TCER Working Paper Series

CONCAVE-MONOTONE TREATMENT RESPONSE AND MONOTONE TREATMENT SELECTION: WITH AN APPLICATION TO THE RETURNS TO SCHOOLING

Tsunao Okumura
Emiko Usui

April 2011

Working Paper E-33
http://tcer.or.jp/wp/pdf/e33.pdf

TOKYO CENTER FOR ECONOMIC RESEARCH
1-7-10-703 Iidabashi, Chiyoda-ku, Tokyo 102-0072, Japan

Abstract

This paper identifies sharp bounds on the mean treatment response and average treatment effect under the assumptions of both concave monotone treatment response (concave-MTR) and monotone treatment selection (MTS). We use our bounds and the US National Longitudinal Survey of Youth to estimate mean returns to schooling. Our upper-bound estimates are substantially smaller than (1) estimates using only the concave-MTR assumption of Manski (1997) and (2) estimates using only the MTR and MTS assumptions of Manski and Pepper (2000). They fall in the lower range of the point estimates given in previous studies that assume linear wage functions. This is because ability bias is corrected by assuming MTS when the functions are close to linear. Our results therefore imply that higher returns reported in previous studies are likely to be overestimated.

Tsunao Okumura
TCER
and
Yokohama National University
International Graduate School of Social Sciences
79-4 Tokiwadai Hodogayaku Yokohama, 240-8501
okumura@ynu.ac.jp

Emiko Usui
Nagoya University
Graduate School of Economics
Nagoya, 464-8601, Japan
usui@soec.nagoya-u.ac.jp

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

TSUNAO OKUMURA[†]
Yokohama National University

EMIKO USUI[‡]
Nagoya University and IZA

Abstract

This paper identifies sharp bounds on the mean treatment response and average treatment effect under the assumptions of both concave monotone treatment response (concave-MTR) and monotone treatment selection (MTS). We use our bounds and the US National Longitudinal Survey of Youth to estimate mean returns to schooling. Our upper-bound estimates are substantially smaller than (1) estimates using only the concave-MTR assumption of Manski (1997) and (2) estimates using only the MTR and MTS assumptions of Manski and Pepper (2000). They fall in the lower range of the point estimates given in previous studies that assume linear wage functions. This is because ability bias is corrected by assuming MTS when the functions are close to linear. Our results therefore imply that higher returns reported in previous studies are likely to be overestimated.

JEL: C14, J24

Keywords: Nonparametric Methods, Partial Identification, Sharp Bounds, Treatment Response, Returns to Schooling

---

*The authors thank Hidehiko Ichimura, Brendan Kline, Elie Tamer, Edward Vytlacil, and especially Charles Manski for their helpful comments. The authors also thank participants at the annual meetings of the Econometric Society, the European Association of Labour Economists, the Midwest Econometrics Group, the Symposium on Econometric Theory and Applications (SETA), and the Western Economic Association; and in seminars at Hitotsubashi University, Kyoto University, Northwestern University, and the University of Tokyo. This research is supported by JSPS grant 21530167, the Kikawada Foundation, and TCER.

[†]Corresponding author: Tsunao Okumura, International Graduate School of Social Sciences, Yokohama National University, 79-4 Tokiwadai Hodogayaku, Yokohama 240-8501, 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 examines the identifying power of the mean treatment response and average 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). We are motivated by the fact that either assumption alone produces bounds that are too wide to have sufficient identifying power for many purposes. We then apply this nonparametric method to estimate the returns to schooling and thus assess the validity of point estimates reported in existing parametric studies.

Manski (1997) studies sharp bounds on the mean treatment response when the response functions are assumed either to satisfy monotone treatment response (MTR) or concave-monotone treatment response (concave-MTR). To enhance the identifying power on the bounds, Manski and Pepper (2000) combine the 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). However, they are still so large that they contain almost all the point estimates of the returns to schooling in the existing empirical literature.

In this paper, we add the assumption of concavity to the assumptions of MTS and MTR. Concavity is a natural assumption, because diminishing marginal returns are commonly assumed in economic analysis. We explore how including this assumption tightens the sharp bounds on the mean treatment response and the average treatment effect.

Using the 2000 wave of the US National Longitudinal Survey of Youth (NLSY), we implement our bounds to estimate returns to schooling. Our sharp upper-bound estimates of the year-by-year returns to schooling are in the range of 0.032 to 0.254. Our estimates are only 5 to 25 percent of Manski's estimates and 7 to 57 percent of Manski and Pepper's estimates. Specifically, our upper-bound estimates on college education are in the range of 0.076 to 0.136 (0.091 to 0.152 for bias-corrected estimates) for local (year-by-year) returns and 0.083 (0.096 for the bias-corrected estimate) for the four-year average. Our upper-bound estimates are therefore substantially smaller than either the estimates using only the concave-MTR as-

[1] Manski and Pepper (2000) introduce the monotone instrumental variable (MIV) assumption, which weakens the instrumental variable assumption by replacing mean independence of outcomes and instruments with 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 asserts that persons who choose more schooling have weakly higher mean wage functions than do those who choose less schooling.

sumption of Manski (1997) or the estimates using the MTR and MTS assumptions of Manski and Pepper (2000). The point estimates of the returns to schooling in the existing literature (e.g., Card (1999)) range between 0.052 and 0.132. Our upper-bound estimates on college education are thus also smaller than many previous point estimates.

To illustrate, Figure 1 depicts the difference among the three sets of bounds in Manski (1997), Manski and Pepper (2000), and this paper. In Figure 1, each individual $j \in J$ has a response function $y_j(\cdot)$ which satisfies the concave-MTR and MTS assumptions. Also, t is a treatment, z_j is j's realized treatment, $y_j = y_j(z_j)$ is j's realized outcome, and $y_j(t)$, $t \neq z_j$, are latent outcomes. Therefore, $E[y(t)|\, z = s]$, $t \neq s$, is the conditional mean of latent outcomes which an individual who chooses treatment s would obtain if he were to choose treatment t. This conditional mean of latent outcome, $E[y(t)|\, z = s]$, is Point A when $t \leq s$, and it is Point A' when $t > s$ (denoted by t' in Figure 1). Manski's lower bound on $E[y(t)|\, z = s]$ for $t \leq s$ is Point B, which is determined as the value of the function describing the straight line joining $(s, E[y|\, z = s])$ (Point C) and the origin (Point O), evaluated at t. Manski and Pepper's lower bound is $E[y|\, z = t]$ (Point D).

In contrast, our lower bound for $t \leq s$ is Point E, which is determined by first collecting all of the functions describing straight lines joining $(s, E[y|\, z = s])$ (Point C) and $(v, E[y|\, z = v])$, for any realized treatment $v \leq t$, and also the straight line which joins $(s, E[y|\, z = s])$ (Point C) and the origin (Point O); and then taking the greatest value of the evaluation of these functions at t. In Figure 1, v takes the values u and t. As a result, when we take the greatest value of Point E (on the line joining Point C and $(u, E[y|\, z = u])$ (Point F)), Point B (on the line joining Points C and O), and Point D (on the line joining Points C and D), our lower bound is Point E. Therefore, our lower bound (Point E) on $E[y(t)|\, z = s]$ for $t \leq s$ is not smaller than the lower bounds in Manski (1997) and Manski and Pepper (2000), namely, Point B and Point D, respectively. Similarly, our upper bound (Point E') on $E[y(t)|\, z = s]$ for $t > s$ is not greater than the upper bounds of Manski (1997) and Manski and Pepper (2000), namely, Point B' and Point D', respectively. Using the law of iterated expectation, our bounds on the mean treatment response $(E[y(t)])$ and the average treatment effect $(E[y(t_2)] - E[y(t_1)]$ for $t_1 < t_2)$ are narrower than or equal to the bounds in Manski (1997) and Manski and Pepper (2000).

Many empirical studies estimate returns to schooling by utilizing the techniques of ordinary least squares (OLS) and instrumental variables (IV). In his survey, Card (1999) shows that point estimates on returns to schooling in previous studies using US data are in the range of 0.052 to 0.132. His study raises two serious concerns about the credibility of these point estimates. First, almost all previous studies assume that the log-wage regression function is linear in years of schooling. However, the assumption that each additional year of schooling has the same proportional effect on log earnings, despite the heterogeneous components of schooling choices, is debatable. Canonical human capital models, as in the work of Card (1999), assume that the log earnings function is concave-increasing in schooling. Second, since years of schooling are considered to be positively correlated with unobserved abilities (because of the ability bias), it is more appropriate to utilize the IV technique. However, the validity of the instrumental variables used in applications has often been questioned. Indeed, IV estimates tend to be greater than OLS estimates, despite the predictions made by the ability-bias hypothesis. As a result, Card (1999) argues that point estimates on returns to schooling in the previous literature are biased upward.

In contrast to these previous studies, the concave-MTR assumption allows for flexible and concave-increasing wage functions. Moreover, the MTS assumption corrects the ability bias, in the sense of the mean-monotonicity of wages and schooling. Estimates of the upper bounds on the returns to schooling under these assumptions are smaller than many point estimates in previous research. This paper shows that our upper-bound on the average treatment effect (i.e., the returns to schooling) is obtained when the conditional-mean response functions (i.e., the log-wage functions) of individuals are the upper envelope (or upper boundary of the convex hull) of the conditional means of realized outcomes. Thus, the log-wage functions that attain the upper-bound estimates are concave but close to linear, and have more gentle slopes than the linear log-wage functions in previous studies; consequently, our upper-bound estimates are smaller than many point estimates. Our results therefore imply that the higher returns reported in previous studies are likely to be overestimated.

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

2 Concave-Monotone Treatment Response and Monotone Treatment Selection

We employ the same setup as Manski (1997) and Manski and Pepper (2000). 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 $y_j(t) \in Y$. Each 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. In combining the distribution of a random sample (z, y) with prior information, we intend to identify the mean treatment response $E[y(t)]$ and the average treatment effect, $E[y(t_2)] - E[y(t_1)]$ for $t_1 < t_2$.

Manski (1997) makes the MTR assumption as follows:
Let T be an ordered set, and t_1 and t_2 be elements of T. For each $j \in J$,

$$t_2 \geq t_1 \Longrightarrow y_j(t_2) \geq y_j(t_1). \tag{1}$$

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

$$\begin{aligned} &\sum_{s \leq t} E[y|\, z = s]\, P(z = s) + y_0 P(z > t) \\ \leq\quad & E[y(t)] \leq \sum_{s \geq t} E[y|\, z = s]\, P(z = s) + y_1 P(z < t), \end{aligned} \tag{2}$$

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), and when $T = [0, \lambda]$ for some $\lambda \in (0, \infty]$ and $Y = [0, \infty]$:

$$\begin{aligned} &\sum_{s < t} E[y|\, z = s]\, P(z = s) + E\left[\left.\frac{y}{z}t \right|\, z \geq t\right] P(z \geq t) \\ \leq\quad & E[y(t)] \leq \sum_{s > t} E[y|\, z = s]\, P(z = s) + E\left[\left.\frac{y}{z}t \right|\, z \leq t\right] P(z \leq t). \end{aligned} \tag{3}$$

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

$$t_2 \geq t_1 \Longrightarrow E[y(t)|\, z = t_2] \geq E[y(t)|\, z = t_1]. \tag{4}$$

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

$$\sum_{s<t} E[y|\, z=s]\, P(z=s) + E[y|\, z=t]\, P(z \geq t) \tag{5}$$

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

In contrast, we assume both concave-MTR and MTS. The following proposition demonstrates the sharp bounds on $E[y(t)]$ under the assumption of both concave-MTR and MTS. The basic idea is as follows: Refer to Figure 1. The MTS assumption implies that, for $u < s$, $E[y|\, z=u]$ (Point F) $\leq E[y(u)|\, z=s]$ (Point G). Also, $E[y(\tau)|\, z=s]$ is concave-MTR in $\tau \in T$. Thus, when $t \leq s$, for $u \leq t$, the value (Point E) of the function describing the straight line traversing $(u, E[y|\, z=u])$ (Point F) and $(s, E[y|\, z=s])$ (Point C), evaluated at t, is a lower bound on $E[y(t)|\, z=s]$ (Point A). Given $(s, E[y|\, z=s])$ (Point C), these lines are drawn for all realized points of $(u, E[y|\, z=u])$ for $u \leq t$ and the origin (Point O) because $Y = [0, \infty]$. The values of these functions evaluated at t (Points B, D, and E) are all lower bounds on $E[y(t)|\, z=s]$ for $t \leq s$ (Point A). These include the lower bound of Manski (1997), i.e., Point B on the line joining Points C and O; and that of Manski and Pepper (2000) ($E[y|\, z=t]$, i.e., Point D on the line joining Points C and D). Our lower bound on $E[y(t)|\, z=s]$ for $t \leq s$ is the greatest among these lower bounds (i.e., Point E is the greatest among Points B, D, and E). Similarly, when $t > s$ (denoted by t' in Figure 1), for any $u < s$, the value (Point E') of the function describing the straight line traversing $(u, E[y|\, z=u])$ (Point F) and $(s, E[y|\, z=s])$ (Point C), evaluated at t, is an upper bound on $E[y(t)|\, z=s]$ (Point A'). Also, the value (Point B') of the function describing the straight line traversing the origin (Point O) and Point C, evaluated at t, is an upper bound on $E[y(t)|\, z=s]$ (Point A'), as is the value $E[y|\, z=t]$ (Point D') – the latter because of the MTS assumption. Our upper bound (Point E') on $E[y(t)|\, z=s]$ for $t > s$ is the smallest among these upper bounds (Points B', D', and E'); note that Point B' corresponds to the upper bound in Manski (1997), and Point D' to the upper bound in Manski and Pepper (2000). Using the law of iterated expectations, our bounds on $E[y(t)]$ are narrower than or equal to those in Manski (1997) and Manski and Pepper (2000).

