Concave-Monotone Treatment Response and Monotone Treatment Selection: With Returns to Schooling Application*

TSUNAO OKUMURA†

International Graduate School of Social Sciences, Yokohama National University, Japan

EMIKO USUI‡

Graduate School of Economics, Nagoya University, Japan

Abstract

This paper identifies sharp bounds on the mean treatment response and average treatment effect under the concave monotone treatment response (concave-MTR) and monotone treatment selection (MTS) assumptions. We use our bounds and the National Longitudinal Survey of Youth to estimate the mean returns to schooling. Our upper bound estimates of the returns to college education are: $0.079 - 0.133$ for the local returns and 0.087 for the four-year average. They are substantially smaller than (1) the estimates using only the concave-MTR assumption of Manski (1997) and (2) the estimates using only MTR and MTS assumptions of Manski and Pepper (2000). They fall in the lower range of the point estimates given in previous studies which assume linear wage functions. This is because sharp upper bounds are obtained when the curve of the concave-MTR function is close to linear. Our results, therefore, imply a possibility that the higher average returns reported in previous studies are attributed to the specification of linear wage functions.

JEL: C14, J24

Keywords: Partial Identification, Sharp Bounds, Treatment Response, Return to Schooling

---

*Sincere gratitude is extended to Hidehiko Ichimura, Daiji Kawaguchi, Elie Tamer, Edward Vytlacil and especially Charles Manski for their helpful comments. The authors thank participants at the seminars from: the 2006 European Meeting of the Econometric Society, the 2006 Annual Meeting of the Japanese Economic Association, the 2008 Annual Conference of the European Association of Labour Economists, the 2009 WEAI conference, Hitotsubashi University, Kyoto University and University of Tokyo. The main part of this paper was written while the first author was visiting Northwestern University, whose hospitality is gratefully acknowledged. A research grant from the Kikawada Foundation is also appreciated.

†Corresponding author: Tsunao Okumura, International Graduate School of Social Sciences, Yokohama National University, 79-4 Tokiwadai Hodogayaku, Yokohama 2408501, Japan. Phone and Fax: 81-45-339-3524. Email: okumura@ynu.ac.jp.

‡Emiko Usui, Graduate School of Economics, Nagoya University, Nagoya, 464-8601, Japan. Phone and Fax: 81-52-789-4940. Email: usui@soec.nagoya-u.ac.jp.

1 Introduction

This paper studies the identifying power on the treatment effect when the concave-monotone treatment response (concave-MTR) assumption of Manski (1997) is combined with the monotone treatment selection (MTS) assumption of Manski and Pepper (2000), since either assumption produces too wide bounds to have sufficient identifying power. We then apply this nonparametric method to estimate the returns to schooling; in order to assess the validity of the point-estimates reported in the existing parametric studies.

Manski (1997) studies sharp bounds on the mean treatment response, when the response functions are assumed to satisfy the monotone treatment response (MTR) and the concave-monotone treatment response (concave-MTR). In order to enhance the identifying power on the bounds, Manski and Pepper (2000) combine the monotone treatment selection (MTS) assumption with the MTR assumption.[1] They apply their bounds to estimate the returns to schooling. Their bound estimates are narrower than those of Manski (1997); yet large enough that "the MTS-MTR assumption does not, ..., have sufficient identifying power" (Manski and Pepper [p.1009, 2000]), since their bound estimates contain almost all the point estimates of the returns to schooling in the existing empirical literature.

In this paper, we introduce the assumption of concave functions into the assumptions of the MTS and MTR. The concavity is a natural assumption, because diminishing marginal returns are commonly assumed in economic analysis. We explore how the inclusion of this assumption tightens the sharp bounds on the mean treatment response and the average treatment effect.

Using the 2000 wave of the National Longitudinal Survey of Youth (NLSY), we implement our bounds to estimate the returns to schooling. The estimates of our bounds on the returns to schooling are compared to those of Manski's bounds and Manski and Pepper's bounds. Our sharp upper-bound estimates are in the range of $0.05-0.25$. These estimates are reduced to $10-20$ percent of Manski's and to $10-60$ percent of Manski and Pepper's. Specifically, upper-bound estimates on the college education are in the range of $0.079-0.133$ for local

[1] Manski and Pepper (2000) introduce the monotone instrumental variable (MIV) assumption, which weakens the instrumental variable assumption by replacing the mean independence of outcomes and instruments with the mean monotonicity of outcomes and instruments. The MTS assumption is the MIV assumption when the instrumental variable is the realized treatment. When the realized schooling years are used as instruments in the estimation, the MTS assumption means that persons who choose more schooling have weakly higher mean wage functions than do those who choose less schooling.

returns (year-by-year returns), and 0.087 for the four-year average. In contrast, only the MTS-MTR assumption produces $0.236 - 0.384$ for the local returns; and 0.138 for the four-year average. Since the point estimates from the existing literature range between 0.052 and 0.132, our upper-bound estimate on college education falls in the range of the point estimates. The estimates, therefore, appear to have identifying power.

Many empirical studies have estimated the returns to schooling by utilizing the parametric approach. First, years of schooling may be positively correlated with unobserved abilities, which bias the OLS point-estimates upward. The Instrumental Variable technique are utilized; however, the IV estimates tend to be greater than the OLS estimates, despite the prediction of the ability bias hypothesis. In our paper, we deal with the issue of the ability bias by assuming the MTS, in the sense of mean-monotonicity of wages and schooling. Second, almost all of the previous studies assume a log-linear wage regression on the years of schooling. In contrast, the concave-MTR is assumed to allow flexible wage regression. In particular, when the curves of the concave-MTR function are close to linear, sharp upper bounds on the average treatment effect are obtained. Our upper-bound estimates on the returns to college education fall in the lower range of the point estimates. Therefore, for the point-estimates around our upper-bound estimates, the linear wage specification appears appropriate. However, when the point-estimates are greater than our estimates, the linear wage function may be mis-specified. Furthermore, if the wage function were strictly concave, the true returns would be lower than our upper-bound estimates. In this case, almost all of the point estimates would be upper-biased.[2]

In Section 2 we study the sharp bounds on the mean treatment response and the average treatment effects under the assumptions of the concave monotone treatment response and the monotone treatment selection. Section 3 applies the bounds to the estimation of the returns to schooling. We conclude with Section 4.

[2]Belzil and Hansen (2002) use spline techniques and a structural dynamic programming model to deal with nonlinear wage functions and the ability bias. They report smaller point-estimates on the local returns to schooling than those from previous studies.

2 Concave-Monotonicity and Monotone Treatment Selection

The same setup as Manski (1997) and Manski and Pepper (2000) is employed. There is a probability space (J, Ω, P) of individuals. Each member j of population J has an individual-specific response function $y_j(\cdot) : T \rightarrow Y$ mapping the mutually exclusive and exhaustive treatments $t \in T$ into outcomes $y_j(t) \in Y$. Individual j has a realized treatment $z_j \in T$ and a realized outcome $y_j \equiv y_j(z_j)$, both of which are observable. The latent outcomes $y_j(t)$, $t \neq z_j$ are not observable. By combining the distribution of a random sample (z, y) with prior information, our purpose is to learn about the mean treatment response $E[y(\cdot)]$.

Manski (1997) assumes the monotone treatment response (MTR):
Let T be an ordered set, and t_1 and t_2 be elements of T. For each $j \in J$,

$$t_2 \geq t_1 \Longrightarrow y_j(t_2) \geq y_j(t_1).$$

Under the MTR assumption, he shows the sharp bounds on $E[y(t)]$:

$$\begin{aligned} & \sum_{s \leq t} E[y|\, z = s] P(z = s) + y_0 P(z > t) \\ \leq \quad & E[y(t)] \leq \sum_{s \geq t} E[y|\, z = s] P(z = s) + y_1 P(z < t), \end{aligned}$$

where $[y_0, y_1]$ is the range of Y.

