, Yildiz N. Dummy Endogenous Variables in Weakly Separable Models. Econometrica. 2007; 75(3):757–779.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

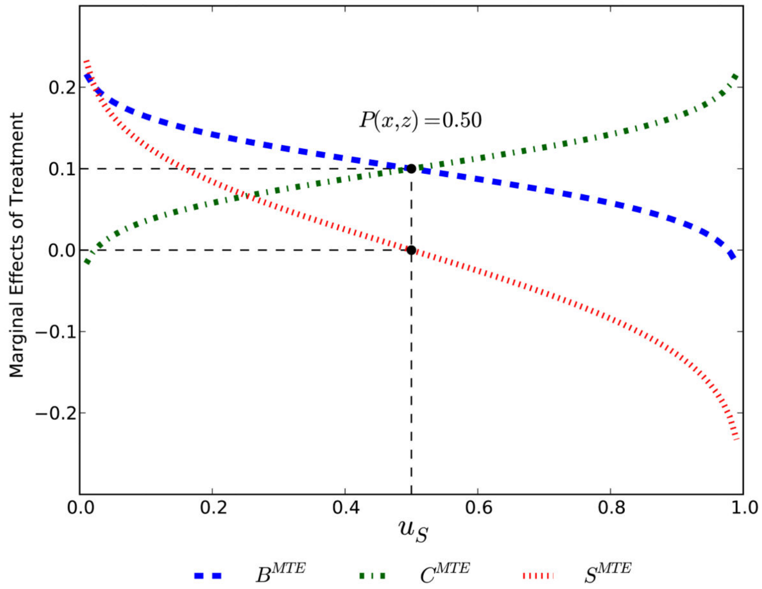
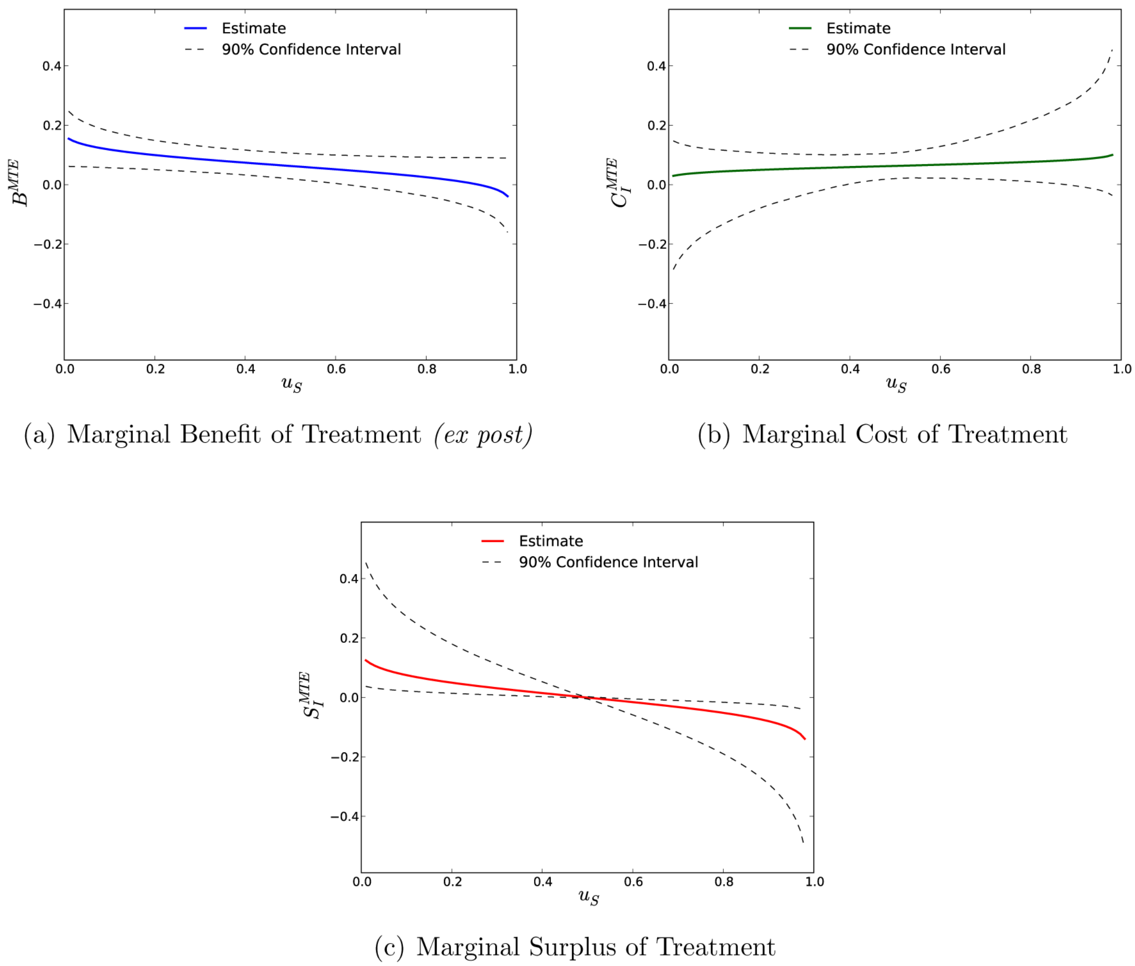