Proposition 1 *Let T be ordered. Let $T = [0, \lambda]$ for some $\lambda \in (0, \infty]$ and $Y = [0, \infty]$. Assume that $y_j(\cdot)$, $j \in J$, satisfies the assumptions of concave-MTR and MTS. We extend*

the set of realized treatments and outcomes by including $(z, y) = (0, 0)$, when there is no realized treatment of zero. We then obtain the following three results:
(1) For $(t, s, s', u) \in T^4$,

$$\begin{aligned}
& \sum_{s<t} E\left[y|\, z=s\right] P\left(z=s\right) \\
& + \sum_{s\geq t} \max_{\{s'|s\geq s'\geq t\}} \left(E\left[y|\, z=s'\right] + \max_{\{u|u<t\}} \left\{ \frac{E\left[y|\, z=s'\right] - E\left[y|\, z=u\right]}{s'-u} \left(t-s'\right) \right\} \right) P\left(z=s\right) \\
\leq \; & E\left[y\left(t\right)\right] \qquad (6) \\
\leq \; & \sum_{s>t} E\left[y|\, z=s\right] P\left(z=s\right) \\
& + \sum_{s\leq t} \min_{\{s'|s\leq s'\leq t\}} \left\{ E\left[y|\, z=s'\right] + \min_{\{u|u<s'\}} \frac{E\left[y|\, z=s'\right] - E\left[y|\, z=u\right]}{s'-u} \left(t-s'\right) \right\} P\left(z=s\right).
\end{aligned}$$

(2) These bounds are sharp. (3) These bounds are narrower than or equal to those using only the concave-MTR assumption of Manski (1997), as well as those using only the MTR and MTS assumptions of Manski and Pepper (2000).

In appendix A, we prove Proposition 1.

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

$$\begin{aligned}
& \sum_{s\geq t} \left\{ \max_{\{s'|s\geq s'\geq t\}} \left(E\left[y|\, z=s'\right] + \max_{\{u|u<t\}} \left\{ \frac{E\left[y|\, z=s'\right] - E\left[y|\, z=u\right]}{s'-u} \left(t-s'\right) \right\} \right) - E\left[y|\, z=t\right] \right\} \\
& \times P\left(z=s\right) \\
& + \sum_{s\leq t} \left(E\left[y|\, z=t\right] - \min_{\{s'|s\leq s'\leq t\}} \left\{ E\left[y|\, z=s'\right] + \min_{\{u|u<s'\}} \frac{E\left[y|\, z=s'\right] - E\left[y|\, z=u\right]}{s'-u} \left(t-s'\right) \right\} \right) \\
& \times P\left(z=s\right). \qquad (7)
\end{aligned}$$

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

The sharp bounds on the average treatment effects ($E\left[y\left(t_2\right)\right] - E\left[y\left(t_1\right)\right]$ for $t_1 < t_2$) are given in Proposition 2. Figure 2 provides an example: When $t_2 = t' \geq s$ and $t_1 < t_2$, let $UB(s, t_2)$ be the sharp upper bound on $E[y(t_2)|\, z = s]$ in Equation (6) (Point E'), and let $AT_1(t_1, s, t_2)$ be the value of the function describing the line joining Points O, F, E, C, and E', evaluated at t_1. When $t_1 < t_2 = t < s$, $E[y|\, z = s]$ (Point H) is the sharp upper bound on $E[y(t_2)|\, z = s]$ in Equation (6). Let $AT_2(t_1, s, t_2)$ be the value of the function describing the line joining Points O, F, and H, evaluated at t_1.[2] Then, our sharp upper bound on $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$ is $UB(s, t_2) - AT_1(t_1, s, t_2)$ for $t_2 = t' \geq s$; and it is $E[y|\, z = s] - AT_2(t_1, s, t_2)$ for $t_2 = t < s$. Using the law of iterated expectations, we obtain the sharp upper bound on the average treatment effect.

Proposition 2 *Let T be ordered. Let $T = [0, \lambda]$ for some $\lambda \in (0, \infty]$ and $Y = [0, \infty]$. Assume that $y_j(\cdot)$, $j \in J$, satisfies the assumptions of concave-MTR and MTS. Suppose (t, s, s', u, v, τ, t_1, t_2) $\in T^8$. We extend the set of realized treatments and outcomes by including $(z, y) = (0, 0)$, when there is no realized treatment of zero. We then define the following functions:*
(1) For $t \geq z = s$:

$$v^m(s) = \arg \min_{\{v | v < v^{m-1}(s)\}} \frac{E[y|\, z = v^{m-1}(s)] - E[y|\, z = v]}{v^{m-1}(s) - v}, \quad \text{for } m = 1, 2, ..., \; M(s), \tag{8}$$

where $v^0(s) = s$ and $M(s)$ satisfies $v^{M(s)}(s) = 0$.
(1.1) For $s \leq \tau \leq t$,

$$AT_1(\tau, s, t) = \begin{cases} E[y|\, z = s] + \frac{UB(s,t) - E[y|z=s]}{t-s}(\tau - s) & \text{if } \; s < t, \\ E[y|\, z = s] & \text{if } \; s = t, \end{cases} \tag{9}$$

where

$$UB(s, t) = \min_{\{s' | s \leq s' \leq t\}} \left\{ E[y|\, z = s'] + \min_{\{u | u < s'\}} \frac{E[y|\, z = s'] - E[y|\, z = u]}{s' - u}(t - s') \right\}. \tag{10}$$

(1.2) For $v^m(s) \leq \tau < v^{m-1}(s)$ $\;(m = 1, 2, ..., \; M(s))$,

$$AT_1(\tau, s, t) = E\left[y|\, z = v^{m-1}(s)\right] + \frac{E[y|\, z = v^{m-1}(s)] - E[y|\, z = v^m(s)]}{v^{m-1}(s) - v^m(s)} \left[\tau - v^{m-1}(s)\right]. \tag{11}$$

(1.3) For $\tau > t$,

$$AT_1(\tau, s, t) = UB(s, t). \tag{12}$$

[2] Proposition 2 provides precise definitions for the quantities $UB(s, t_2)$, $AT_1(t_1, s, t_2)$, and $AT_2(t_1, s, t_2)$.

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

$$v(s,t) = \arg\min_{\{v|v<t\}} \frac{E[y|z=s] - E[y|z=v]}{t-v}, \tag{13}$$

$$v^m(s,t) = \arg\min_{\{v|v<v^{m-1}(s,t)\}} \frac{E[y|z=v^{m-1}(s,t)] - E[y|z=v]}{v^{m-1}(s,t)-v} \quad \text{for } m = 2,3,...,M(s,t), \tag{14}$$

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

(2.1) For $v(s,t) \leq \tau \leq t < z = s$,

$$AT_2(\tau,s,t) = E[y|z=s] + \frac{E[y|z=s] - E[y|z=v(s,t)]}{t-v(s,t)}[\tau - t]. \tag{15}$$

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

$$\begin{aligned} & AT_2(\tau,s,t) \\ = \; & E\left[y|z=v^{m-1}(s,t)\right] + \frac{E[y|z=v^{m-1}(s,t)] - E[y|z=v^m(s,t)]}{v^{m-1}(s,t)-v^m(s,t)}\left[\tau - v^{m-1}(s,t)\right]. \end{aligned} \tag{16}$$

(2.3) For $\tau > t$,

$$AT_2(\tau,s,t) = E[y|z=s]. \tag{17}$$

Then, we obtain the following results:

(i) For $t_1 < t_2$,

$$\begin{aligned} 0 &\leq E[y(t_2)] - E[y(t_1)] \\ &\leq \sum_{s \leq t_2} [UB(s,t_2) - AT_1(t_1,s,t_2)] P(z=s) \\ &\quad + \sum_{s>t_2} \{E[y|z=s] - AT_2(t_1,s,t_2)\} P(z=s). \end{aligned} \tag{18}$$

These bounds are sharp.

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

In appendix B, we prove Proposition 2.

Proposition 2 shows that our sharp upper bound on $E[y(t_2)] - E[y(t_1)]$ is attained when $AT_1(\tau, s, t_2)$ is the mean response function of individuals whose realized treatment (s) is not greater than t_2 (i.e., $E[y(\tau)|\, z = s]$ for $s \le t_2$); and when $AT_2(\tau, s, t_2)$ is the mean response function of individuals whose realized treatment (s) is greater than t_2 (i.e., $E[y(\tau)|\, z = s]$ for $s > t_2$). The function $AT_1(\tau, s, t_2)$ is a function in τ describing the upper envelope of the points $(u, E[y|\, z = u])$ for all $u \le s$ and the point $(t_2, UB(s, t_2))$ (i.e., a function in τ describing the upper boundary of the convex hull for a set formed by these points). The function $AT_2(\tau, s, t_2)$ is a function in τ describing the upper envelope of the points $(u, E[y|\, z = u])$ for all $u \le t_2$ and the point $(t_2, E[y|\, z = s])$.[3] Therefore, the curves of the functions $AT_1(\tau, s, t_2)$ and $AT_2(\tau, s, t_2)$ are close to linear between t_1 and t_2. Furthermore, the value $\{UB(s, t_2) - AT_1(t_1, s, t_2)\}/(t_2 - t_1)$, which is an approximation of the slope of $AT_1(\tau, s, t_2)$, is most probably below the estimated coefficient obtained in many previous studies that have used linear regression (specifically, those in which the realized outcomes y are regressed on realized treatments s).[4]

[3]We will now provide intuitive explanations of why the conditional-mean response functions $AT_1(\tau, s, t_2)$ and $AT_2(\tau, s, t_2)$ satisfy the concave-MTR and MTS assumptions. The convex hull is the smallest convex set containing elements. Because a convex set has a concave upper boundary, these functions are concave-MTR. Because the convex hull for a set formed by $(u, E[y\,|z = u])$ for all $u \le s$ is included in the convex hull for a set formed by $(u, E[y\,|z = u])$ for all $u \le s'$ and $s < s'$; and because $UB(s, t_2)$ and $E[y\,|z = s]$ both weakly increase in s, it follows that $AT_k(\tau, s, t_2) \le AT_k(\tau, s', t_2)$ for $k = 1, 2$ and for $s < s'$. That is, the functions $AT_1(\tau, s, t_2)$ and $AT_2(\tau, s, t_2)$ satisfy the MTS assumption. To illustrate, Figure 2 shows that when $t_2 = t'$, the function $AT_1(\tau, s, t_2)$ is a function describing the line joining Points O, F, E, C, and E'; that is, a function describing the upper boundary of the convex hull for the set of Points O, F, D, C, and E'. When $t_2 = t$, the function $AT_2(\tau, s, t_2)$ is a function describing the line joining Points O, F, and H; that is, a function describing the upper boundary of the convex hull for the set of Points O, F, H, and D. These functions are both concave-MTR and MTS.

[4]Manski (1995) and Manski and Pepper (2009) study the identifying power of the homogeneous linear response (HLR) assumption and the exogenous treatment selection (ETS) assumption, which are imposed on the OLS regressions. The HLR assumption asserts that $y_j(t) = \beta t + \delta_j$, where δ_j is an unobserved covariate and β is a slope parameter that takes the same value for all j. The ETS assumption asserts that for $(t, t_1, t_2) \in T^3$, $E[y(t)|\, z = t_1] = E[y(t)|\, z = t_2]$. Under the HLR and ETS assumptions, Manski (1995) and Manski and Pepper (2009) identify $\beta = \{E[y|\, z = t_2] - E[y|\, z = t_1]\}/(t_2 - t_1)$. This β is not smaller than $\{UB(s, t_2) - AT_1(t_1, s, t_2)\}/(t_2 - t_1)$ because $UB(s, t_2) \le E[y|\, z = t_2]$ and $E[y|\, z = t_1] \le AT_1(t_1, s, t_2)$. Also, under the HLR and MTS assumptions, which are more restrictive than the concave-MTR and MTS assumptions, Manski and Pepper (2009) identify the sharp upper bound on β. This upper bound is not greater than $\{UB(s, t_2) - AT_1(t_1, s, t_2)\}/(t_2 - t_1)$. To illustrate, Figure 2 shows that when $t_2 = t'$ and $t_1 = t$, the quantity $\{UB(s, t_2) - AT_1(t_1, s, t_2)\}/(t_2 - t_1)$ is the slope of the line joining Points E and E'. The coefficient β under the HLR and ETS assumptions is the slope of the line joining Points D and D'. The upper bound on β under the HLR and MTS assumptions is the slope of the line joining Points D and F.

3 Estimation of Returns to Schooling

We use the 2000 wave of the US National Longitudinal Survey of Youth (NLSY), which is representative of the US noninstitutionalized civilian population who were between the ages of 14 and 22 in 1979. Like Manski and Pepper (2000), who use the 1994 wave of the NLSY, we use a random sample of white men who reported that they were full-time, year-round workers and not self-employed nor involved in a family business. The sample size is $1,221$. Their hourly rate of pay and realized years of schooling were observed. In our application to returns to schooling, z represents the realized years of schooling; $y_j(t)$ is the logarithm of the 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.[5] Therefore, the average treatment effect, $E[y(t_2)] - E[y(t_1)]$ for $t_1 < t_2$, is the expected return to completing t_2 years of schooling relative to t_1 years.