Manski (1997) also shows the sharp bounds on $E[y(t)]$ when $y_j(\cdot)$ is a concave and monotone treatment response (concave-MTR); $T = [0, \lambda]$ for some $\lambda \in (0, \infty]$ and $Y = [0, \infty]$:

$$\begin{aligned} & \sum_{s < t} E[y|\, z = s] P(z = s) + E\left[\left.\frac{y}{z} t \right|\, z \geq t\right] P(z \geq t) \\ \leq \quad & E[y(t)] \leq \sum_{s \geq t} E[y|\, z = s] P(z = s) + E\left[\left.\frac{y}{z} t \right|\, z < t\right] P(z < t). \end{aligned}$$

Manski and Pepper (2000) introduce the assumption of monotone treatment selection (MTS):

$$t_2 \geq t_1 \Longrightarrow E[y(t)|\, z = t_2] \geq E[y(t)|\, z = t_1].$$

Under MTS and MTR assumptions, they show the sharp bounds on $E[y(t)]$:

$$\sum_{s<t} E[y|\, z=s]\, P(z=s) + E[y|\, z=t]\, P(z \geq t)$$

$$\leq \quad E[y(t)] \leq \sum_{s>t} E[y|\, z=s]\, P(z=s) + E[y|\, z=t]\, P(z \leq t)\,.$$

In contrast, we assume the concave-MTR and MTS. The following proposition demonstrates the sharp bounds on $E[y(t)]$ under the concave-MTR and MTS assumptions. The basic idea is the following: The MTS assumption implies that, for $u < s$, $E[y|\, z=u] \leq E[y(u)|\, z=s]$. The concave-MTR assumption implies $E[y(\tau)|\, z=s]$ is concave-MTR in $\tau \in T$. Thus, the straight line traversing $(u, E[y|\, z=u])$ and $(s, E[y|\, z=s])$ is the lower bound on $E[y(t)|\, z=s]$ for $u \leq t \leq s$ and the upper bound for $t \geq s$ (Refer to Figure 1). These lines are drawn for all realized $u \leq t$ and the origin. Therefore, the greatest amongst the points for $\tau = t$ on these lines is the lower bound on $E[y(t)|\, z=s]$ for $t \leq s$. As a result, our lower bound on $E[y(t)|\, z=s]$ for $t \leq s$ is not smaller than Manski's (1997) (the line joining $(s, E[y|\, z=s])$ and the origin); and Manski and Pepper's (2000) $(E[y|\, z=t])$. Similarly, our upper bound on $E[y(t)|\, z=s]$ for $t > s$ is not greater than Manski's (1997) and Manski and Pepper's (2000).

Proposition 1 *Let T be ordered. Let $T = [0, \lambda]$ for some $\lambda \in (0, \infty]$ and $Y = [0, \infty]$. Assume that $y_j()$, $j \in J$, satisfies the concave-monotone treatment and MTS assumptions. We extend the set of realized treatments and outcomes by including $(z, y) = (0, 0)$, when there is no realized treatment of 0.*

Then, for $(t, s, s', u) \in T \times T \times T \times T$*,*

$$
\begin{aligned}
& \sum_{s<t} E\left[y\right| z=s] \, P(z=s) \\
& + \sum_{s \geq t} \max_{\{s' | s \geq s' \geq t\}} \left(E\left[y\right| z=s'] + \max_{\{u | u<t\}} \left\{ \frac{E\left[y\right| z=s'] - E\left[y\right| z=u]}{s'-u} (t-s') \right\} \right) P(z=s) \\
\leq \; & E[y(t)] \\
\leq \; & \sum_{s>t} E\left[y\right| z=s] \, P(z=s) \\
& + \sum_{s \leq t} \min_{\{s' | s \leq s' \leq t\}} \left\{ E\left[y\right| z=s'] + \min_{\{u | u<s'\}} \frac{E\left[y\right| z=s'] - E\left[y\right| z=u]}{s'-u} (t-s') \right\} P(z=s) \, .
\end{aligned}
$$

These bounds are sharp; and narrower than or equal to those using only the concave-MTR assumption of Manski (1997), as well as those using only MTR and MTS assumptions of Manski and Pepper (2000).

Proof.

For $u < s$, $E\left[y(u)\right| z=s] \geq E\left[y(u)\right| z=u] = E\left[y\right| z=u]$ by MTS.
$E\left[y\right| z=s] = E\left[y(s)\right| z=s] \geq E\left[y(u)\right| z=s]$ by MTR.
Hence, $E\left[y\right| z=s] \geq E\left[y(u)\right| z=s] \geq E\left[y\right| z=u]$.
Since $y_j(t)$ is concave-monotone for all $j \in J$, $E\left[y(t)\right| z=s]$ is concave-monotone in t.
Therefore, for $t \geq s > u$,

$$
E\left[y(t)\right| z=s] \leq E\left[y\right| z=s] + \frac{E\left[y\right| z=s] - E\left[y\right| z=u]}{s-u} (t-s) \, , \tag{1}
$$

and for $s \geq t > u$,

$$
E\left[y(t)\right| z=s] \geq E\left[y\right| z=s] + \frac{E\left[y\right| z=s] - E\left[y\right| z=u]}{s-u} (t-s) \, . \tag{2}
$$

Since Equation (1) holds for any $u < s \leq t$, for $t \geq s$,

$$
E\left[y(t)\right| z=s] \leq E\left[y\right| z=s] + \min_{\{u | u<s\}} \frac{E\left[y\right| z=s] - E\left[y\right| z=u]}{s-u} (t-s) \, . \tag{3}
$$

Similarly, since Equation (2) holds for any $u < t \leq s$, for $t \leq s$,

$$
E\left[y(t)\right| z=s] \geq E\left[y\right| z=s] + \max_{\{u | u<t\}} \left\{ \frac{E\left[y\right| z=s] - E\left[y\right| z=u]}{s-u} (t-s) \right\} . \tag{4}
$$

MTS implies that the upper and lower bounds on $E\left[\left.y(t)\right|z=s\right]$ are weakly increasing in s. Thus, Equations (3) and (4) imply that, for $t \geq s$,

$$E\left[\left.y(t)\right|z=s\right] \leq \min_{\{s'|s\leq s'\leq t\}} \left\{ E\left[\left.y\right|z=s'\right] + \min_{\{u|u<s'\}} \frac{E\left[\left.y\right|z=s'\right]-E\left[\left.y\right|z=u\right]}{s'-u}(t-s') \right\}, \tag{5}$$

and for $t \leq s$,

$$E\left[\left.y(t)\right|z=s\right] \geq \max_{\{s'|s\geq s'\geq t\}} \left(E\left[\left.y\right|z=s'\right] + \max_{\{u|u<t\}} \left\{ \frac{E\left[\left.y\right|z=s'\right]-E\left[\left.y\right|z=u\right]}{s'-u}(t-s') \right\} \right). \tag{6}$$

Applying Equations (5) and (6) to the Law of Iterated Expectations yields the second terms of the upper and lower bounds on $E\left[y(t)\right]$ in the Proposition, respectively.

Manski (1997) and Manski and Pepper (2000) show that under either concave-MTR or MTS-MTR assumptions, for $s \geq t$,

$$E\left[\left.y(t)\right|z=s\right] \leq E\left[\left.y\right|z=s\right], \tag{7}$$

and for $s < t$,

$$E\left[\left.y(t)\right|z=s\right] \geq E\left[\left.y\right|z=s\right]. \tag{8}$$

This implies the first terms of the upper and lower bounds on $E\left[y(t)\right]$ in the Proposition. Thus, these results yield the bounds on $E\left[y(t)\right]$ in the Proposition.

The proof of the sharpness of the bounds is provided in the Appendix.

As Manski (1997) shows, because $T = [0, \lambda]$ and $Y = [0, \infty]$, for $t \geq s$,

$$E\left[\left.y(t)\right|z=s\right] \leq E\left[\left.y\right|z=s\right] + \frac{E\left[\left.y\right|z=s\right]}{s}(t-s) = E\left[\left.\frac{y}{z}t\right|z=s\right],$$

and for $t \leq s$,

$$E\left[\left.y(t)\right|z=s\right] \geq E\left[\left.y\right|z=s\right] + \frac{E\left[\left.y\right|z=s\right]}{s}(t-s) = E\left[\left.\frac{y}{z}t\right|z=s\right].$$

Since $(z, y) = (0, 0)$ in the set of the observations yields $E\left[\left.y\right|z=0\right] = 0$, these bounds have been included in Equations (3) and (4).