Figure 1.
Marginal Effects of Treatment

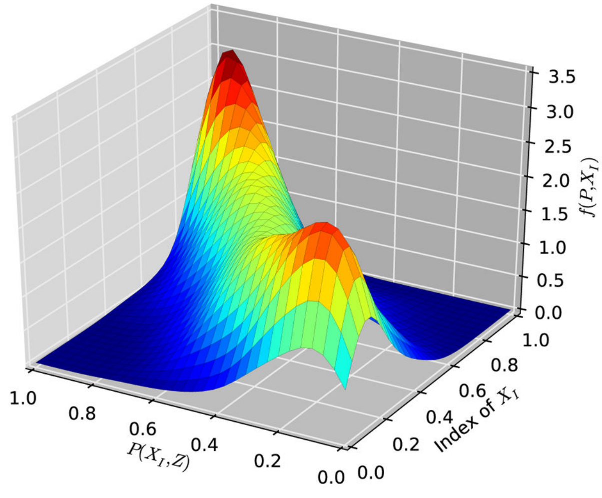


(a) Marginal Benefit of Treatment *(ex post)*

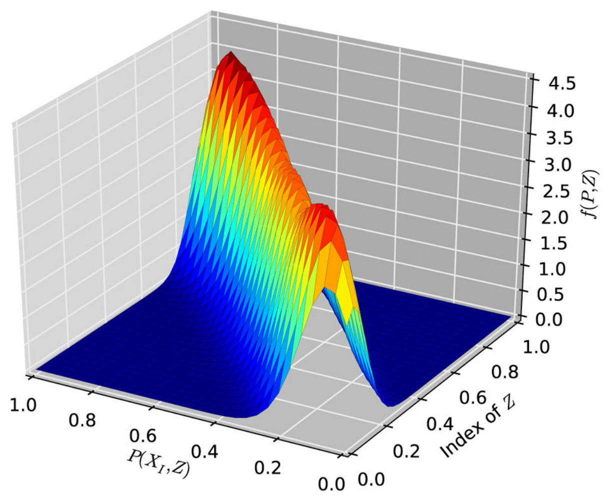
(b) Marginal Cost of Treatment

(c) Marginal Surplus of Treatment

Figure 2.
Marginal Effects of Treatment, Parametric

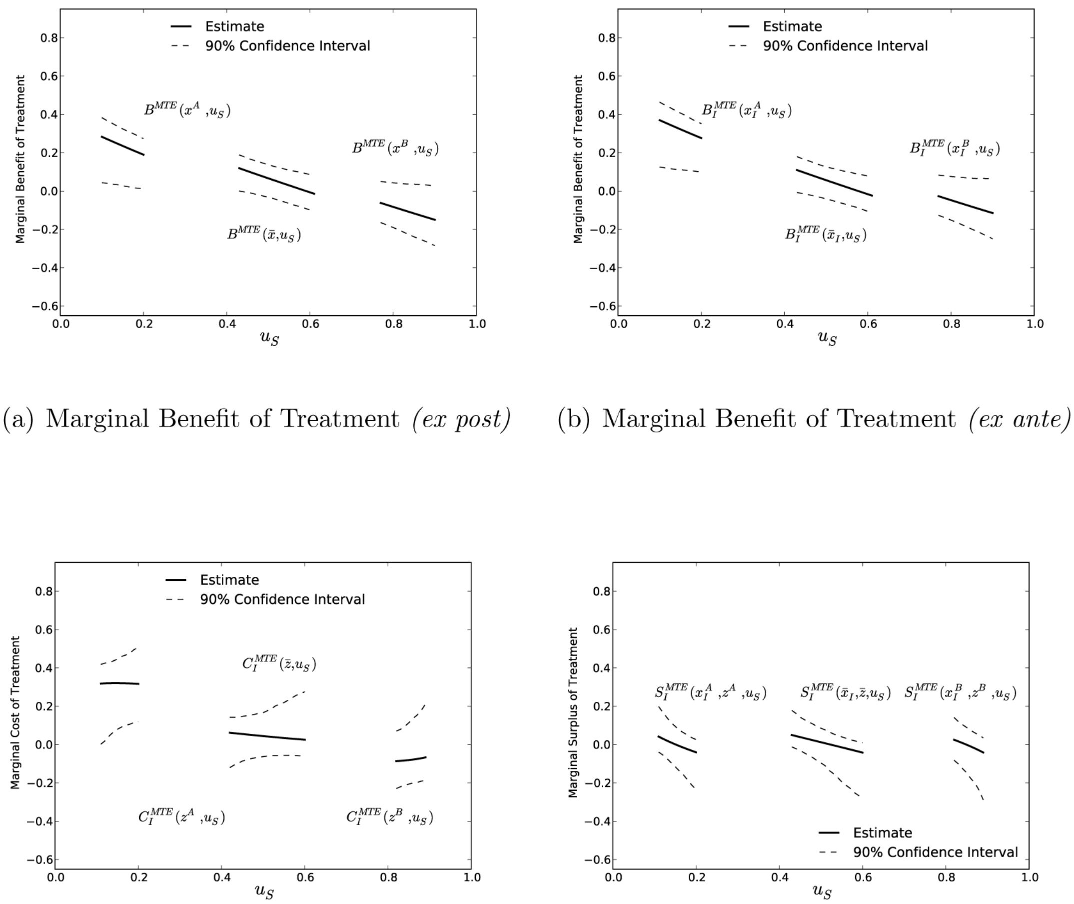


(a) Support of $P(X_I, Z)$ conditional on $X_I(\alpha_1 - \alpha_0)$



(b) Support of $P(X_I, Z)$ conditional on $Z\gamma$

Figure 3.
Conditional Support



(a) Marginal Benefit of Treatment *(ex post)* (b) Marginal Benefit of Treatment *(ex ante)*

(c) Marginal Cost of Treatment (d) Marginal Surplus of Treatment

Figure 4.

Marginal Effects of Treatment, Semiparametric

Table 1

Specification

	X	X_I	Z	Common
Years of Experience (in 1991)	✓			
Current Local Wages (in 1991)	✓			
Permanent Local Wages	✓	✓		
AFQT Scores				✓
Mother's Education				✓
Number of Siblings				✓
Cohort Dummies				✓
Urban Residence			✓	
Local Presence of Public College (age 14)			✓	
Local Tuition at Public College (age 17)			✓	
Local Wages (age 17)			✓	

Notes: Our main specification includes years of experience (linear and squared), current local wages (linear), permanent local wages (linear and squared), AFQT scores (linear and squared), mother's education (linear and squared), number of siblings (linear and squared), urban residence (linear), cohort dummies (linear), local presence of public colleges (linear), local tuition of public college (linear), and local wages (linear). All exclusions from the benefit equation are interacted with AFQT scores, mother's education, and number of siblings.

Table 2

Slope Parameters

Parameter	Estimate	90% Confi.	*p* -val.
$\sigma_{(U_1-U_0),\tilde{V}}$	−0.042	−0.216 / 0.001	0.06
$\sigma_{U_C,\tilde{V}}$	0.015	−0.020 / 0.579	0.29
$\sigma^2_{\tilde{V}}$	0.058	0.005 / 0.769	0.00

Notes: Confi. = Confidence Interval, *p* - val. = *p* -values.