Table 1 shows estimates of $E(y|z)$ and $P(z)$ for all samples. Forty percent of the NLSY respondents have 12 years of schooling and 18 percent have 16 years of schooling. For the most part, estimates of $E(y|z)$ increase with z. However, we find three decreases of the estimates of $E[y|\, z = s_2]$ from the estimates of $E[y|\, z = s_1]$ for $s_1 < s_2$ when s_2 is equal to nine, ten, and nineteen years of schooling. The decreases conflict with the MTS and MTR assumptions. Like Manski and Pepper (2000), who also find four decreases in their data, we compute the uniform 95-percent confidence intervals for the estimates of $E(y|z)$ and find that the intervals contain everywhere monotone functions. Therefore, it appears that the MTS and MTR assumptions are consistent with the empirical evidence. However, the upper bounds in Equations (6) and (18) have the term $\min_{\{u|u<s'\}} \{E[y|\, z = s'] - E[y|\, z = u]\} \big/ (s' - u)$, which is negative when $E[y|\, z = s']$ is a decrease. The upper bounds in these Equations are likely to be underestimated, as compared to the case in which there is no decrease, because this negative term affects the upper bounds in or after schooling years which show decreases.[6]

Therefore, we deal with the problem associated with the decreases as follows. First, to

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

[6] When there is a decrease from $E[y|\, z = s_1]$ to $E[y|\, z = s_2]$ $(s_1 < s_2)$, for all s, t_1, and t_2 such that $s_1 \leq s \leq s_2$, $s \leq t_1$, $s_2 \leq t_2$, and $t_1 < t_2$, the terms of $UB(s, t_2) - AT_1(t_1, s, t_2)$ in Equation (18) are negative. Therefore, the decreases reduce the upper bounds on the returns to schooling not only for the schooling years of the decreases, but also for the subsequent years. The decreases in nine and ten years of schooling have a significant effect on estimates of the returns to schooling; however, the decrease in nineteen years of schooling has little effect. When we implement subsampling to derive confidence intervals and the bias-corrected estimates, we have decreases in the subsamples, with which we deal in the same way.

deal with the anomalous wage decrease for individuals with nine or ten years of schooling, compared to those with just eight, we exclude individuals with eight years of schooling. Because of the small sample size, we also exclude individuals with seven years of schooling, following Manski and Pepper (2000). When the samples of seven and eight years of schooling are included, the upper-bound estimates are smaller (see Footnote 12). Second, for the decreases from $E[y|\, z = s_1]$ to $E[y|\, z = s_2]$ (s_1 and s_2 are eighteen and nineteen years of schooling, respectively), we replace $(s_2, E[y|\, z = s_2])$ with $(s_2, E[y|\, z = s_1])$, so that the decreases are smoothed out.[7]

Table 2 reports estimates of the bounds on $E[y(t)]$ in Equation (6). The subsampling method is utilized to derive a 90 percent confidence interval, with the 0.05 subsampling quantile as the lower bound and the 0.95 subsampling quantile as the upper bound, as shown in Table 2.[8]

For comparison, Table 3 reports (i) the estimates of the bounds using only the MTR and MTS assumptions (i.e., the bounds in Manski and Pepper (2000)), and (ii) the estimates of the bounds using only the concave-MTR assumption (i.e., the bounds in Manski (1997)). Estimates of our bounds are much narrower than those of both Manski and Pepper (2000) and Manski (1997). In particular, by adding the assumption of concavity to the MTR and MTS assumptions, the widths of the estimated bounds on the mean treatment response $E[y(t)]$ are reduced to three-quarters when $t = 12$ (high-school completion) and to half when $t = 16$ (college completion). Also, compared to Manski and Pepper's (2000) estimates using only the MTR and MTS assumptions, our lower-bound estimates are larger than their estimates for lower years of schooling, while our upper-bound estimates are smaller than their estimates for higher years of schooling. Equation (7) 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 4 reports the estimates of the upper bounds on the average treatment effect, $E[y(t_2)] - E[y(t_1)]$ for $t_1 < t_2$ (denoted by $\Delta(t_1, t_2)$) in Equation (18), namely, the estimates of the returns to schooling. Our bound estimates and the 95 percent subsampling

[7] Alternatively, we replace $\{E[y|\, z = s'] - E[y|\, z = u]\} \,/\, (s' - u)$ in the upper bounds in Equations (6) and (18) with $Max(\{E[y|\, z = s'] - E[y|\, z = u]\} \,/\, (s' - u)\,,\, 0)$. This specification produces fairly similar estimates to the specification which smooths out decreases.

[8] As Politis, Romano, and Wolf (1999) and Chernozhukov, Hong, and Tamer (2007) indicate, even if the estimates of the bounds are asymptotically biased, the subsampling confidence intervals (reported in Tables 2, 3, 4, and 5) are still consistent.

quantile are listed in Columns 1 and 2, respectively. Table 5 presents the bound estimates and the confidence intervals using only the MTR and MTS assumptions of Manski and Pepper (2000), as well as those using only the concave-MTR assumption of Manski (1997). As Manski's (1997) upper-bound estimates also decrease with years of schooling, our upper-bound estimates decrease with years of schooling, because of the shared concave-MTR assumption. Our upper-bound estimates are in the range of 0.032 to 0.254 for year-by-year returns to schooling, except for $\Delta(9,10)$. Our upper-bound estimates are only 7 to 57 percent as large as those in Manski and Pepper (2000) and only 5 to 25 percent as large as those in Manski (1997). Upper-bound estimates on local returns to college education ($\Delta(12,13)$, $\Delta(13,14)$, $\Delta(14,15)$, and $\Delta(15,16)$) are between 0.076 and 0.136. Upper-bound estimates on $\Delta(12,16)$ imply that the completion of four years of college yields, at most, an increase of 0.331 in mean log wage relative to the completion of high school. This implies that the average yearly returns from completing four years of college is at most $\Delta(12,16)/4 = 0.083$.[9]

As Manski and Pepper (2009) point out, when the *max* and *min* operations are used in Equations (6) and (18), estimates of these bounds may suffer from finite-sample bias. This bias makes bound estimates narrower than the true bounds. Kreider and Pepper (2007) and Manski and Pepper (2009) propose a bias-corrected estimator which is based on the bootstrap method; however, we incorporate a bias-corrected estimator which is based on the subsampling method.[10] Columns 4 and 8 in Table 2 report the bias-corrected bound estimates of $E[y(t)]$.[11] Column 3 in Table 4 reports the bias-corrected upper-bound estimates on the average treatment effect. For the returns to college education, the bias-corrected estimates of the bounds are between 0.091 and 0.152 for local returns and 0.096 for the four-year average. Therefore, both the bias-corrected estimates and the 95 percent subsampling quantile of the bounds on the returns to schooling are still significantly smaller than the estimates of the

[9] The upper-bound estimates on the returns to college education using only the MTR and MTS assumptions are between 0.249 and 0.386 for local returns and 0.138 for the four-year average, while the estimates using only the concave-MTR assumption are between 0.608 and 0.705 for local returns and 0.608 for the four-year average.

[10] Let T_n be the analog estimate of the bound and let $E^*(T_{n,b})$ be the mean of the estimates using the subsampling distribution; then the subsampling bias-corrected estimator is $\left(1+\sqrt{b/n}\right)T_n - \sqrt{b/n}E^*(T_{n,b})$, where n and b are the size of the sample and subsample, respectively (see Politis, Romano, and Wolf (1999)).

[11] Lower-bound estimates are corrected downward at lower years of schooling, and upper-bound estimates are corrected upward at higher years of schooling, because when t is small (large), the lower (upper) bound in Equation (6) is dominated by the second term, which has the *max* (*min*) operations.

bounds in either Manski and Pepper (2000) or Manski (1997).[12]

Many empirical studies regress log wages on years of schooling by using OLS and IV techniques and report various point estimates of the returns to schooling. Card (1999) shows that these point estimates (using US data) are between 0.052 and 0.132. He also raises two serious concerns about the credibility of these high returns to schooling.[13] First, years of schooling may be positively correlated with unobserved abilities, which bias OLS estimates upward (the effect known as the ability bias). And although the IV technique is utilized often in the existing empirical literature, the estimates obtained using the IV technique tend to be greater than OLS estimates, despite the predictions made by the ability-bias hypothesis. Second, almost all previous studies assume linear log-wage functions on years of schooling. There is, however, little a priori reason to assume that log wage varies linearly with years of schooling.

To address these two problems, we apply the MTS and concave-MTR assumptions. The MTS assumption corrects the ability bias, in the sense of the mean-monotonicity of wages and schooling. The concave-MTR assumption is consistent with conventional theories of human capital accumulation. Estimates of the upper bounds on the returns to schooling under these assumptions are within the range of the point estimates of previous research. When our data and the OLS technique are used to regress log wages on years of schooling, the point estimate on returns to schooling is 0.101 (and the standard error is 0.0052). This estimate is greater than the four-year average and some local returns of the bias-corrected upper-bound estimates on college education in Table 4. Proposition 2 implies that the conditional-mean response functions that attain our upper-bound estimates are the upper envelope of the conditional means of realized outcomes. As a result, the upper-bound estimates are smaller than the point estimates.[14] Thus, our estimation results imply that higher point estimates on returns to schooling in previous studies, particularly those utilizing the IV technique, are

[12] When the samples of seven and eight years of schooling are included, the upper-bound estimates on the returns to schooling are in the range of 0.055 to 0.085 for the local returns to college education and 0.058 for the four-year average. The 95 percent subsampling quantile and the bias-corrected estimates are $0.063-0.112$ and $0.065-0.094$ for the local returns, and 0.071 and 0.066 for the four-year average.

[13] Manski and Pepper (2000) and Belzil and Hansen (2002) also question the validity of these point estimates.

[14] Footnote 4 explains that the quantity $UB(s,t_2) - AT_1(t_1,s,t_2)$ in Equation (18) is not greater than $(t_2 - t_1)$ multiplied by the estimated slope-parameter of a linear response function under the HLR and ETS assumptions. When t_2 is large, the upper bound in Equation (18) is dominated by the first term, which contains the terms of $UB(s,t_2) - AT_1(t_1,s,t_2)$. Therefore, when t_2 is large, the upper-bound estimates are smaller than the point estimates.

biased upward.

4 Conclusion

We identify sharp bounds on the mean treatment response and average treatment effect under the assumptions of both the concave-MTR and MTS. Bounds on the returns to schooling are estimated by utilizing our bounds and the NLSY data. Our upper-bound estimates of returns to college education are 0.076 to 0.136 (0.078 to 0.173 for the 95 percent subsampling quantile) for local returns and 0.083 (0.091) for the four-year average. Estimates of our bounds are therefore substantially tighter than estimates using only the MTR and MTS assumptions of Manski and Pepper (2000).

Our upper-bound estimates are also smaller than many point estimates reported in the previous literature. In addition, other researchers have claimed that returns to schooling have probably been overestimated as result of both the ability bias and misspecifications of the wage function. Our estimation results now provide evidence that casts further doubt on the validity of high average returns to schooling.

This approach can be applied to other important economic models that assume concave-MTR and MTS. For example, the concave-MTR and MTS assumptions are consistent with conventional production theory. Production theory often asserts that the output of a product increases with input. This assertion has dual interpretations. The concave-MTR assumption asserts that a production function weakly increases and marginal product weakly decreases with input. The MTS assumption asserts that firms which select greater levels of output have weakly greater average production functions than do those which select smaller levels of output.[15] Therefore, under the concave-MTR and MTS assumptions about the production process, the bound approach in this paper could be applied to reveal the average production function by estimating the average increase in a firm's production as input increases: $E\left[y\left(t_2\right)\right] - E\left[y\left(t_1\right)\right]$.

[15] Consider firm i's concave-MTR production function $y_i\left(t\right)$, where t is input, y_i is output, and $y_i\left(0\right) \geq 0$. It is assumed that for all firms i, realized input z_i is determined optimally where the marginal product equals the marginal cost, $y_i'\left(z_i\right) = r$ (where r is the cost of input and the price of the output is set to unity). If firm i's production function has higher productivity than firm j's, defined as $y_i'\left(t\right) \geq y_j'\left(t\right)$ for all t and $y_i\left(0\right) \geq y_j\left(0\right)$, then $z_i \geq z_j$ and $y_i\left(t\right) \geq y_j\left(t\right) \left(= \int_0^t y_j'\left(s\right) ds + y_j\left(0\right)\right)$. It is reasonable to assume that $E\left[y\left(t\right)\middle| z_i\right] \geq E\left[y\left(t\right)\middle| z_j\right]$ for $z_i \geq z_j$. In this case, $y_i\left(t\right)$ satisfies the concave-MTR and MTS assumptions.

5 Appendix

5.1 Appendix A: Proof of Proposition 1

Proof of Part (1):

For $u < s : E\left[y|\, z = s\right] = E\left[y\left(s\right)|\, z = s\right] \geq E\left[y\left(u\right)|\, z = s\right]$ by the MTR assumption.
$E\left[y\left(u\right)|\, z = s\right] \geq E\left[y\left(u\right)|\, z = u\right] = E\left[y|\, z = u\right]$ by the MTS assumption. Hence,

$$E\left[y|\, z = s\right] \geq E\left[y\left(u\right)|\, z = s\right] \geq E\left[y|\, z = u\right]. \tag{A.1}$$

Since $y_j\left(\tau\right)$ is concave-MTR in $\tau \in T$ for all $j \in J$, $E\left[y\left(\tau\right)|\, z = s\right]$ is concave-MTR in τ. Compare $E\left[y\left(t\right)|\, z = s\right]$ with the value of the function describing the straight line joining the points $\left(s, E\left[y|\, z = s\right]\right)$ and $\left(u, E\left[y|\, z = u\right]\right)$, evaluated at t.
Because Equation $(A.1)$ holds and $E\left[y\left(\tau\right)|\, z = s\right]$ is concave-MTR in τ, for $t \geq s > u$,

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

and for $s \geq t > u$,

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

Since Equation $(A.2)$ holds for any u which is smaller than s when s is not greater than t, then for $t \geq s$,

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

Similarly, since Equation $(A.3)$ holds for any u which is smaller than t when t is not greater than s, then for $t \leq s$,

$$E\left[y\left(t\right)|\, z = s\right] \geq E\left[y|\, z = s\right] + \max_{\{u|u<t\}} \left\{\frac{E\left[y|\, z = s\right] - E\left[y|\, z = u\right]}{s - u}\left(t - s\right)\right\}. \tag{A.5}$$

The MTS assumption implies that for all $s' \geq s$, $E\left[y\left(t\right)|\, z = s\right] \leq E\left[y\left(t\right)|\, z = s'\right]$. Also, for all $s' \leq t$, Equation $(A.4)$ can be applied to the upper bound on $E\left[y\left(t\right)|\, z = s'\right]$; for $s' \leq t$,

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

Therefore, for $t \geq s$,

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

Similarly, for $t \leq s$, by the MTS assumption and Equation $(A.5)$,

$$E\left[y\left(t\right)|\, z=s\right] \geq \max_{\{s'|s\geq s'\geq t\}}\left(E\left[y|\, z=s'\right]+\max_{\{u|u<t\}}\left\{\frac{E\left[y|\, z=s'\right]-E\left[y|\, z=u\right]}{s'-u}\left(t-s'\right)\right\}\right). \tag{A.7}$$

Applying Equations $(A.6)$ and $(A.7)$ to the law of iterated expectations yields the second terms of the upper and lower bounds, respectively, on $E\left[y\left(t\right)\right]$ in Equation (6).