Manski and Pepper (2000) show that the MTR-MTS assumption implies that for $t > s$,

$$E\left[\left.y(t)\right|z=s\right] \leq E\left[\left.y\right|z=t\right],$$

and for $t \leq s$,

$$E\left[\left.y(t)\right|z=s\right] \geq E\left[\left.y\right|z=t\right].$$

These bounds are included in the case of $s' = t$ in Equations (5) and (6).

Manski's (1997) and Manski and Pepper's (2000) bounds are included in the brackets of Equations (5) and (6), and Equations (7) and (8). Equations (5) and (6) take minimum and maximum of the objects within these brackets, respectively. Therefore, the bounds in the Proposition are narrower than or equal to those of Manski (1997) and Manski and Pepper (2000). ∎

The introduction of the assumption of concavity into the MTR and MTS assumptions narrows the width of the bounds on $E[y(t)]$ by,

$$\sum_{s\geq t}\left\{\max_{\{s'|s\geq s'\geq t\}}\left(E[y|\,z=s'] + \max_{\{u|u<t\}}\left\{\frac{E[y|\,z=s'] - E[y|\,z=u]}{s'-u}(t-s')\right\}\right) - E[y|\,z=t]\right\}$$
$$\times P(z=s)$$

$$+\sum_{s\leq t}\left(E[y|\,z=t] - \min_{\{s'|s\leq s'\leq t\}}\left\{E[y|\,z=s'] + \min_{\{u|u<s'\}}\frac{E[y|\,z=s'] - E[y|\,z=u]}{s'-u}(t-s')\right\}\right)$$
$$\times P(z=s). \tag{9}$$

The first term shows the increase in the lower bound and the second term demonstrates the decrease in the upper bound.

The sharp bounds on the average treatment effects, $E[y(t_2)] - E[y(t_1)]$ for $t_1 < t_2$, are given in the following Proposition. The response functions satisfying the sharp bounds are represented in the graph of the boundary of the convex hull (or convex envelope) of the observations which are not greater than the realized treatments. The convex hull is the smallest convex set containing observations (with the concave boundary) and thus, a subset of the convex hull containing more observations. Therefore, the response functions represented in its boundary satisfy the concave-MTR, MTS and sharp bounds.

Proposition 2 *Let T be ordered. Let $T = [0,\lambda]$ for some $\lambda \in (0,\infty]$ and $Y = [0,\infty]$. Assume that $y_j()$, $j \in J$, satisfies the concave-monotone treatment and MTS assumptions. Suppose $t \in T$, $s \in T$, $v \in T$, $w \in T$, $t_1 \in T$, $t_2 \in T$. We extend the set of realized treatments and outcomes by including $(z,y) = (0,0)$, when there is no realized treatment of 0. We then define the following functions.*

(1) For $t \geq z = s$:

$$v^m(s) = \arg\min_{\{v|v<v^{m-1}(s)\}}\frac{E[y|\,z=v^{m-1}(s)] - E[y|\,z=v]}{v^{m-1}(s)-v}, \quad \textit{for } m = 1,2,...,\ M(s),$$

where $v^0(s) = s$ *and* $M(s)$ *satisfies* $v^{M(s)}(s) = 0$.

(i) For $s \leq w < t$,

$$LBAT_1(w,s,t) = E[y|\, z = s] + \frac{UB(s,t) - E[y|\, z = s]}{t-s}(w-s),$$

where

$$UB(s,t) = \min_{\{s'|s \leq s' \leq t\}} \left\{ E[y|\, z = s'] + \min_{\{u|u<s'\}} \frac{E[y|\, z = s'] - E[y|\, z = u]}{s' - u}(t - s') \right\}.$$

(ii) For $v^m(s) \leq w < v^{m-1}(s) \quad (m = 1, 2, ..., M(s))$,

$$LBAT_1(w,s,t) = E\left[y|\, z = v^{m-1}(s)\right] + \frac{E[y|\, z = v^{m-1}(s)] - E[y|\, z = v^m(s)]}{v^{m-1}(s) - v^m(s)}\left[w - v^{m-1}(s)\right].$$

(2) For $t < z = s$:

$$v(s,t) = \arg\min_{\{v|v<t\}} \frac{E[y|\, z = s] - E[y|\, z = v]}{t - v},$$

$$v^m(s,t) = \arg\min_{\{v|v<v^{m-1}(s,t)\}} \frac{E[y|\, z = v^{m-1}(s,t)] - E[y|\, z = v]}{v^{m-1}(s,t) - v} \quad \text{for } m = 2, 3, ..., M(s,t),$$

where $M(s,t)$ *satisfies* $v^{M(s,t)}(s,t) = 0$.

(i) For $v(s,t) \leq w < t < z = s$,

$$LBAT_2(w,s,t) = E[y|\, z = s] + \frac{E[y|\, z = s] - E[y|\, z = v(s,t)]}{t - v(s,t)}[w - t].$$

(ii) For $v^m(s,t) \leq w < v^{m-1}(s,t) \quad (m = 2, 3, ..., M(s,t))$,

$$LBAT_2(w,s,t) = E\left[y|\, z = v^{m-1}(s,t)\right] + \frac{E[y|\, z = v^{m-1}(s,t)] - E[y|\, z = v^m(s,t)]}{v^{m-1}(s,t) - v^m(s,t)}\left[w - v^{m-1}(s,t)\right].$$

Then,

$$\begin{aligned} 0 &\leq E[y(t_2)] - E[y(t_1)] \\ &\leq \sum_{s \leq t_2} [UB(s,t_2) - LBAT_1(t_1,s,t_2)]\, P(z = s) \\ &\quad + \sum_{s > t_2} \{E[y|\, z = s] - LBAT_2(t_1,s,t_2)\}\, P(z = s). \end{aligned} \tag{10}$$

These bounds are sharp.

Proof. The lower bound on $E[y(t_2)] - E[y(t_1)]$ holds because $y_j(\tau)$ is monotone. It is sharp because the hypothesis $\{y_j(t_1) = y_j(t_2) = y_j, j \in J\}$ satisfies the concave-MTR and MTS (since $E[y|\, z = s]$ is increasing in s).
To obtain the sharp upper bound on $E[y(t_2)] - E[y(t_1)]$, let us obtain the sharp upper bound on $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$.
Proof of Proposition 1 implies:

$$E[y(t_2)|\, z = s] \leq UB(s, t_2) \qquad \text{for } z = s \leq t_2,$$

$$E[y(t_2)|\, z = s] \leq E[y|\, z = s] \quad \text{for } t_2 < s = z,$$

and these upper bounds are sharp.
Therefore, to obtain the sharp upper bound on $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$, (1) for $z = s \leq t_2$, given $E[y(t_2)|\, z = s] = UB(s, t_2)$, minimize $E[y(t_1)|\, z = s]$ subject to the condition that $E[y(\tau)|\, z = s]$ for $\tau \in T$ traverses $(t_2, UB(z, t_2))$, $(t_1, E[y(t_1)|\, z = s])$ and $(z, E[y|\, z = s])$; and satisfies the concave-MTR and MTS. (2) For $t_2 < s = z$, given $E[y(t_2)|\, z = s] = E[y|\, z = s]$, minimize $E[y(t_1)|\, z = s]$ subject to the condition that $E[y(\tau)|\, z = s]$ for $\tau \in T$ traverses $(t_2, E[y|\, z = s])$, $(t_1, E[y(t_1)|\, z = s])$ and $(z, E[y|\, z = s])$; and satisfies the concave-MTR and MTS.
(1) For $z = s \leq t_2$;
(i) for $z = s \leq t_1 < t_2$, given $E[y(t_2)|\, z = s] = UB(s, t_2)$, $E[y(t_1)|\, z = s] = LBAT_1(t_1, z, t_2)$ yielding the maximal value of $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$. If $E[y(t_1)|\, z = s] < LBAT_1(t_1, z, t_2)$, $E[y(\tau)|\, z = s]$ traversing $(z, E[y|\, z = s])$ and $(t_2, UB(s, t_2))$ is not concave-MTR. The Appendix proves that $LBAT_1(\tau, z, t_2)$ is concave-MTR and MTS in $\tau \in T$. Since $LBAT_1(\tau, z, t_2)$ traverses $(t_2, UB(z, t_2))$, $(t_1, LBAT_1(t_1, z, t_2))$ and $(z, E[y|\, z = s])$, one can take $LBAT_1(\tau, z, t_2)$ as $E[y(\tau)|\, z = s]$.
(ii) for $t_1 < z = s \leq t_2$, given $E[y(t_2)|\, z = s] = UB(s, t_2)$, $E[y(t_1)|\, z = s] = LBAT_1(t_1, z, t_2)$ yielding the maximal value of $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$. If $E[y(t_1)|\, z = s] < LBAT_1(t_1, z, t_2)$, the concave-MTR function $E[y(\tau)|\, z = s]$ implies $E[y(v^m(s))|\, z = s] < E[y|\, z = v^m(s)]$ for $v^m(s) \leq t_1 < v^{m-1}(s) \leq s$; thus, violating the MTS of $E[y(\tau)|\, z = s]$. As the Appendix shows, $LBAT_1(\tau, z, t_2)$ is concave-MTR and MTS; and traverses $(t_2, UB(z, t_2))$, $(t_1, LBAT_1(t_1, z, t_2))$ and $(z, E[y|\, z = s])$. Therefore, one can take $LBAT_1(\tau, z, t_2)$ as $E[y(\tau)|\, z = s]$.
(2) $t_2 < s = z$;

(i) For $v(s,t_2) \le t_1 < t_2 < z = s$; given $E[y(t_2)|\, z=s] = E[y|\, z=s]$, $E[y(t_1)|\, z=s] = LBAT_2(t_1,z,t_2)$ yields the maximal value of $E[y(t_2)|\, z=s] - E[y(t_1)|\, z=s]$. If $E[y(t_1)|\, z=s] < LBAT_2(t_1,z,t_2)$, the concave-MTR function $E[y(\tau)|\, z=s]$ implies $E[y(v(s,t_2))|\, z=s] < E[y|\, z=v(s,t_2)]$; thereby, violating the MTS of $E[y(\tau)|\, z=s]$. The Appendix proves that $LBAT_2(\tau,z,t_2)$ is concave-MTR and MTS in $\tau \in T$. Since $LBAT_2(\tau,z,t_2)$ traverses $(t_2, E[y|\, z=s])$, $(t_1, LBAT_2(t_1,z,t_2))$ and $(z, E[y|\, z=s])$, one can take $LBAT_2(\tau,z,t_2)$ as $E[y(\tau)|\, z=s]$.

(ii) For $v^m(s,t_2) \le t_1 < v^{m-1}(s,t_2) \le t_2 < z = s$; given $E[y(t_2)|\, z=s] = E[y|\, z=s]$, $E[y(t_1)|\, z=s] = LBAT_2(t_1,z,t_2)$ yields the maximal value of $E[y(t_2)|\, z=s] - E[y(t_1)|\, z=s]$. If $E[y(t_1)|\, z=s] < LBAT_2(t_1,z,t_2)$, the concave-MTR function $E[y(\tau)|\, z=s]$ implies $E[y(v^m(s,t_2))|\, z=s] < E[y|\, z=v^m(s,t_2)]$; thus, violating the MTS of $E[y(\tau)|\, z=s]$. As the Appendix shows, $LBAT_2(\tau,z,t_2)$ is concave-MTR and MTS; and traverses $(t_2, E[y|\, z=s])$, $(t_1, LBAT_2(t_1,z,t_2))$ and $(z, E[y|\, z=s])$. Therefore, one can take $LBAT_2(\tau,z,t_2)$ as $E[y(\tau)|\, z=s]$.

In conclusion,

(1) For $z = s \le t_2$, and $t_1 < t_2$,

$$0 \le E[y(t_2)|\, z=s] - E[y(t_1)|\, z=s] \le UB(s,t_2) - LBAT_1(t_1,s,t_2).$$

(2) For $t_1 < t_2 < s = z$,

$$0 \le E[y(t_2)|\, z=s] - E[y(t_1)|\, z=s] \le E[y|\, z=s] - LBAT_2(t_1,s,t_2).$$

These bounds are sharp.

By employing the Law of Iterated Expectations, the sharp bounds on $E[y(t_2)] - E[y(t_1)]$ in Proposition 2 are obtained. ■

Proposition 2 shows that our sharp upper bound on $\Delta(t_1,t_2)$ is attained when $LBAT_1(t_1,z,t_2)$ is the mean response function of the individuals whose realized treatment (z) is not greater than t_2; and $LBAT_2(t_1,z,t_2)$ is the mean response function of the individuals whose realized treatment (z) is greater than t_2. The $LBAT_1(t_1,z,t_2)$ function is represented in the graph of the boundary of the convex hull of the observations which are not greater than z; and a line which joins $(t_2, UB(t_1,t_2))$ and $(z, E[y|\, z])$. The $LBAT_2(t_1,z,t_2)$ function is represented in the graph of the boundary of the convex hull of the observations whose treatments are not greater than t_2. Therefore, the curves of the $LBAT_1(t_1,z,t_2)$ and $LBAT_2(t_1,z,t_2)$ functions are close to linear between t_1 and t_2; so are the curves of the individual response functions.

3 Estimation of the Returns to Schooling

We use the 2000 wave of the National Longitudinal Survey of Youth (NLSY), which is representative of the U.S. non-institutionalized civilian population between the ages of 14 and 22 in 1979. As Manski and Pepper (2000) who use the 1994 wave of the NLSY, a random sample of white men is utilized; who reported that they were full-time, year-round workers and not self-employed. The sample size is 1240. Their hourly rate of pay and realized years of schooling were observed. In our application to the returns to schooling, z represents the realized years of schooling; the response variable $y_j(t)$ is the logarithm of hourly rate of pay a person j would obtain if he were to have t years of schooling; and $y_j = y_j(z)$ is the logarithm of the observed hourly wage.[3]

Table 1 shows the estimates of $E(y|z)$ and $P(z)$: Table 1(a) for all samples and Table 1(b) for the sample excluding the observations with seven and eight years of schooling. Forty-one percent of the NLSY respondents have 12 years of schooling and 18 percent have 16 years of schooling in Table 1(a). For the most part, the estimates of $E(y|z)$ in Table 1(a) increase with z. However, there are three dips, which conflict with the monotone treatment selection assumption. Because of these dips, we exclude the observations with seven and eight years of schooling (Table 1(b)) and the negative $\{E[y|\, z = s'] - E[y|\, z = u]\} \diagup (s' - u)$ when computing $UB(s,t)$ for estimation of the bounds.[4]

Table 2 reports the estimates of the bounds on $E[y(t)]$ in Proposition 1. The 5 percent

[3] We exclude three individuals whose wages are less than one dollar. Thus, the support of Y is $[0, \infty]$.