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

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

and for $s < t$,

$$E\left[y\left(t\right)|\, z=s\right] \geq E\left[y|\, z=s\right]. \tag{A.9}$$

Applying Equations $(A.8)$ and $(A.9)$ to the law of iterated expectations yields the first terms of the upper and lower bounds, respectively, on $E\left[y\left(t\right)\right]$ in Equation (6).

These results thus yield the bounds on $E\left[y\left(t\right)\right]$ in Equation (6). *Q.E.D.*

Proof of Part (2):

To show that the bounds on $E\left[y\left(t\right)\right]$ in Equation (6) are sharp, it suffices to demonstrate (i) that there exists a set of functions of $y_j\left(\tau\right)$ for $\tau \in T$ that satisfy the concave-MTR and MTS assumptions and that attain the upper bound, and (ii) that there also exists a set of functions of $y_j\left(\tau\right)$ for $\tau \in T$ that satisfy the concave-MTR and MTS assumptions and that attain the lower bound.

(i) *Proof of the existence of the functions $y_j\left(\tau\right)$ for $\tau \in T$ that satisfy the concave-MTR and MTS assumptions and that attain the upper bound:*

The proof is organized in the following eight steps: Step 1 specifies functions $E\left[y\left(\tau\right)|\, z=s\right]$ for $\tau \in T$. Step 2 proves that these functions satisfy the concave-MTR assumption. Step 3 proves that these functions satisfy the MTS assumption. Steps 4 and 5 prove that these functions are equal to $E\left[y|\, z=s\right]$ when $\tau = s$. Steps 6, 7, and 8 prove that these functions attain the upper bound in Equation (6).

Step 1: Define

$$UB\left(s,t\right) = \min_{\{s'|s\leq s'\leq t\}}\left\{E\left[y|\, z=s'\right]+\min_{\{u|u<s'\}}\frac{E\left[y|\, z=s'\right]-E\left[y|\, z=u\right]}{s'-u}\left(t-s'\right)\right\}. \tag{A.10}$$

For $s < t$, let the function $E\left[y\left(\tau\right)|\, z = s\right]$, for $\tau \in T$, be:

$$\min\left(\min_{\{\widetilde{s}|s\leq\widetilde{s}<t\}}\left\{E\left[y|\, z=\widetilde{s}\right]+\frac{UB\left(\widetilde{s},t\right)-E\left[y|\, z=\widetilde{s}\right]}{t-\widetilde{s}}\left(\tau-\widetilde{s}\right)\right\},\ E\left[y|\, z=t\right]\right). \qquad (A.11)$$

For $s \geq t$, let the function $E\left[y\left(\tau\right)|\, z = s\right]$, for $\tau \in T$, be:

$$E\left[y|\, z=s\right]. \qquad (A.12)$$

Notice that the functions $(A.10)$, $(A.11)$, and $(A.12)$ weakly increase in s. $UB\left(s,t\right)$ weakly increases in s, because in Equation $(A.10)$ the object is minimized over the set $\{s'\,|s \leq s' \leq t\}$ such that the set $\{s'\,|s_1 \leq s' \leq t\}$ includes the set $\{s'\,|s_2 \leq s' \leq t\}$ for $s_1 \leq s_2$; the minimal value over the former set is therefore not greater than that over the latter set. Similarly, the function $(A.11)$ weakly increases in s, as does the function $(A.12)$, the latter because of the MTR and MTS assumptions.

Step 2: The function $(A.11)$ satisfies the concave-MTR assumption, since by definition its graph is the boundary of the convex hull, that is, the intersection of the subgraph of the weakly increasing linear functions in τ.[16] The function $(A.12)$ is concave-MTR in τ.

Step 3: The functions $(A.11)$ and $(A.12)$ satisfy the MTS assumption, since these functions weakly increase in s.

Step 4: We will now prove that when $s < t$ and $\tau = s$, the function $(A.11)$ is equal to $E\left[y|\, z=s\right]$.

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

$$E\left[y|\, z=s\right]\leq E\left[y|\, z=\widetilde{s}\right]+\min_{\{u|u<\widetilde{s}<t\}}\left\{\frac{E\left[y|\, z=\widetilde{s}\right]-E\left[y|\, z=u\right]}{\widetilde{s}-u}\right\}\left(s-\widetilde{s}\right), \qquad (A.13)$$

since for $s < \widetilde{s} < t$, $E\left[y|\, z=s\right] = E\left[y|\, z=\widetilde{s}\right] + \{E\left[y|\, z=\widetilde{s}\right] - E\left[y|\, z=s\right]\}/\left(\widetilde{s}-s\right)\left(s-\widetilde{s}\right)$, and for $s = \widetilde{s}$, $E\left[y|\, z=s\right] = E\left[y|\, z=\widetilde{s}\right]$.

Second, for $\widetilde{s} < t$,

$$\min_{\{u|u<\widetilde{s}<t\}}\frac{E\left[y|\, z=\widetilde{s}\right]-E\left[y|\, z=u\right]}{\widetilde{s}-u}\geq\frac{UB\left(\widetilde{s},t\right)-E\left[y|\, z=\widetilde{s}\right]}{t-\widetilde{s}}. \qquad (A.14)$$

This is because by Equation $(A.10)$,
$UB\left(\widetilde{s},t\right) \leq E\left[y|\, z=\widetilde{s}\right] + \min_{\{u|u<\widetilde{s}<t\}}\{E\left[y|\, z=\widetilde{s}\right] - E\left[y|\, z=u\right]\}/\left(\widetilde{s}-u\right)\ \left(t-\widetilde{s}\right)$.
Third, for $s \leq \widetilde{s} < t$, by Equations $(A.13)$ and $(A.14)$,

$$E\left[y|\, z=s\right]\leq E\left[y|\, z=\widetilde{s}\right]+\frac{UB\left(\widetilde{s},t\right)-E\left[y|\, z=\widetilde{s}\right]}{t-\widetilde{s}}\left(s-\widetilde{s}\right). \qquad (A.15)$$

[16] The subgraph of $f\left(\tau\right)$ is defined as $\{\left(\tau,y\right)|\, y\leq f\left(\tau\right)\}$.

By Equation $(A.15)$, when $s < t$ and $\tau = s$, the function $(A.11)$ is equal to $\min(E[y|\, z = s]$, $E[y|\, z = t])$. Because $E[y|\, z = s] \leq E[y|\, z = t]$, it follows that the function $(A.11)$ is equal to $E[y|\, z = s]$, when $s < t$ and $\tau = s$.

Step 5: The function $(A.12)$ is $E[y|\, z = s]$ when $s \geq t$ and $\tau = s$, because of its definition.

Step 6: We will now prove that when $s < t$ and $\tau = t$, the function $(A.11)$ is equal to $UB(s,t)$; and that when $s \geq t$ and $\tau = t$, the function $(A.12)$ is equal to $E[y|\, z = s]$.

When $s < t$ and $\tau = t$, the function $(A.11)$ is equal to $\min\left(\min_{\{\widetilde{s}|s\leq\widetilde{s}<t\}} UB(\widetilde{s},t),\ E[y|\, z = t]\right)$. Because $UB(s,t)$ weakly increases in s, $\min_{\{\widetilde{s}|s\leq\widetilde{s}<t\}} UB(\widetilde{s},t) = UB(s,t)$, and $UB(s,t) \leq UB(t,t) = E[y|\, z = t]$ for $s < t$. Hence, when $s < t$ and $\tau = t$, the function $(A.11)$ is equal to $UB(s,t)$.

When $s \geq t$ and $\tau = t$, the function $(A.12)$ is equal to $E[y|\, z = s]$ by its definition.

Step 7: In the case where the function $E[y(\tau)|\, z = s]$ is $(A.11)$ for $s < t$ and $(A.12)$ for $s \geq t$, it follows from Step 6 and the law of iterated expectations, and from the fact that $UB(s,t) = E[y|\, z = s]$ for $s = t$, that:

$$E[y(t)] = \sum_{s>t} E[y|\, z = s]\, P(z = s) + \sum_{s\leq t} UB(s,t)\, P(z = s)\,. \tag{A.16}$$

The quantity $(A.16)$ is the upper bound in Equation (6). Therefore, these functions attain the upper bound in Equation (6).

Step 8: By combining Steps 1 to 7, we conclude that the functions $(A.11)$ and $(A.12)$ satisfy the concave-MTR and MTS assumptions and attain the upper bound in Equation (6). Hence, there exists a set of functions of $y_j(\tau)$ for $\tau \in T$ that satisfy the concave-MTR and MTS assumptions and that attain the upper bound in Equation (6).

For $s < t$, $UB(s,t)$ is the sharp upper bound on $E[y(t)|\, z = s]$, and for $s \geq t$, $E[y|\, z = s]$ is the sharp upper bound on $E[y(t)|\, z = s]$. Hence, the sharp joint upper bound on $\{E[y(t)|\, z = s]$, $s \in T\}$ is obtained by setting each of the quantities $E[y(t)|\, z = s]$, $s \in T$ at $UB(s,t)$ for $s < t$, and at $E[y|\, z = s]$ for $s \geq t$. Therefore, the upper bound in Equation (6) is the sharp upper bound on $E[y(t)]$.

(ii) *Proof of the existence of the functions $y_j(\tau)$ for $\tau \in T$ that satisfy the concave-MTR and MTS assumptions and that attain the lower bound:*

The proof is organized in the following five steps: Step 1 specifies the functions $E[y(\tau)|\, z = s]$, for $\tau \in T$. Step 2 proves that these functions satisfy the concave-MTR assumption. Step 3 proves that these functions satisfy the MTS assumption. Step 4 proves that these functions are equal to $E[y|\, z = s]$ when $\tau = s$. Steps 5 and 6 prove that these functions attain the lower bound in Equation (6).

Step 1: Define the following four functions: For $k = s,\, \widetilde{s},\, t$ and $l = s,\, t$,

$$s'^{*}(k,t) = \arg\max_{\{s'|k\geq s'\geq t\}}\left(E[y|\, z=s'] + \max_{\{u|u<t\}}\left\{\frac{E[y|\, z=s'] - E[y|\, z=u]}{s'-u}(t-s')\right\}\right). \tag{A.17}$$

$$u^{*}(k,t) = \arg\min_{\{u|u<t\}}\frac{E[y|\, z=s'^{*}(k,t)] - E[y|\, z=u]}{s'^{*}(k,t)-u}. \tag{A.18}$$

$$LB(\tau,k,t) = E[y|\, z=s'^{*}(k,t)] + \frac{E[y|\, z=s'^{*}(k,t)] - E[y|\, z=u^{*}(k,t)]}{s'^{*}(k,t)-u^{*}(k,t)}[\tau - s'^{*}(k,t)]. \tag{A.19}$$

$$LF(\tau,l,t) = \min_{\{\widetilde{s}|\widetilde{s}\geq l\}}\min\{LB(\tau,\widetilde{s},t),\ E[y|\, z=\widetilde{s}]\}. \tag{A.20}$$

For $s \geq t$, let the function $E[y(\tau)|\, z = s]$ be

$$LF(\tau,s,t). \tag{A.21}$$

For $s < t$, let the function $E[y(\tau)|\, z = s]$ be

$$\min\{E[y|\, z=s], LF(\tau,t,t)\}. \tag{A.22}$$

Notice that for $k \geq t$, the function $LB(t,k,t)$ in Equation $(A.19)$ where $\tau = t$ weakly increases in k. This is because in Equation $(A.17)$, the object is maximized over the set $\{s'\,|k \geq s' \geq t\}$ such that the set $\{s'\,|k_1 \geq s' \geq t\}$ includes the set $\{s'\,|k_2 \geq s' \geq t\}$ for $k_1 \geq k_2$; the maximal value over the former set is therefore not smaller than that over the latter set. $LB(t,k,t)$ is the maximal value in Equation $(A.17)$. Notice also that $LF(\tau,l,t)$ weakly increases in l. This is because in Equation $(A.20)$, the object is minimized over the set $\{\widetilde{s}\,|\widetilde{s} \geq l\}$ such that the set $\{\widetilde{s}\,|\widetilde{s} \geq l_1\}$ includes the set $\{\widetilde{s}\,|\widetilde{s} \geq l_2\}$ for $l_1 \leq l_2$; the minimal value over the former set is therefore not greater than that over the latter set.

Step 2: The functions $(A.21)$ and $(A.22)$ satisfy the concave-MTR assumption, since their graphs are the boundaries of the convex hulls, and they weakly increase in τ.

Step 3: The functions $(A.21)$ and $(A.22)$ satisfy the MTS assumption, since $LF(\tau,s,t)$ and $E[y|\, z = s]$ weakly increase in s.

Step 4: We will now prove that when $s \geq t$ and $\tau = s$, the function $(A.21)$ is equal to $E[y|\, z = s]$.

First, for $\widetilde{s} \geq s \geq t$, by Equations $(A.17)$, $(A.18)$, and $(A.19)$ where $k = \widetilde{s}$ and $\tau = t$,

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

Therefore, for $\widetilde{s} \geq s \geq t > u^*\left(\widetilde{s},t\right)$,

$$\frac{E\left[y \middle| z = s'^*\left(\widetilde{s},t\right)\right] - E\left[y \middle| z = u^*\left(\widetilde{s},t\right)\right]}{s'^*\left(\widetilde{s},t\right) - u^*\left(\widetilde{s},t\right)} \geq \frac{E\left[y \middle| z = s\right] - E\left[y \middle| z = u^*\left(\widetilde{s},t\right)\right]}{s - u^*\left(\widetilde{s},t\right)} \geq 0. \qquad (A.23)$$

For $\widetilde{s} \geq s \geq t > u^*\left(\widetilde{s},t\right)$, because of Equations $(A.19)$ and $(A.23)$,

$$\begin{aligned} & LB\left(s,\widetilde{s},t\right) = E\left[y \middle| z = u^*\left(\widetilde{s},t\right)\right] + \frac{E\left[y \middle| z = s'^*\left(\widetilde{s},t\right)\right] - E\left[y \middle| z = u^*\left(\widetilde{s},t\right)\right]}{s'^*\left(\widetilde{s},t\right) - u^*\left(\widetilde{s},t\right)}\left[s - u^*\left(\widetilde{s},t\right)\right] \\ \geq \;& E\left[y \middle| z = u^*\left(\widetilde{s},t\right)\right] + \frac{E\left[y \middle| z = s\right] - E\left[y \middle| z = u^*\left(\widetilde{s},t\right)\right]}{s - u^*\left(\widetilde{s},t\right)}\left[s - u^*\left(\widetilde{s},t\right)\right] = E\left[y \middle| z = s\right]. \end{aligned} \qquad (A.24)$$

Second, for $\widetilde{s} \geq s$, because of the MTS-MTR assumption,

$$E\left[y \middle| z = \widetilde{s}\right] \geq E\left[y \middle| z = s\right]. \qquad (A.25)$$

Hence, for $s \geq t$, Equation $(A.20)$ where $\tau = s$ and $l = s$, and Equations $(A.24)$ and $(A.25)$ imply that $LF\left(s,s,t\right) = E\left[y \middle| z = s\right]$. That is, when $s \geq t$ and $\tau = s$, the function $(A.21)$ is equal to $E\left[y \middle| z = s\right]$.

Step 5: We will now prove that when $s < t$ and $\tau = s$, the functions $(A.22)$ is equal to $E\left[y \middle| z = s\right]$.

First, for $s < t \leq s'^*\left(\widetilde{s},t\right) \leq \widetilde{s}$, by Equation $(A.18)$ where $k = \widetilde{s}$, $0 \leq \{E\left[y \middle| z = s'^*\left(\widetilde{s},t\right)\right] - E\left[y \middle| z = u^*\left(\widetilde{s},t\right)\right]\}/\left[s'^*\left(\widetilde{s},t\right) - u^*\left(\widetilde{s},t\right)\right] \leq \{E\left[y \middle| z = s'^*\left(\widetilde{s},t\right)\right] - E\left[y \middle| z = s\right]\}/\left[s'^*\left(\widetilde{s},t\right) - s\right]$.

Thus,

$$\begin{aligned} LB\left(s,\widetilde{s},t\right) &\geq E\left[y \middle| z = s'^*\left(\widetilde{s},t\right)\right] + \frac{E\left[y \middle| z = s'^*\left(\widetilde{s},t\right)\right] - E\left[y \middle| z = s\right]}{s'^*\left(\widetilde{s},t\right) - s}\left[s - s'^*\left(\widetilde{s},t\right)\right] \\ &= E\left[y \middle| z = s\right]. \end{aligned} \qquad (A.26)$$

Second, for $s < t \leq \widetilde{s}$, Equation $(A.25)$ holds.

By Equations $(A.25)$ and $(A.26)$, $LF\left(s,t,t\right) \geq E\left[y \middle| z = s\right]$.

Therefore, when $s < t$ and $\tau = s$, the function $(A.22)$ is equal to $E[y|\, z = s]$.

Step 6: We will now prove that the functions $(A.21)$ and $(A.22)$ attain the lower bound in Equation (6). The proof is organized in the following five substeps: Substeps 6.1 to 6.3 prove that the function $(A.21)$ is equal to $LB(t, s, t)$, when $s \geq t$ and $\tau = t$. Substep 6.4 proves that the function $(A.22)$ is equal to $E[y|\, z = s]$ when $s < t$ and $\tau = t$. Substep 6.5 uses the previous substeps and the law of iterated expectations to prove that the functions $(A.21)$ and $(A.22)$ attain the lower bound in Equation (6).

Substep 6.1: Since the function $LB(t, k, t)$ weakly increases in k, for $\widetilde{s} \geq s$,

$$LB(t, s, t) \leq LB(t, \widetilde{s}, t). \tag{A.27}$$

Substep 6.2: When $k = \widetilde{s}$ in Equations $(A.17)$ and $(A.18)$, $\widetilde{s} \geq s'^{*}(\widetilde{s}, t) \geq t > u^{*}(\widetilde{s}, t)$. Therefore, the MTS-MTR assumption implies that $E[y|\, z = \widetilde{s}] \geq E[y|\, z = s'^{*}(\widetilde{s}, t)]$ and $E[y|\, z = s'^{*}(\widetilde{s}, t)] \geq E[y|\, z = u^{*}(\widetilde{s}, t)]$. Hence, for $\widetilde{s} \geq t$,

$$LB(t, \widetilde{s}, t) \leq E[y|\, z = \widetilde{s}]. \tag{A.28}$$

Substep 6.3: By Equations $(A.20)$ where $\tau = t$ and $l = s$, and by Equations $(A.27)$ and $(A.28)$, it follows that for $s \geq t$,

$$LF(t, s, t) = LB(t, s, t). \tag{A.29}$$

Hence, when $s \geq t$ and $\tau = t$, the function $(A.21)$ is equal to $LB(t, s, t)$.

Substep 6.4: The proof of Substep 6.4 will be constructed along lines that are similar to the proofs of Substeps 6.1 to 6.3. For $\widetilde{s} \geq t$, because (i) Equations $(A.17)$, $(A.18)$, and $(A.19)$ hold, (ii) $LB(t, k, t)$ weakly increases in k for $k \geq t$, and (iii) $s'^{*}(t, t) = t$, it follows that:

$$LB(t, \widetilde{s}, t) \geq LB(t, t, t) = E[y|\, z = t]. \tag{A.30}$$

Also, for $\widetilde{s} \geq t$,

$$E[y|\, z = \widetilde{s}] \geq E[y|\, z = t]. \tag{A.31}$$

By Equation $(A.20)$ where $\tau = t$ and $l = t$, and by Equations $(A.30)$ and $(A.31)$,

$$LF(t, t, t) = E[y|\, z = t]. \tag{A.32}$$

For $s < t$,

$$E[y|\, z = s] \leq E[y|\, z = t]. \tag{A.33}$$

Hence, by Equations $(A.32)$ and $(A.33)$, when $s < t$ and $\tau = t$, the function $(A.22)$ is equal to $E\left[y|\, z = s\right]$.

Substep 6.5: When the function $E\left[y\left(\tau\right)|\, z = s\right]$ is $(A.21)$ for $s \geq t$ and $(A.22)$ for $s < t$, by Substeps 6.3 and 6.4, together with the law of iterated expectations,

$$E\left[y\left(t\right)\right] = \sum_{s<t} E\left[y|\, z = s\right] P\left(z = s\right) + \sum_{s\geq t} LB\left(t, s, t\right) P\left(z = s\right). \qquad (A.34)$$

The quantity $(A.34)$ is the lower bound in Equation (6). Therefore, these functions attain the lower bound in Equation (6).

Step 7: By combining Steps 1 to 6, we conclude that the functions $(A.21)$ and $(A.22)$ satisfy the concave-MTR and MTS assumptions and attain the lower bound in Equation (6). Hence, there exists a set of functions of $y_j\left(\tau\right)$ for $\tau \in T$ that satisfy the concave-MTR and MTS assumptions and that attain the lower bound in Equation (6).

For $s \geq t$, $LB\left(t, s, t\right)$ is the sharp lower bound on $E\left[y\left(t\right)|\, z = s\right]$, and for $s < t$, $E\left[y|\, z = s\right]$ is the sharp lower bound on $E\left[y\left(t\right)|\, z = s\right]$. Hence, the sharp joint lower bound on $\left\{E\left[y\left(t\right)|\, z = s\right],\ s \in T\right\}$ is obtained by setting each of the quantities $E\left[y\left(t\right)|\, z = s\right],\ s \in T$ at $LB\left(t, s, t\right)$ for $s \geq t$, and at $E\left[y|\, z = s\right]$ for $s < t$. Therefore, the lower bound in Equation (6) is the sharp lower bound on $E\left[y\left(t\right)\right]$. *Q.E.D.*

Proof of Part (3):

We will now prove that our bounds in Equation (6) are narrower than or equal to both Manski's (1997) bounds, as represented in Equation (3) above; and Manski and Pepper's (2000) bounds, as represented in Equation (5) above.

The first terms of the upper bounds in Equations (3), (5), and (6) are the same, and the first terms of the lower bounds in these equations are also the same. Therefore, we will now compare the second terms of the bounds in these equations.

(i) Comparison with the bounds in Manski (1997) (Equation (3)):

Because the set of realized treatments and outcomes includes $(z, y) = (0, 0)$ and $E\left[0|\, z = 0\right] = 0$, it follows that, for $s \leq t$:

$$\begin{aligned} &\min_{\{s'|s\leq s'\leq t\}} \left\{E\left[y|\, z = s'\right] + \min_{\{u|u<s'\}} \frac{E\left[y|\, z = s'\right] - E\left[y|\, z = u\right]}{s' - u} \left(t - s'\right)\right\} \\ \leq\ & E\left[y|\, z = s\right] + \frac{E\left[y|\, z = s\right]}{s} \left(t - s\right) = E\left[\left.\frac{y}{z} t \right|\, z = s\right]; \qquad (A.35) \end{aligned}$$

and for $s \geq t$,

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

$$\geq \; E\left[y|\, z=s\right] + \frac{E\left[y|\, z=s\right]}{s}(t-s) = E\left[\left.\frac{y}{z}t\right|\, z=s\right]. \tag{A.36}$$

(The right-hand sides of Equations $(A.35)$ and $(A.36)$ are included in the brackets of the left-hand sides of Equations $(A.35)$ and $(A.36)$, respectively, for the case where $u=0$ and $s'=s$.)

Taking Equations $(A.35)$ and $(A.36)$ together with the law of iterated expectations implies that the second term of the upper bound in Equation (6) is smaller than or equal to that in Equation (3), and that the second term of the lower bound in Equation (6) is greater than or equal to that in Equation (3). Therefore, our bounds in Equation (6) are narrower than or equal to Manski's (1997) bounds as shown in Equation (3).

(ii) Comparison with the bounds in Manski and Pepper (2000) (Equation (5)):

For $s \leq t$,

$$\min_{\{s'|s \leq s' \leq t\}} \left\{ E\left[y|\, z=s'\right] + \min_{\{u|u<s'\}} \frac{E\left[y|\, z=s'\right] - E\left[y|\, z=u\right]}{s'-u} (t-s') \right\} \leq E\left[y|\, z=t\right]; \tag{A.37}$$

and for $s \geq t$,

$$\max_{\{s'|s \geq s' \geq t\}} \left(E\left[y|\, z=s'\right] + \max_{\{u|u<t\}} \left\{ \frac{E\left[y|\, z=s'\right] - E\left[y|\, z=u\right]}{s'-u} (t-s') \right\} \right) \geq E\left[y|\, z=t\right]. \tag{A.38}$$

(The right-hand sides of Equations $(A.37)$ and $(A.38)$ are included in the case of $s'=t$ in the brackets of the left-hand sides of Equations $(A.37)$ and $(A.38)$, respectively.)

Taking Equations $(A.37)$ and $(A.38)$ together with the law of iterated expectations implies that the second term of the upper bound in Equation (6) is smaller than or equal to that in Equation (5), and that the second term of the lower bound in Equation (6) is greater than or equal to that in Equation (5). Therefore, our bounds in Equation (6) are narrower than or equal to Manski and Pepper's (2000) bounds as shown in Equation (5). *Q.E.D.*

5.2 Appendix B: Proof of Proposition 2

Proof of Part (i):

The lower bound on $E[y(t_2)] - E[y(t_1)]$ in Equation (18) 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 assumptions (since $E[y|\, z = s]$ increases in s).

In order to obtain the sharp upper bound on $E[y(t_2)] - E[y(t_1)]$, let us first obtain the sharp upper bound on $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$.

For $(s, t_1, t_2) \in T^3$, in order to obtain the sharp upper bound on $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$, hold $E[y(t_2)|\, z = s]$ fixed and minimize $E[y(t_1)|\, z = s]$ subject to these three conditions: (a) that the function $E[y(\tau)|\, z = s]$ for $\tau \in T$ traverses the three points $(t_2, E[y(t_2)|\, z = s])$, $(t_1, E[y(t_1)|\, z = s])$, and $(s, E[y|\, z = s])$; (b) that this function satisfies the concave-MTR assumption; and (c) that this function satisfies the MTS assumption. This yields the maximum of $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$ as a function of $E[y(t_2)|\, z = s]$. Then maximize this expression over $E[y(t_2)|\, z = s]$.

To implement this strategy, we use the following seven-step process: In Steps 1 and 2, we set $E[y(t_2)|\, z = s]$ at the sharp upper bound on $E[y(t_2)|\, z = s]$. In Steps 3 and 4, given $E[y(t_2)|\, z = s]$ which is equal to its sharp upper bound, we minimize $E[y(t_1)|\, z = s]$ subject to conditions (a), (b), and (c). Steps 1 to 4 therefore determine the value $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$. In Step 5, we show that this value is greater than or equal to other values $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$ such that $E[y(\tau)|\, z = s]$ satisfies conditions (a), (b), and (c). In Step 6, by combining Steps 1 to 5, we show the sharp upper bound on $E[y(t_2)|\, z = s] - E[y(t_1)|\, z = s]$. In Step 7, we conclude that the sharp upper bound on $E[y(t_2)] - E[y(t_1)]$ is the upper bound in Equation (18).

Step 1: Equations $(A.6)$ and $(A.8)$ in Appendix A and Equation (10) imply:

$$E[y(t_2)|\, z = s] \le UB(s, t_2) \qquad \text{for } z = s \le t_2, \tag{A.39}$$

and

$$E[y(t_2)|\, z = s] \le E[y|\, z = s] \quad \text{for } t_2 < s = z. \tag{A.40}$$

The proof of Part (2) of Proposition 1 in Appendix A implies that these upper bounds are sharp.