[4] First, due to the small sample size, we exclude the observations with seven years of schooling following Manski and Pepper (2000). Including these observations does not change the conclusion. Second, three dips are found at nine, fourteen and nineteen years of schooling. Manski and Pepper (2000) also find three dips in their data. However, they assume that the MTS and MTR assumption is consistent with empirical evidence and use all samples (except for the observations with seven years of schooling). When they compute the uniform 95 percent confidence intervals for the estimates of $E(y|z)$, the intervals contain everywhere monotone functions. We also find that the uniform 95 percent confidential intervals for the estimates of $E(y|z)$ (using our data) contain everywhere monotone functions. Nevertheless, we deal with the problem associated with the dips by utilizing the following: (i) the sample excluding the observations with eight years of schooling, and (ii) the sample excluding the observations with eight, nineteen and twenty years of schooling. In terms of the bound calculation, we take two approaches: (a) when computing $UB(s,t)$, given $s' \in [s,t]$, if there is $u < s'$ such that the slope $\{E[y|\, z = s'] - E[y|\, z = u]\} \diagup (s' - u)$ is negative, we take the minimum slope among the nonnegative slopes by excluding the only negative slopes given by this (s', u), and (b) if the slope is negative, we drop all of the (negative and nonnegative) slopes given by this s' and take the minimum slope for other $s' \in [s,t]$. The Sample (i) and Approach (a) are employed in Tables 2 and 3. This combination is utilized because (1) when the sample of the eight years of schooling is included, the bound estimates on the returns to schooling are much tighter (especially for $8 - 12$ years of schooling); however, including the nineteen and twenty years of schooling has little effect on the bound estimate; and (2) the difference in estimates between using approaches (a) and (b) are negligible.

and 95 percent bootstrap quantiles are also shown. For comparison, Table 2 reports the estimates of the bounds using only the MTR and MTS (Manski and Pepper's (2000) bounds), and the estimates of the bounds using only the concave-MTR (Manski's (1997) bounds). The estimates of our bounds are much narrower than the estimates of Manski and Pepper (2000) and Manski (1997). Specifically, our lower-bound estimates increase more with less years of schooling, whereas our upper-bound estimates decrease more with greater years of schooling, compared to the estimates of Manski and Pepper (2000). Equation (9) shows that when t is small, the first term (the increase in the lower bound) dominates the reduction in the width of the bounds, whereas, when t is large, the second term (the decrease in the upper bound) dominates it.

Table 3 reports the estimates of the upper bounds on the average treatment effect, $E[y(t)] - E[y(t-1)]$ (denoted by $\Delta(t-1, t)$, for $t = 10, ..., 20$) in Proposition 2. Our bound estimates and the 95 percent bootstrap quantiles are listed in Columns 1 and 2; the bound estimates and the quantiles using only MTR and MTS of Manski and Pepper (2000) in Columns 3 and 4; and the bound estimates and the quantiles using only the concave-MTR of Manski (1997) in Columns 5 and 6. Our upper-bound estimates on the average treatment effect are in the range of $0.045 - 0.165$ for year-by-year estimates for schooling greater than 11 years. Our upper-bound estimates are reduced to $10 - 60$ percent as compared to those using only MTR and MTS from Manski and Pepper (2000). Our upper-bound estimates are also reduced to $10 - 20$ percent as compared to those using only concave-MTR from Manski (1997). The upper-bound estimates on local returns to college education ($\Delta(12, 13)$, $\Delta(13, 14)$, $\Delta(14, 15)$, and $\Delta(15, 16)$) are between 0.079 and 0.133. The upper-bound estimates on $\Delta(12, 16)$ imply that the completion of four-years of college yields, at most, an increase of 0.346 in mean log wage relative to the completion of high school. The average value of the year-by-year college education treatment effect is at most 0.087.

Many empirical studies have estimated the returns to schooling by utilizing the ordinary least squares (OLS) and Instrumental Variables (IV) techniques. Card (1999) in his survey shows that the point-estimates on the returns to schooling in the previous studies of US data are in the range of $0.052 - 0.132$; and raises two serious questions on the credibility of these point-estimates.[5] First, almost all of the previous studies assume that the log wage

[5] The point-estimate on the returns to schooling is 0.1009 (and the standard error is 0.0052), when log wages are regressed on years of schooling utilizing our data and the OLS technique.

regression function is linear in years of schooling. However, it is controversial to assume that each additional year of schooling has the same proportional effect on earnings, despite heterogeneous components of schooling choices. The canonical human capital models assume that the human capital production function is concave-increasing in schooling. Second, since years of schooling are considered positively correlated with unobserved abilities (the ability bias and selection problems), it is more appropriate to utilize the Instrumental Variable technique. However, the validity of Instrumental Variables used in applications often are questioned. Indeed, the IV estimates tend to be greater than the OLS estimates, despite the prediction of the ability bias hypothesis. As a result, Card (1999) argues the possibility that the point-estimates on the returns to schooling from the previous literature are over-estimated. In contrast, the MTS and concave-MTR assumptions which we impose deal with the ability bias/selection problems and the concave-increasing wage (human capital production) functions. Our estimates of the upper bound fall in the lower range of the point estimates of previous research. When the curve of the concave-MTR response functions is close to linear, our upper bounds are sharp. Therefore, the point-estimates around our sharp upper-bound estimates are consistent with the assumption of linear wage function. The returns which are smaller than the upper-bound estimates are obtained, when the functions are strictly concave. However, the returns which are greater than the upper-bound estimates may be overestimated. Therefore, our estimation result implies the possibility that higher point-estimates on the returns to schooling of previous studies are attributed to the linearity of wage functions.

Belzil and Hansen (2002) use spline techniques and a structural dynamic programing model to deal with nonlinear wage functions and the ability bias. They report smaller point-estimates on the local returns to schooling than those from previous studies, and cast doubts on the validity of linear wage functions which previous studies assume.[6] Our approach also does not assume the linearity of wage functions and finds smaller estimates. Therefore, our estimation result corroborates this critique.

[6]Belzil and Hansen (2002) find that the log wage regression is convex in schooling. They state that the convexity of the log wage regression function implies that those endowed with higher school ability experience higher returns. This is consistent with the MTS.

4 Conclusion

When the assumption of a concave response function is introduced into Manski and Pepper's (2000) assumption of monotone treatment response and monotone treatment selection, sharp bounds on the mean treatment response and the average treatment effect are obtained. Bounds on the returns to schooling are estimated by utilizing our bounds and the NLSY data. The estimates of our bounds are found to be substantially tighter than those of Manski and Pepper (2000). Our upper bound estimates of the returns to college education are: $0.079 - 0.133$ for local returns, and 0.087 for the four-year average.

The returns to schooling have been estimated in many empirical studies. Suffering from ability bias and selection problems, the returns are overestimated in any regression model. Almost all studies assume that the wage regression is log linear as regards schooling. Improper inference is placed on the returns to schooling due to a misspecification of the wage function. The MTS assumption deals with ability bias and selection problems. The concave-MTR assumption allows flexible wage regression. Our upper bound estimates on the returns to schooling (obtained under the MTS and concave-MTR assumptions) are smaller than many point-estimates reported in previous literature. Therefore, our estimation results cast doubt on the validity of the high average returns to schooling.

5 Appendix: Proof of the Sharp Bounds in Proposition 1

In order to show that the bounds are sharp, it suffices to demonstrate that the concave-MTR and MTS functions of $y_j(\tau)$ for $\tau \in T$ attain lower and upper bounds.

1. Existence of the concave-MTR and MTS functions of $y_j(\tau)$ for $\tau \in T$ that attain the *upper* bound.

Denote,

$$UB(s,t) = \min_{\{s'|s \le s' \le t\}} \left\{ E[y|\, z=s'] + \min_{\{u|u<s'\}} \frac{E[y|\, z=s'] - E[y|\, z=u]}{s'-u}(t-s') \right\}.$$

For $s < t$, take,

$$E[y(\tau)|\, z=s] = \min\left(\min_{\{\widetilde{s}|s \le \widetilde{s} < t\}} \left\{ E[y|\, z=\widetilde{s}] + \frac{UB(\widetilde{s},t) - E[y|\, z=\widetilde{s}]}{t-\widetilde{s}}(\tau - \widetilde{s}) \right\},\ E[y|\, z=t] \right),$$

for $\tau \in T$.

For $s \ge t$, take,

$$E[y(\tau)|\, z=s] = E[y|\, z=s],$$

for $\tau \in T$.

These functions attain upper bounds in the Proposition since $UB(s,t)$ and $E[y|\, z=s]$ weakly increase in s, and $E[y|\, z=s] \le UB(s,t) \le E[y|\, z=t]$ for $s \le t$.

$E[y(\tau)|\, z=s]$ satisfies the concave-MTR, as by definition its graph is the boundary of the convex hull; that is, the intersection of the subgraphs of the weakly increasing linear functions in τ.[7] $E[y(\tau)|\, z=s]$ also satisfies the MTS, since by definition the object is minimized over the set of $\{\widetilde{s}|s \le \widetilde{s} < t\}$ such that $\{\widetilde{s}|s_1 \le \widetilde{s} < t\} \supseteq \{\widetilde{s}|s_2 \le \widetilde{s} < t\}$ for $s_1 \le s_2$.

We show $E[y(s)|\, z=s] = E[y|\, z=s]$.

First, for $\zeta \le \widehat{s} < t$,

$$E[y|\, z=\zeta] \le E[y|\, z=\widehat{s}] + \min_{\{u|u<\widehat{s}<t\}} \left\{ \frac{E[y|\, z=\widehat{s}] - E[y|\, z=u]}{\widehat{s}-u} \right\}(\zeta - \widehat{s}).$$

It is because,

$$\begin{aligned} E[y|\, z=\zeta] &= E[y|\, z=\widehat{s}] + \frac{E[y|\, z=\widehat{s}] - E[y|\, z=\zeta]}{\widehat{s}-\zeta}(\zeta - \widehat{s}) \\ &\le E[y|\, z=\widehat{s}] + \min_{\{u|u<\widehat{s}<t\}} \left\{ \frac{E[y|\, z=\widehat{s}] - E[y|\, z=u]}{\widehat{s}-u} \right\}(\zeta - \widehat{s}). \end{aligned}$$

[7] The subgraph of $f(\tau)$ is defined as $\{(\tau,y)|\, y \le f(\tau)\}$.

Second,

$$\min_{\{u|u<s<t\}} \frac{E\left[y|\,z=s\right]-E\left[y|\,z=u\right]}{s-u} \geq \frac{UB\left(s,t\right)-E\left[y|\,z=s\right]}{t-s}, \tag{11}$$

since,

$$E\left[y|\,z=s\right]+\min_{\{u|u<s<t\}} \frac{E\left[y|\,z=s\right]-E\left[y|\,z=u\right]}{s-u}\left(t-s\right) \geq UB\left(s,t\right).$$

Therefore, for $s \leq \widetilde{s} < t$,

$$E\left[y|\,z=s\right] \leq E\left[y|\,z=\widetilde{s}\right]+\frac{UB\left(\widetilde{s},t\right)-E\left[y|\,z=\widetilde{s}\right]}{t-\widetilde{s}}\left(s-\widetilde{s}\right).$$

Thus,

$$E\left[y\left(s\right)|\,z=s\right]=E\left[y|\,z=s\right].$$

2. Existence of the concave-MTR and MTS functions of $y_j\left(\tau\right)$ for $\tau \in T$ that attain the *lower* bound.

Denote,

$$\begin{aligned} s'^{*}\left(s,t\right) &\equiv \arg\max_{\{s'|s\geq s'\geq t\}}\left(E\left[y|\,z=s'\right]+\max_{\{u|u<t\}}\left\{\frac{E\left[y|\,z=s'\right]-E\left[y|\,z=u\right]}{s'-u}\left(t-s'\right)\right\}\right). \\ u^{*}\left(s,t\right) &\equiv \arg\max_{\{u|u<t\}}\left\{\frac{E\left[y|\,z=s'^{*}\left(s,t\right)\right]-E\left[y|\,z=u\right]}{s'^{*}\left(s,t\right)-u}\left[t-s'^{*}\left(s,t\right)\right]\right\} \end{aligned}$$

Denote,

$$LB\left(\tau,s,t\right) \equiv \min_{\{\widetilde{s}|\widetilde{s}\geq s\}}\min\left\{\begin{array}{c} E\left[y|\,z=s'^{*}\left(\widetilde{s},t\right)\right]+\frac{E\left[y|z=s'^{*}\left(\widetilde{s},t\right)\right]-E\left[y|z=u^{*}\left(\widetilde{s},t\right)\right]}{s'^{*}\left(\widetilde{s},t\right)-u^{*}\left(\widetilde{s},t\right)}\left[\tau-s'^{*}\left(\widetilde{s},t\right)\right], \\ E\left[y|\,z=\widetilde{s}\right] \end{array}\right\}.$$

For $s \geq t$, take $E\left[y\left(\tau\right)|\,z=s\right]=LB\left(\tau,s,t\right)$.

For $s<t$, take,

$$E\left[y\left(\tau\right)|\,z=s\right]=\min\left\{E\left[y|\,z=s\right],LB\left(\tau,t,t\right)\right\}.$$

$E\left[y\left(\tau\right)|\,z=s\right]$ satisfies the concave-MTR, since its graph is the boundary of the convex hull, and weakly increases in τ. $E\left[y\left(\tau\right)|\,z=s\right]$ also satisfies MTS, due to minimizing over the set of $\{\widetilde{s}\,|\widetilde{s}\geq s\}$. ($\{\widetilde{s}\,|\widetilde{s}\geq s_1\} \supseteq \{\widetilde{s}\,|\widetilde{s}\geq s_2\}$ for $s_1 \leq s_2$, and $E\left[y|\,z=s\right]$ weakly increases in s.) These functions attain the lower bounds in the Proposition for the following reasons. First, for $\widetilde{s} \geq s$,

$$\begin{aligned} & E\left[y|\,z=s'^{*}\left(\widetilde{s},t\right)\right]+\frac{E\left[y|\,z=s'^{*}\left(\widetilde{s},t\right)\right]-E\left[y|\,z=u^{*}\left(\widetilde{s},t\right)\right]}{s'^{*}\left(\widetilde{s},t\right)-u^{*}\left(\widetilde{s},t\right)}\left[t-s'^{*}\left(\widetilde{s},t\right)\right] \\ \geq\ & E\left[y|\,z=s'^{*}\left(s,t\right)\right]+\frac{E\left[y|\,z=s'^{*}\left(s,t\right)\right]-E\left[y|\,z=u^{*}\left(s,t\right)\right]}{s'^{*}\left(s,t\right)-u^{*}\left(s,t\right)}\left[t-s'^{*}\left(s,t\right)\right], \end{aligned}$$

because of the definitions of $s'^*(\widetilde{s},t)$ and $u^*(\widetilde{s},t)$. Second,

$$E[y|\, z=\widetilde{s}] \geq E[y|\, z=s'^*(\widetilde{s},t)] + \frac{E[y|\, z=s'^*(\widetilde{s},t)] - E[y|\, z=u^*(\widetilde{s},t)]}{s'^*(\widetilde{s},t) - u^*(\widetilde{s},t)}[t - s'^*(\widetilde{s},t)];$$

since $E[y|\, z=\widetilde{s}]$ is weakly increasing in $\widetilde{s}$, $\widetilde{s} \geq s'^*(\widetilde{s},t) > u^*(\widetilde{s},t)$ and $t \leq s'^*(\widetilde{s},t)$.

Therefore, for $s \geq t$,

$$\begin{aligned} E[y(t)|\, z=s] &= LB(t,s,t) \\ &= E[y|\, z=s'^*(s,t)] + \frac{E[y|\, z=s'^*(s,t)] - E[y|\, z=u^*(s,t)]}{s'^*(s,t) - u^*(s,t)}[t - s'^*(s,t)], \end{aligned}$$

and for $s < t$,

$$E[y(t)|\, z=s] = E[y|\, z=s]$$

by the definitions of $s'^*(t,t)$ and $u^*(t,t)$.

It suffices to show that $E[y(s)|\, z=s] = E[y|\, z=s]$.