Step 2: Set $E[y(t_2)|\, z = s]$ at $UB(s, t_2)$ in Case (1) where $z = s \le t_2$, and at $E[y|\, z = s]$

in Case (2) where $t_2 < s = z$. Then, find the minimal value of $E[y(t_1)|\, z = s]$ subject to conditions (a), (b), and (c). Steps 3 and 4 obtain these minimal values.

Step 3: The claim of this step is the following:

In Case (1) where $z = s \leq t_2$: Given $E[y(t_2)|\, z = s] = UB(s, t_2)$, then $AT_1(t_1, s, t_2)$ is the minimal value of $E[y(t_1)|\, z = s]$, subject to conditions (a), (b), and (c). We prove this claim using four substeps: 3.1, 3.2, 3.3, and 3.4.

Substep 3.1: Proof that $AT_1(\tau, s, t_2)$ satisfies condition (a):
Equation (9) implies that $AT_1(t_2, s, t_2) = UB(s, t_2)$ and $AT_1(s, s, t_2) = E[y|\, z = s]$. Therefore, $AT_1(\tau, s, t_2)$ traverses the three points $(t_2, UB(s, t_2))$, $(t_1, AT_1(t_1, s, t_2))$, and $(s, E[y|\, z = s])$, thereby satisfying condition (a). Note that $UB(s, t_2) = E[y|\, z = s]$ if $s = t_2$.

Substep 3.2: Proof that $AT_1(\tau, s, t_2)$ satisfies condition (b):
When $s < t_2$, by Equation (8) where $m = 1$ and Equation $(A.14)$ where $\widetilde{s} = s$, $t = t_2$, and $u = v(s)$, it follows that

$$0 \leq \frac{UB(s, t_2) - E[y|\, z = s]}{t_2 - s} \leq \frac{E[y|\, z = s] - E[y|\, z = v(s)]}{s - v(s)}. \qquad (A.41)$$

When $s = t_2$, by Equation (9),

$$AT_1(t_2, s, t_2) = E[y|\, z = s]. \qquad (A.42)$$

Because Equation (8) holds, $v^{m+1}(s) < v^m(s) < v^{m-1}(s)$, and $E[y|\, z = s]$ weakly increases in s, then for $s \leq t_2$ and $m = 1, 2, ..., M(s)$,

$$0 \leq \frac{E[y|\, z = v^{m-1}(s)] - E[y|\, z = v^m(s)]}{v^{m-1}(s) - v^m(s)} \leq \frac{E[y|\, z = v^{m-1}(s)] - E[y|\, z = v^{m+1}(s)]}{v^{m-1}(s) - v^{m+1}(s)}. \qquad (A.43)$$

Now consider 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 $(v^{m+1}(s), E[y|\, z = v^{m+1}(s)])$, where $v^{m+1}(s) < v^m(s) < v^{m-1}(s)$. The slope of the line linking the two vertices $(v^{m-1}(s), E[y|\, z = v^{m-1}(s)])$ and $(v^m(s), E[y|\, z = v^m(s)])$ and the slope of the line linking the two vertices $(v^{m-1}(s), E[y|\, z = v^{m-1}(s)])$ and $(v^{m+1}(s), E[y|\, z = v^{m+1}(s)])$ satisfy Equation $(A.43)$. Then the slope of the line linking the two vertices $(v^m(s), E[y|\, z = v^m(s)])$ and $(v^{m+1}(s), E[y|\, z = v^{m+1}(s)])$ satisfies the following equation: for $s \leq t_2$ and $m = 1, 2, ..., M(s)$,

$$0 \leq \frac{E[y|\, z = v^{m-1}(s)] - E[y|\, z = v^m(s)]}{v^{m-1}(s) - v^m(s)} \leq \frac{E[y|\, z = v^m(s)] - E[y|\, z = v^{m+1}(s)]}{v^m(s) - v^{m+1}(s)}. \qquad (A.44)$$

And for $\tau > t_2$, by Equation (12), the slope of $AT_1(\tau, s, t_2)$ is zero.
Therefore, by Equations $(A.41)$, $(A.42)$, and $(A.44)$, $AT_1(\tau, s, t_2)$ is concave-MTR in $\tau \in T$.
$Q.E.D.$

Substep 3.3: The proof that $AT_1(\tau, s, t_2)$ satisfies condition (c):
Because of Equations (8), (9), and (11), the subgraph of $AT_1(\tau, s, t_2)$ includes all observations whose treatments are not greater than s (i.e., $(k, E[y|\, z = k])$ for $k \leq s$). Therefore, this includes the segments linking any two of those observations. This result and Equation (11) imply that for $\tau \leq s' \leq s \leq t_2$,

$$AT_1(\tau, s', t_2) \leq AT_1(\tau, s, t_2). \tag{A.45}$$

Therefore, $AT_1(s', s', t_2) \leq AT_1(s', s, t_2)$ for $s' \leq s$. The proof of Part (2) in Appendix A shows that $UB(s, t_2)$ weakly increases in s; that is, $UB(s', t_2) \leq UB(s, t_2)$ for $s' \leq s$. Thus, between s' and t_2, the function describing the segment linking the point $(s',\, AT_1(s', s, t_2))$ and the point $(t_2,\, UB(s, t_2)) = (t_2, AT_1(t_2, s, t_2))$ is not smaller than the function describing the segment linking the point $(s',\, AT_1(s', s', t_2))$ and the point $(t_2,\, UB(s', t_2))$. Hence, because Equation (9) holds, and because $AT_1(\tau, s, t_2)$ is concave-MTR in τ, it follows that Equation $(A.45)$ holds for $s' \leq \tau \leq t_2$ and $s' \leq s \leq t_2$.
In addition, for $s' \leq s \leq t_2 < \tau$, by Equation (12), Equation $(A.45)$ holds.
Because Equation $(A.45)$ holds for $\tau \in T$, $AT_1(\tau, s, t_2)$ is MTS. $Q.E.D.$

Substep 3.4: The claim of this substep is that given $E[y(t_2)|\, z = s] = UB(s, t_2)$, $AT_1(t_1, s, t_2)$ is smaller than or equal to the value $E[y(t_1)|\, z = s]$ for any function $E[y(\tau)|\, z = s]$ which satisfies the conditions (a), (b), and (c), identified previously. We divide Case (1) where $z = s \leq t_2$ into two subcases: (1.1), where $z = s \leq t_1 < t_2$, and (1.2), where $t_1 < z = s \leq t_2$; we then prove this claim for each of these two subcases.

Subcase (1.1), where $z = s \leq t_1 < t_2$: If the concave-MTR function $E[y(\tau)|\, z = s]$ traverses the points $(t_2, UB(s, t_2))$ and $(s, E[y|\, z = s])$, then for $t_1 \in [s, t_2)$, the value $E[y(t_1)|\, z = s]$ is not smaller than the value $AT_1(t_1, s, t_2)$ because the function $AT_1(\tau, s, t_2)$ for $\tau \in [s, t_2)$ in Equation (9) describes the segment linking the points $(t_2, UB(s, t_2))$ and $(s, E[y|\, z = s])$.

Subcase (1.2), where $t_1 < z = s \leq t_2$: For $v^m(s) \leq t_1 < v^{m-1}(s) \leq s$ $(m = 1, 2, ..., M(s))$, because the function $E[y(\tau)|\, z = s]$ is concave-MTR, the value $E[y(t_1)|\, z = s]$ is not smaller than the value of the function describing the segment linking the points

$\left(v^{m-1}(s), E\left[y\left(v^{m-1}(s)\right) \middle| z=s\right]\right)$ and $\left(v^{m}(s), E\left[y\left(v^{m}(s)\right) \middle| z=s\right]\right)$, evaluated at t_1. Because $E\left[y(\tau)\middle| z=s\right]$ is MTS, $E\left[y\middle| z=v^{m-1}(s)\right] \leq E\left[y\left(v^{m-1}(s)\right)\middle| z=s\right]$ and $E\left[y\middle| z=v^{m}(s)\right] \leq E\left[y\left(v^{m}(s)\right)\middle| z=s\right]$. Therefore, the value $E\left[y(t_1)\middle| z=s\right]$ is not smaller than the value $AT_1(t_1,s,t_2)$ because the function $AT_1(\tau,s,t_2)$ for $\tau \in \left[v^{m}(s), v^{m-1}(s)\right)$ in Equation (11) describes the segment linking the points $\left(v^{m-1}(s), E\left[y\middle| z=v^{m-1}(s)\right]\right)$ and $\left(v^{m}(s), E\left[y\middle| z=v^{m}(s)\right]\right)$.

Therefore, the claim of this substep is true.

Combining Substeps 3.1 to 3.4, we conclude that in Case (1) where $z=s\leq t_2$, given $E\left[y(t_2)\middle| z=s\right] = UB(s,t_2)$, then $AT_1(t_1,s,t_2)$ is the minimal value of $E\left[y(t_1)\middle| z=s\right]$, subject to conditions (a), (b) and (c).

Step 4: The claim of this step is the following:
In Case (2) where $z=s>t_2$: Given $E\left[y(t_2)\middle| z=s\right] = E\left[y\middle| z=s\right]$, then $AT_2(t_1,s,t_2)$ is the minimal value of $E\left[y(t_1)\middle| z=s\right]$, subject to conditions (a), (b), and (c) as specified earlier. Like our proof of Case (1) in Step 3, we prove this claim using five substeps.

Substep 4.1: By Equations (15), (16), and (17), $AT_2(\tau,s,t_2)$ traverses the points $(t_2, E\left[y\middle| z=s\right])$, $(t_1, AT_2(t_1,s,t_2))$, and $(s, E\left[y\middle| z=s\right])$, and therefore satisfies condition (a).

Substep 4.2: $AT_2(\tau,s,t_2)$ satisfies condition (b).

Substep 4.3: $AT_2(\tau,s,t_2)$ satisfies condition (c).

Substep 4.4: For $s' \leq t_2 < s$, $AT_1(\tau,s',t_2)$ and $AT_2(\tau,s,t_2)$ satisfy condition (c).
The proofs of Substeps 4.2, 4.3, and 4.4 can be constructed along lines that are similar to the proof in Substeps 3.2 and 3.3 of Step 3 that $AT_1(\tau,s,t_2)$ satisfies conditions (b) and (c). (For the proof of Substep 4.4, we use the fact that for $s' \leq t_2 < s$, $UB(s',t_2) \leq UB(t_2,t_2) = E\left[y\middle| z=t_2\right] \leq E\left[y\middle| z=s\right]$ to show that $AT_1(\tau,s',t_2) \leq AT_2(\tau,s,t_2)$.)

Substep 4.5: The claim of this substep is that given $E\left[y(t_2)\middle| z=s\right] = E\left[y\middle| z=s\right]$, it follows that $AT_2(t_1,s,t_2)$ is smaller than or equal to the value of $E\left[y(t_1)\middle| z=s\right]$ for any function $E\left[y(\tau)\middle| z=s\right]$ which satisfies conditions (a), (b), and (c). We divide Case (2) where $z=s>t_2$ into two subcases: (2.1), where $v(s,t_2) \leq t_1 < t_2 < z=s$, and (2.2), where $v^{m}(s,t_2) \leq t_1 < v^{m-1}(s,t_2) \leq t_2 < z=s$ $(m=2,3,...,M(s,t_2))$; we then prove this claim for each subcase.

Subcase (2.1), where $v(s,t_2) \leq t_1 < t_2 < z=s$: Because the function $E\left[y(\tau)\middle| z=s\right]$ is concave-MTR and traverses the points $(t_2, E\left[y\middle| z=s\right])$ and $\left(v(s,t_2), E\left[y\left(v(s,t_2)\right)\middle| z=s\right]\right)$, the value $E\left[y(t_1)\middle| z=s\right]$ is not smaller than the value of the function describing the segment linking these two points, evaluated at t_1. Because $E\left[y(\tau)\middle| z=s\right]$ is MTS, $E\left[y\middle| z=v(s,t_2)\right] \leq$

$E\left[y\left(v\left(s,t_2\right)\right)\mid z=s\right]$. The function $AT_2\left(\tau,s,t_2\right)$ for $\tau\in\left[v\left(s,t_2\right),t_2\right)$ describes the segment linking the points $\left(t_2,E\left[y\mid z=s\right]\right)$ and $\left(v\left(s,t_2\right),E\left[y\mid z=v\left(s,t_2\right)\right]\right)$. Therefore, the value $E\left[y\left(t_1\right)\mid z=s\right]$ is not smaller than the value $AT_2\left(t_1,s,t_2\right)$.

Subcase (2.2), where $v^m\left(s,t_2\right)\leq t_1<v^{m-1}\left(s,t_2\right)\leq t_2<z=s$: The value $E\left[y\left(t_1\right)\mid z=s\right]$ is not smaller than the value $AT_2\left(t_1,s,t_2\right)$. The proof can be constructed along lines that are similar to the proof in Subcase (1.2) of Substep 3.4.

Therefore, the claim of Substep 4.5 is true.

Combining Substeps 4.1 to 4.5, we conclude that given $E\left[y\left(t_2\right)\mid z=s\right]=E\left[y\mid z=s\right]$, then $AT_2\left(t_1,s,t_2\right)$ is the minimal value of $E\left[y\left(t_1\right)\mid z=s\right]$, subject to conditions (a), (b), and (c).