The definitions of $s'^*(\widetilde{s},t)$ and $u^*(\widetilde{s},t)$ and the fact that for $\widetilde{s} \geq s \geq t$,

$$\begin{aligned} & E[y|\, z=s'^*(\widetilde{s},t)] + \frac{E[y|\, z=s'^*(\widetilde{s},t)] - E[y|\, z=u^*(\widetilde{s},t)]}{s'^*(\widetilde{s},t) - u^*(\widetilde{s},t)}[t - s'^*(\widetilde{s},t)] \\ = \;& E[y|\, z=u^*(\widetilde{s},t)] + \frac{E[y|\, z=s'^*(\widetilde{s},t)] - E[y|\, z=u^*(\widetilde{s},t)]}{s'^*(\widetilde{s},t) - u^*(\widetilde{s},t)}[t - u^*(\widetilde{s},t)] \\ \geq \;& E[y|\, z=u^*(\widetilde{s},t)] + \frac{E[y|\, z=s] - E[y|\, z=u^*(\widetilde{s},t)]}{s - u^*(\widetilde{s},t)}[t - u^*(\widetilde{s},t)] \end{aligned}$$

imply that for $\widetilde{s} \geq s \geq t$,

$$\frac{E[y|\, z=s'^*(\widetilde{s},t)] - E[y|\, z=u^*(\widetilde{s},t)]}{s'^*(\widetilde{s},t) - u^*(\widetilde{s},t)} \geq \frac{E[y|\, z=s] - E[y|\, z=u^*(\widetilde{s},t)]}{s - u^*(\widetilde{s},t)} \geq 0.$$

Thus,

$$\begin{aligned} & E[y|\, z=s'^*(\widetilde{s},t)] + \frac{E[y|\, z=s'^*(\widetilde{s},t)] - E[y|\, z=u^*(\widetilde{s},t)]}{s'^*(\widetilde{s},t) - u^*(\widetilde{s},t)}[s - s'^*(\widetilde{s},t)] \\ = \;& E[y|\, z=u^*(\widetilde{s},t)] + \frac{E[y|\, z=s'^*(\widetilde{s},t)] - E[y|\, z=u^*(\widetilde{s},t)]}{s'^*(\widetilde{s},t) - u^*(\widetilde{s},t)}[s - u^*(\widetilde{s},t)] \\ \geq \;& E[y|\, z=u^*(\widetilde{s},t)] + \frac{E[y|\, z=s] - E[y|\, z=u^*(\widetilde{s},t)]}{s - u^*(\widetilde{s},t)}[s - u^*(\widetilde{s},t)] \\ = \;& E[y|\, z=s] \quad \text{for } \widetilde{s} \geq s \geq t. \end{aligned}$$

Hence, $E[y(s)|\, z=s] = LB(s,s,t) = E[y|\, z=s]$ for $s \geq t$. Similarly, $LB(s,t,t) = E[y|\, z=t]$. Therefore, $E[y(s)|\, z=s] = E[y|\, z=s]$ for $s<t$.

6 Appendix: Proof of the Concave-MTR and MTS of $LBAT_1(\tau,s,t_2)$ and $LBAT_2(\tau,s,t_2)$ in Proposition 2

(1) $LBAT_1(\tau,s,t_2)$:

(i) The concave-MTR of $LBAT_1(\tau,s,t_2)$:

By the definitions of $UB(s,t_2)$ and $v(s)$; and Equation (11);

$$0 \leq \frac{UB(s,t_2) - E[y|\, z=s]}{t_2 - s} \leq \frac{E[y|\, z=s] - E[y|\, z=v(s)]}{s - v(s)}. \tag{12}$$

By the definition of $v^m(s)$ and MTS, for $m = 1,2,...,\ M(s)$,

$$0 \leq \frac{E[y|\, z=v^{m-1}(s)] - E[y|\, z=v^m(s)]}{v^{m-1}(s) - v^m(s)} \leq \frac{E[y|\, z=v^{m-1}(s)] - E[y|\, z=v^{m+1}(s)]}{v^{m-1}(s) - v^{m+1}(s)}. \tag{13}$$

Show for $m = 1,2,...,\ M(s)$,

$$\frac{E[y|\, z=v^{m-1}(s)] - E[y|\, z=v^{m+1}(s)]}{v^{m-1}(s) - v^{m+1}(s)} \leq \frac{E[y|\, z=v^m(s)] - E[y|\, z=v^{m+1}(s)]}{v^m(s) - v^{m+1}(s)}. \tag{14}$$

Suppose,

$$\frac{E[y|\, z=v^{m-1}(s)] - E[y|\, z=v^{m+1}(s)]}{v^{m-1}(s) - v^{m+1}(s)} > \frac{E[y|\, z=v^m(s)] - E[y|\, z=v^{m+1}(s)]}{v^m(s) - v^{m+1}(s)} \geq 0.$$

As one looks at a triangle whose vertices are $(v^{m-1}(s), E[y|\, z=v^{m-1}(s)])$, $(v^m(s), E[y|\, z=v^m(s)])$ and $(v^{m+1}(s), E[y|\, z=v^{m+1}(s)])$,

$$\frac{E[y|\, z=v^{m-1}(s)] - E[y|\, z=v^{m+1}(s)]}{v^{m-1}(s) - v^{m+1}(s)} < \frac{E[y|\, z=v^{m-1}(s)] - E[y|\, z=v^m(s)]}{v^{m-1}(s) - v^m(s)}.$$

This contradicts the definition of $v^m(s)$. Therefore, Equation (13) is shown.

Hence, by Equations (12), (13), and (14), for $m = 1,2,...,\ M(s)$,

$$\begin{aligned} 0 &\leq \frac{UB(s,t_2) - E[y|\, z=s]}{t_2 - s} \\ &\leq \frac{E[y|\, z=v^{m-1}(s)] - E[y|\, z=v^m(s)]}{v^{m-1}(s) - v^m(s)} \leq \frac{E[y|\, z=v^m(s)] - E[y|\, z=v^{m+1}(s)]}{v^m(s) - v^{m+1}(s)}. \end{aligned}$$

That is, $LBAT_1(\tau, s, t_2)$ is concave-MTR in $\tau \in T$.

(ii) MTS of $LBAT_1(\tau, s, t_2)$:

The subgraph of $LBAT_1(\tau, s, t_2)$ includes all observations whose treatments are not greater than s $((k, E[y|\, z = k])$ for $k \leq s)$. Therefore, this includes the segments between any two among those observations. Hence, $LBAT_1(\tau, s, t_2) \geq LBAT_1(\tau, k, t_2)$ for $k \leq s$. That is, $LBAT_1(\tau, s, t_2)$ is MTS.

(2) The concave-MTR and MTS of $LBAT_2(\tau, s, t_2)$: Similarly, it is shown that $LBAT_2(\tau, s, t_2)$ is concave-MTR and MTS. (When proving the MTS, we utilize the property that $E[y|\, z = s] \geq E[y|\, z = k]$ for $(t_2 \leq)\, k < s$.)

References

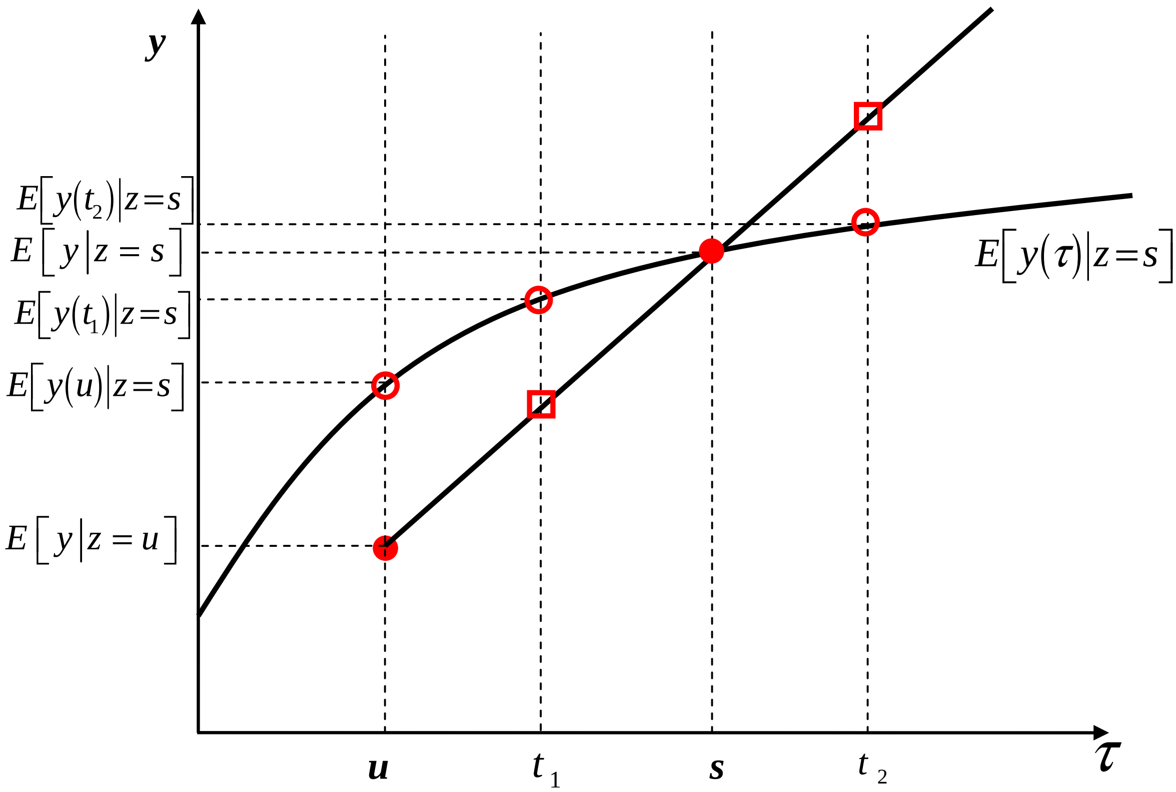
[1] Belzil, C., and J. Hansen, 2002, Unobserved Ability and the Return to Schooling. *Econometrica* **70,** 2075-91.

[2] Card, D, 1999, Casual Effect of Education on Earnings, In *Handbook of Labor Economics, Vol. 3A.* Ashenfelter, O.C. and D. Card (eds.). North-Holland, Amsterdam.

[3] Manski, C.F., 1997, Monotone Treatment Response. *Econometrica* **65**, 1311-34.

[4] Manski, C.F. and J. Pepper, 2000, Monotone Instrumental Variables: With an Application to the Returns to Schooling. *Econometrica* **68**, 997-1010.

Figure 1



y
$E\left[y(t_2)|z=s\right]$
$E\left[y|z=s\right]$
$E\left[y(t_1)|z=s\right]$
$E\left[y(u)|z=s\right]$
$E\left[y|z=u\right]$
$E\left[y(\tau)|z=s\right]$
u
t_1
s
t_2
τ

Table 1: Mean Log(Wages) and Distribution of Schooling: Wave 2000, NLSY

Table 1(a)

z	E(y | z)	P(z)	Sample Size
7	2.228	0.006	7
8	2.541	0.016	20
9	2.449	0.018	23
10	2.515	0.017	21
11	2.637	0.019	24
12	2.715	0.407	508
13	2.985	0.071	88
14	2.979	0.087	109
15	3.062	0.037	46
16	3.244	0.180	225
17	3.266	0.037	46
18	3.381	0.051	63
19	3.358	0.025	31
20	3.368	0.029	36
Total		1	1247

Table 1 (b)

z	E(y | z)	P(z)	Sample Size
9	2.449	0.019	23
10	2.515	0.017	21
11	2.637	0.020	24
12	2.715	0.416	508
13	2.985	0.072	88
14	2.979	0.089	109
15	3.062	0.038	46
16	3.244	0.184	225
17	3.266	0.038	46
18	3.381	0.052	63
19	3.358	0.025	31
20	3.368	0.030	36
Total		1	1220

Table 2: Upper and Lower Bounds on E[*y*(*t*)]

This Paper

	Lower Bound on E[*y*(*t*)]			Upper Bound on E[*y*(*t*)]		
t	Estimate	0.05 Bootstrap Quantile	0.95 Bootstrap Quantile	Estimate	0.05 Bootstrap Quantile	0.95 Bootstrap Quantile
9	2.449	2.299	2.618	2.949	2.925	2.975
10	2.562	2.496	2.713	2.951	2.925	2.977
11	2.673	2.637	2.774	2.953	2.929	2.980
12	2.777	2.749	2.828	2.957	2.931	2.983
13	2.847	2.813	2.896	2.993	2.949	3.021
14	2.877	2.852	2.910	3.029	2.956	3.065
15	2.906	2.884	2.937	3.079	2.969	3.112
16	2.935	2.912	2.964	3.123	2.983	3.156
17	2.943	2.920	2.972	3.171	2.996	3.200
18	2.950	2.926	2.978	3.218	3.014	3.247
19	2.949	2.925	2.976	3.249	3.019	3.282
20	2.949	2.925	2.975	3.282	3.026	3.301

Manski and Pepper (2000)

	Lower Bound on E[*y*(*t*)]			Upper Bound on E[*y*(*t*)]		
t	Estimate	0.05 Bootstrap Quantile	0.95 Bootstrap Quantile	Estimate	0.05 Bootstrap Quantile	0.95 Bootstrap Quantile
9	2.449	2.299	2.618	2.949	2.925	2.975
10	2.513	2.302	2.693	2.951	2.925	2.977
11	2.631	2.514	2.749	2.955	2.930	2.981
12	2.705	2.675	2.738	2.959	2.934	2.985
13	2.847	2.804	2.896	3.087	3.041	3.136
14	2.845	2.806	2.884	3.084	3.032	3.135
15	2.875	2.832	2.920	3.136	3.063	3.209
16	2.935	2.908	2.962	3.259	3.214	3.305
17	2.938	2.908	2.968	3.277	3.161	3.392
18	2.950	2.924	2.977	3.380	3.276	3.480
19	2.949	2.923	2.975	3.358	3.200	3.514
20	2.949	2.925	2.975	3.368	3.222	3.523

Manski (1997)

	Lower Bound on E[*y*(*t*)]			Upper Bound on E[*y*(*t*)]		
t	Estimate	0.05 Bootstrap Quantile	0.95 Bootstrap Quantile	Estimate	0.05 Bootstrap Quantile	0.95 Bootstrap Quantile
9	0.594	0.578	0.610	2.949	2.925	2.975
10	1.141	1.120	1.162	2.996	2.968	3.024
11	1.667	1.641	1.692	3.063	3.025	3.104
12	2.175	2.143	2.206	3.148	3.096	3.205
13	2.401	2.374	2.429	3.516	3.450	3.587
14	2.584	2.560	2.609	3.927	3.843	4.015
15	2.722	2.700	2.746	4.382	4.282	4.488
16	2.844	2.820	2.868	4.854	4.734	4.978
17	2.892	2.868	2.916	5.400	5.265	5.542
18	2.925	2.900	2.950	5.960	5.808	6.118
19	2.941	2.916	2.967	6.538	6.369	6.713
20	2.949	2.925	2.975	7.123	6.938	7.315

Table 3: Upper Bounds on Returns to Schooling

		Upper Bounds on $\Delta(s, t)$					
		Our Estimate		Manski and Pepper's Estimate		Manski's Estimate	
		Estimate	0.95 Bootstrap Quantile	Estimate	0.95 Bootstrap Quantile	Estimate	0.95 Bootstrap Quantile
s	***t***	(1)	(2)	(3)	(4)	(5)	(6)
9	10	0.502	0.655	0.502	0.655	1.498	1.512
10	11	0.252	0.316	0.442	0.654	1.021	1.035
11	12	0.165	0.199	0.328	0.449	0.787	0.801
12	13	0.133	0.159	0.382	0.439	0.703	0.717
13	14	0.103	0.121	0.236	0.304	0.654	0.669
14	15	0.093	0.102	0.292	0.373	0.626	0.641
15	16	0.079	0.087	0.384	0.447	0.607	0.622
16	17	0.061	0.072	0.343	0.461	0.600	0.616
17	18	0.056	0.065	0.442	0.545	0.596	0.612
18	19	0.048	0.055	0.408	0.566	0.594	0.610
19	20	0.045	0.047	0.419	0.574	0.594	0.610
12	16	0.346	0.378	0.554	0.609	2.427	2.489
Average Effect		0.087	0.094	0.138	0.152	0.607	0.622