Step 5: In Steps 1 to 4, we have shown that when $s\leq t_2$ and $E\left[y\left(t_2\right)\mid z=s\right]=UB\left(s,t_2\right)$, the maximum of $E\left[y\left(t_2\right)\mid z=s\right]-E\left[y\left(t_1\right)\mid z=s\right]$ subject to conditions (a), (b), and (c) is $UB\left(s,t_2\right)-AT_1\left(t_1,s,t_2\right)$. Also, when $s>t_2$ and $E\left[y\left(t_2\right)\mid z=s\right]=E\left[y\mid z=s\right]$, the maximum of $E\left[y\left(t_2\right)\mid z=s\right]-E\left[y\left(t_1\right)\mid z=s\right]$ subject to these conditions is $E\left[y\mid z=s\right]-AT_2\left(t_1,s,t_2\right)$. In Step 5, we will show that these maxima are greater than or equal to the maxima of $E\left[y\left(t_2\right)\mid z=s\right]-E\left[y\left(t_1\right)\mid z=s\right]$, such that $E\left[y\left(t_2\right)\mid z=s\right]$ is different from $UB\left(s,t_2\right)$ for $s\leq t_2$, or it is different from $E\left[y\mid z=s\right]$ for $s>t_2$; and such that $E\left[y\left(\tau\right)\mid z=s\right]$ satisfies conditions (a), (b), and (c).

In Case (1) where $z=s\leq t_2$: Suppose that we set $E\left[y\left(t_2\right)\mid z=s\right]$ at a value which is smaller than $UB\left(s,t_2\right)$. Let this value be $VB\left(s,t_2\right)$ where $VB\left(s,t_2\right)<UB\left(s,t_2\right)$. (Note that $UB\left(s,t_2\right)$ is the sharp upper bound on $E\left[y\left(t_2\right)\mid z=s\right]$ in Case (1) in Equation ($A.39$).) Given $E\left[y\left(t_2\right)\mid z=s\right]=VB\left(s,t_2\right)$, we will now minimize $E\left[y\left(t_1\right)\mid z=s\right]$ such that $E\left[y\left(\tau\right)\mid z=s\right]$ satisfies conditions (a), (b), and (c). The process for obtaining the minimal value of $E\left[y\left(t_1\right)\mid z=s\right]$ is similar to that in Step 3.

In Subcase (1.1), where $z=s\leq t_1<t_2$: Given $E\left[y\left(t_2\right)\mid z=s\right]=VB\left(s,t_2\right)$, the minimal value of $E\left[y\left(t_1\right)\mid z=s\right]$ such that $E\left[y\left(\tau\right)\mid z=s\right]$ satisfies conditions (a), (b), and (c) is the value of the function describing the segment linking the points $\left(t_2,VB\left(s,t_2\right)\right)$ and $\left(s,E\left[y\mid z=s\right]\right)$, evaluated at t_1. The function $AT_1\left(\tau,s,t_2\right)$ for $\tau\in\left[s,t_2\right)$ describes the segment linking the points $\left(t_2,UB\left(s,t_2\right)\right)$ and $\left(s,E\left[y\mid z=s\right]\right)$. Therefore, the maximum of $VB\left(s,t_2\right)-E\left[y\left(t_1\right)\mid z=s\right]$ subject to conditions (a), (b), and (c) is smaller than $UB\left(s,t_2\right)-AT_1\left(t_1,s,t_2\right)$.

In Subcase (1.2), where $t_1<z=s\leq t_2$: Given $E\left[y\left(t_2\right)\mid z=s\right]=VB\left(s,t_2\right)$, the

minimal value of $E\left[y\left(t_1\right)\right|z=s]$ such that $E\left[y\left(\tau\right)\right|z=s]$ satisfies conditions (a), (b), and (c) is $AT_1\left(t_1,s,t_2\right)$. Therefore, the value $VB\left(s,t_2\right)-AT_1\left(t_1,s,t_2\right)$ is smaller than the value $UB\left(s,t_2\right)-AT_1\left(t_1,s,t_2\right)$.

Thus, in Case (1) $z=s\leq t_2:UB\left(s,t_2\right)-AT_1\left(t_1,s,t_2\right)$ is greater than or equal to a value $E\left[y\left(t_2\right)\right|z=s]-E\left[y\left(t_1\right)\right|z=s]$ for a function $E\left[y\left(\tau\right)\right|z=s]$ which satisfies conditions (a), (b), and (c).

In Case (2), where $z=s>t_2$: Suppose that we set $E\left[y\left(t_2\right)\right|z=s]$ at a value which is smaller than $E\left[y\right|z=s]$. Let this value be $WB\left(s,t_2\right)$ where $WB\left(s,t_2\right)<E\left[y\right|z=s]$. (Note that $E\left[y\right|z=s]$ is the sharp upper bound on $E\left[y\left(t_2\right)\right|z=s]$ in Case (2) in Equation $(A.40)$.) A process similar to that in Step 4 can now apply to obtain the minimal value of $E\left[y\left(t_1\right)\right|z=s]$ subject to conditions (a), (b), and (c), given $E\left[y\left(t_2\right)\right|z=s]=WB\left(s,t_2\right)$.

In Subcase (2.1), where $v\left(s,t_2\right)\leq t_1<t_2<z=s$: Given $E\left[y\left(t_2\right)\right|z=s]=WB\left(s,t_2\right)$, the minimal value of $E\left[y\left(t_1\right)\right|z=s]$ subject to conditions (a), (b), and (c) is not smaller than the value of the function describing the segment linking the points $\left(t_2,WB\left(s,t_2\right)\right)$ and $\left(v\left(s,t_2\right),E\left[y\right|z=v\left(s,t_2\right)]\right)$, evaluated at t_1. The proof of this claim can be constructed along lines that are similar to the proof of Subcase (2.1) in Substep 4.5. The function $AT_2\left(\tau,s,t_2\right)$ for $\tau\in\left[v\left(s,t_2\right),t_2\right)$ describes the segment linking the points $\left(t_2,E\left[y\right|z=s]\right)$ and $\left(v\left(s,t_2\right),E\left[y\right|z=v\left(s,t_2\right)]\right)$. Therefore, the maximum of $WB\left(s,t_2\right)-E\left[y\left(t_1\right)\right|z=s]$ subject to conditions (a), (b), and (c) is smaller than $E\left[y\right|z=s]-AT_2\left(t_1,s,t_2\right)$.

In Subcase (2.2), where $v^m\left(s,t_2\right)\leq t_1<v^{m-1}\left(s,t_2\right)\leq t_2<z=s$ $\left(m=2,3,...,M\left(s,t_2\right)\right)$: Given $E\left[y\left(t_2\right)\right|z=s]=WB\left(s,t_2\right)$, the minimal value of $E\left[y\left(t_1\right)\right|z=s]$ subject to conditions (a), (b), and (c) is not smaller than the value $AT_2\left(t_1,s,t_2\right)$. The proof of this claim can be constructed along lines that are similar to the proof in Subcase (2.2) of Substep 4.5. Therefore, the maximum value of $WB\left(s,t_2\right)-E\left[y\left(t_1\right)\right|z=s]$ is smaller than the value $E\left[y\right|z=s]-AT_2\left(t_1,s,t_2\right)$.

Thus, in Case (2), where $z=s>t_2:E\left[y\right|z=s]-AT_2\left(t_1,s,t_2\right)$ is greater than or equal to a value $E\left[y\left(t_2\right)\right|z=s]-E\left[y\left(t_1\right)\right|z=s]$ for a function $E\left[y\left(\tau\right)\right|z=s]$ which satisfies conditions (a), (b), and (c).

Step 6: By combining Steps 1 to 5, we conclude the following:

In Case (1) where $z=s\leq t_2$ and $t_1<t_2$,

$$0\leq E\left[y\left(t_2\right)\right|z=s]-E\left[y\left(t_1\right)\right|z=s]\leq UB\left(s,t_2\right)-AT_1\left(t_1,s,t_2\right). \qquad (A.46)$$

In Case (2) where $t_1 < t_2 < s = z$,

$$0 \leq E\left[y\left(t_2\right)\middle| z = s\right] - E\left[y\left(t_1\right)\middle| z = s\right] \leq E\left[y\middle| z = s\right] - AT_2\left(t_1, s, t_2\right). \qquad (A.47)$$

These bounds are sharp.

Step 7: By Step 6, the sharp joint upper bound on $\{E\left[y\left(t_2\right)\middle| z = s\right] - E\left[y\left(t_1\right)\middle| z = s\right], s \in T\}$ is obtained by setting each of the quantities $E\left[y\left(t_2\right)\middle| z = s\right] - E\left[y\left(t_1\right)\middle| z = s\right]$, $s \in T$, at its upper bound in Equation $(A.46)$ for $s \leq t_2$, and at its upper bound in Equation $(A.47)$ for $s > t_2$. Therefore, by the law of iterated expectations, we conclude that the sharp upper bound on $E\left[y\left(t_2\right)\right] - E\left[y\left(t_1\right)\right]$ is the upper bound in Equation (18). *Q.E.D.*

Proof of Part (ii):

We will now prove that our bounds in Equation (18) are narrower than or equal to the bounds in Manski (1997) and Manski and Pepper (2000).

(1) Comparison with the bounds in Manski (1997):

The sharp bounds on the average treatment effects obtained using only the concave-MTR assumption of Manski (1997) are:

$$\begin{aligned} 0 &\leq E\left[y\left(t_2\right)\right] - E\left[y\left(t_1\right)\right] && (A.48)\\ &\leq \sum_{s \leq t_2} E\left[\frac{y}{z}\middle| z = s\right]\left(t_2 - t_1\right)P\left(z = s\right) + \sum_{s > t_2}\left\{E\left[y\middle| z = s\right] - E\left[\frac{y}{t_2}t_1\middle| z = s\right]\right\}P\left(z = s\right). \end{aligned}$$

We will now compare the upper bounds in Equations (18) and $(A.48)$ using the two steps: Step (1.1) compares the first terms of the upper bounds in Equations (18) and $(A.48)$, and Step (1.2) compares the second terms.

Step (1.1): For $s \leq t_2$, the function $AT_1\left(\tau, s, t_2\right)$ is concave-MTR in τ and traverses the points $\left(t_2, UB\left(s, t_2\right)\right)$ and $\left(v^{M(s)}\left(s\right), E\left[y\middle| z = v^{M(s)}\left(s\right)\right]\right) = \left(0, E\left[y\middle| z = 0\right]\right)$ where $E\left[y\middle| z = 0\right] \geq 0$; this follows from the proof of Part (i) and Equations (9) and (11). Then

$$AT_1\left(t_1, s, t_2\right) \geq \frac{UB\left(s, t_2\right)}{t_2}t_1. \qquad (A.49)$$

Because Equation (10) holds, the set of realized treatments and outcomes includes $(0, 0)$, and $E\left[0\middle| z = 0\right] = 0$, for $s \leq t_2$,

$$UB\left(s, t_2\right) \leq E\left[y\middle| z = s\right] + \frac{E\left[y\middle| z = s\right]}{s}\left(t_2 - s\right) = E\left[\frac{y}{z}t_2\middle| z = s\right]. \qquad (A.50)$$

By Equations $(A.49)$ and $(A.50)$, for $s \leq t_2$, $UB\left(s, t_2\right) - AT_1\left(t_1, s, t_2\right) \leq E\left[y/z\middle| z = s\right]\left(t_2 - t_1\right)$.

Step (1.2): For $s > t_2$, the function $AT_2(\tau, s, t_2)$ is concave-MTR in τ and traverses the points $(t_2, E[y|\, z = s])$ and $\left(v^{M(s,t)}(s,t), E\left[y|\, z = v^{M(s,t)}(s,t)\right]\right) = (0, E[y|\, z = 0])$ where $E[y|\, z = 0] \geq 0$; this follows from Equations (15) and (16). Then $AT_2(t_1, s, t_2) \geq E[(y/t_2)\, t_1|\, z = s]$. Thus, for $s > t_2$, $E[y|\, z = s] - AT_2(t_1, s, t_2) \leq E[y|\, z = s] - E[(y/t_2)\, t_1|\, z = s]$.

Therefore, by the law of iterated expectations, the upper bound in Equation (18) is smaller than or equal to that in Equation ($A.48$).

(2) Comparison with the bounds in Manski and Pepper (2000):
The sharp bounds on the average treatment effects using only the MTR and MTS assumptions of Manski and Pepper (2000) are:

$$\begin{aligned} 0 &\leq E[y(t_2)] - E[y(t_1)] \qquad (A.51) \\ &\leq \sum_{s<t_1} \{E[y|\, z = t_2] - E[y|\, z = s]\} P(z = s) + \sum_{t_1 \leq s \leq t_2} \{E[y|\, z = t_2] - E[y|\, z = t_1]\} P(z = s) \\ &\quad + \sum_{s>t_2} \{E[y|\, z = s] - E[y|\, z = t_1]\} P(z = s). \end{aligned}$$

We will now compare the upper bounds in Equations (18) and ($A.51$).

The upper bound in Equation ($A.51$) is equal to the upper bound on $E[y(t_2)]$ minus the lower bound on $E[y(t_1)]$ in Equation (5). Proposition 1 shows that the bounds in Equation (6) is narrower than or equal to the bounds in Equation (5). Therefore, the upper bound on $E[y(t_2)]$ minus the lower bound on $E[y(t_1)]$ in Equation (6) is smaller than or equal to the upper bound in Equation ($A.51$). Furthermore, as the proof of Part (i) shows, the upper bound in Equation (18) is sharp; thereby, this upper bound is smaller than or equal to the upper bound on $E[y(t_2)]$ minus the lower bound on $E[y(t_1)]$ in Equation (6). Hence, the upper bound in Equation (18) is smaller than or equal to the upper bound in Equation ($A.51$).[17] *Q.E.D.*

[17] Alternatively, this can be proven by using the following facts: (1) For $s < t_1$, $UB(s, t_2) \leq E[y|\, z = t_2]$ because of Proposition 1, and $AT_1(t_1, s, t_2) \geq E[y|\, z = s]$ because $AT_1(\tau, s, t_2)$ is MTR in τ and $AT_1(s, s, t_2) = E[y|\, z = s]$. (2) For $t_1 \leq s \leq t_2$, $UB(s, t_2) \leq E[y|\, z = t_2]$, and $AT_1(t_1, s, t_2) \geq E[y|\, z = t_1]$ because $AT_1(\tau, s, t_2)$ is MTS. (3) For $s > t_2$, $AT_2(t_1, s, t_2) \geq E[y|\, z = t_1]$ because $AT_2(\tau, s, t_2)$ is MTS.

References

[1] Belzil, C., and J. Hansen (2002): "Unobserved Ability and the Return to Schooling," *Econometrica*, 70, 2075-2091.

[2] Card, D. (1999): "The Causal Effect of Education on Earnings," in *Handbook of Labor Economics*, Vol. 3A, ed. by O. Ashenfelter and D. Card. Amsterdam: North-Holland, 1801-1863.

[3] Chernozhukov, V., H. Hong, and E. Tamer (2007): "Estimation and Confidence Regions for Parameter Sets in Econometric Models," *Econometrica*, 75, 1243-1284.

[4] Kreider, B., and J. V. Pepper (2007): "Disability and Employment: Reevaluating the Evidence in Light of Reporting Errors," *Journal of the American Statistical Association,* 102, 432-441.

[5] Manski, C. F. (1995): *Identification Problems in the Social Sciences.* Cambridge, Mass.: Harvard University Press.

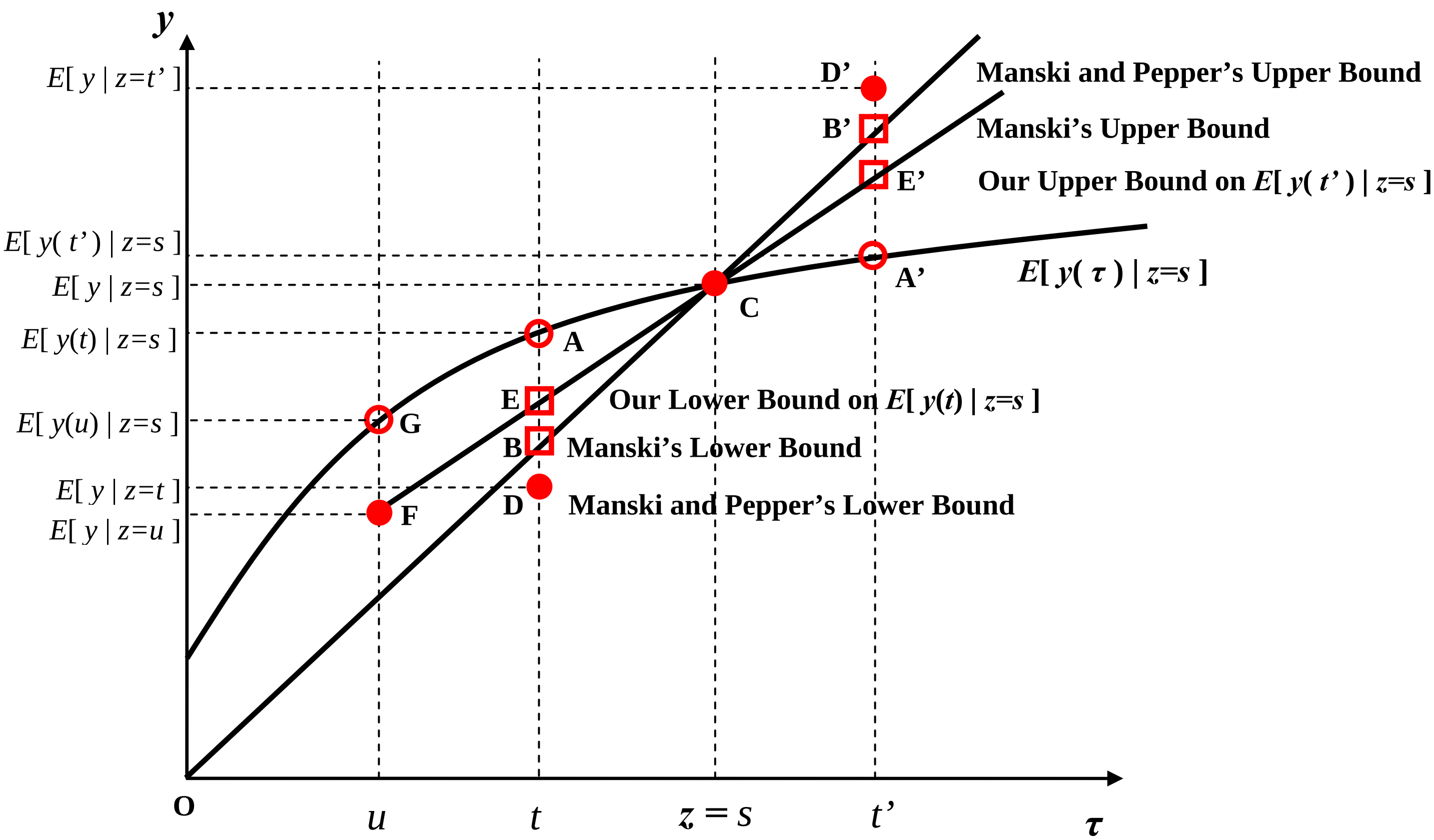
[6] Manski, C. F. (1997): "Monotone Treatment Response," *Econometrica*, 65, 1311-1334.

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

[8] Manski, C. F. and J. V. Pepper (2009): "More on Monotone Instrumental Variables," *The Econometrics Journal*, 12, S200-S216.

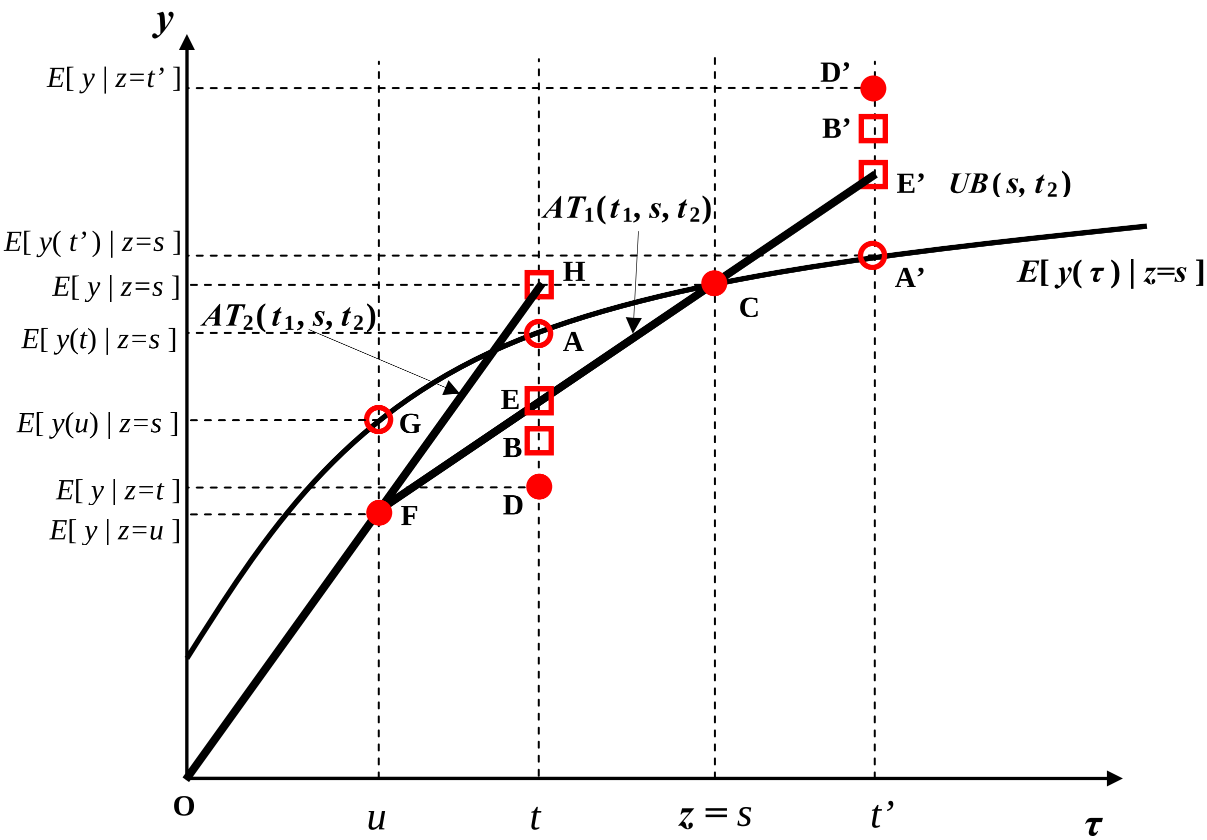
[9] Politis, D. N., J. P. Romano, and M. Wolf (1999): *Subsampling.* New York: Springer-Verlag.

Figure 1



Note:
Filled circles indicate the realized treatments and the conditional means of realized outcomes.
Open circles indicate the latent treatments and the conditional means of latent outcomes.
Open squares indicate the bounds on the conditional-mean treatment responses.

Figure 2



Note:
Filled circles indicate the realized treatments and the conditional means of realized outcomes.
Open circles indicate the latent treatments and the conditional means of latent outcomes.
Open squares indicate the bounds on the conditional-mean treatment responses.

Table 1: Mean Log(Wages) and Distribution of Schooling

z	*E* (*y* | *z*)	*P* (*z*)	Sample Size
7	2.228	0.006	7
8	2.541	0.016	20
9	2.449	0.019	23
10	2.515	0.017	21
11	2.637	0.020	24
12	2.722	0.404	493
13	2.954	0.070	86
14	2.975	0.087	106
15	3.062	0.038	46
16	3.247	0.181	221
17	3.266	0.038	46
18	3.386	0.051	62
19	3.358	0.025	31
20	3.393	0.029	35
Total		1	1221

Table 2: Upper and Lower Bounds on $E[y(t)]$: Assumptions of Concave-MTR and MTS

	Lower Bound on $E[y(t)]$				Upper Bound on $E[y(t)]$			
	Estimate	0.05 Subsampling Quantile	0.95 Subsampling Quantile	Bias Corrected Estimate	Estimate	0.05 Subsampling Quantile	0.95 Subsampling Quantile	Bias Corrected Estimate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
t								
9	2.449	2.126	2.809	2.449	2.953	2.929	3.112	2.931
10	2.558	2.440	2.936	2.509	2.954	2.931	3.119	2.931
11	2.665	2.606	2.991	2.616	2.957	2.933	3.120	2.934
12	2.767	2.734	3.012	2.734	2.961	2.935	3.121	2.940
13	2.834	2.801	3.040	2.806	3.000	2.947	3.121	2.991
14	2.871	2.847	3.062	2.843	3.041	2.956	3.128	3.044
15	2.905	2.880	3.082	2.879	3.084	2.961	3.140	3.100
16	2.937	2.910	3.101	2.915	3.108	2.965	3.158	3.131
17	2.946	2.922	3.106	2.923	3.129	2.973	3.185	3.155
18	2.953	2.926	3.110	2.931	3.150	2.979	3.216	3.180
19	2.953	2.929	3.111	2.931	3.170	2.983	3.237	3.205
20	2.953	2.929	3.112	2.931	3.189	2.984	3.262	3.230

Table 3: Upper and Lower Bounds on *E*[*y*(*t*)]:
Manski and Pepper's (2000) Bounds and Manski's (1997) Bounds

	Manski and Pepper's (2000) Bounds: Assumptions of MTR and MTS					
	Lower Bound on *E*[*y*(*t*)]			Upper Bound on *E*[*y*(*t*)]		
	Estimate	0.05 Subsampling Quantile	0.95 Subsampling Quantile	Estimate	0.05 Subsampling Quantile	0.95 Subsampling Quantile
	(1)	(2)	(3)	(4)	(5)	(6)
t						
9	2.449	2.126	2.809	2.952	2.904	3.005
10	2.513	2.025	2.835	2.954	2.903	3.005
11	2.631	2.350	2.912	2.958	2.908	3.011
12	2.711	2.652	2.771	2.963	2.912	3.016
13	2.834	2.745	2.923	3.072	2.982	3.161
14	2.844	2.765	2.920	3.083	2.979	3.181
15	2.876	2.789	2.973	3.138	2.993	3.303
16	2.937	2.882	2.991	3.262	3.169	3.358
17	2.940	2.879	2.999	3.279	3.036	3.512
18	2.953	2.900	3.005	3.385	3.174	3.601
19	2.951	2.902	3.005	3.359	3.041	3.704
20	2.952	2.904	3.005	3.393	3.087	3.697

	Manski's (1997) Bounds: Assumption of Concave-MTR					
	Lower Bound on *E*[*y*(*t*)]			Upper Bound on *E*[*y*(*t*)]		
	Estimate	0.05 Subsampling Quantile	0.95 Subsampling Quantile	Estimate	0.05 Subsampling Quantile	0.95 Subsampling Quantile
	(1)	(2)	(3)	(4)	(5)	(6)
t						
9	0.594	0.566	0.627	2.952	2.904	3.005
10	1.141	1.103	1.182	3.000	2.945	3.056
11	1.666	1.617	1.715	3.069	2.995	3.151
12	2.174	2.113	2.236	3.156	3.058	3.267
13	2.400	2.347	2.452	3.524	3.401	3.668
14	2.584	2.538	2.633	3.934	3.781	4.113
15	2.723	2.681	2.767	4.389	4.203	4.603
16	2.846	2.799	2.892	4.860	4.635	5.109
17	2.894	2.848	2.942	5.407	5.150	5.688
18	2.928	2.879	2.977	5.967	5.684	6.282
19	2.944	2.896	2.995	6.545	6.233	6.899
20	2.952	2.904	3.005	7.131	6.793	7.520

Table 4: Upper Bounds on Returns to Schooling: Assumptions of Concave-MTR and MTS

		Upper Bounds on $\Delta(t_1, t_2)$		
		Estimate	0.95 Subsampling Quantile	Bias Corrected Estimate
		(1)	(2)	(3)
t_1	t_2			
9	10	0.505	0.878	0.482
10	11	0.254	0.407	0.257
11	12	0.168	0.227	0.182
12	13	0.136	0.173	0.152
13	14	0.103	0.129	0.117
14	15	0.091	0.101	0.107
15	16	0.076	0.078	0.091
16	17	0.051	0.064	0.061
17	18	0.044	0.052	0.053
18	19	0.034	0.043	0.042
19	20	0.032	0.039	0.040
12	16	0.331	0.365	0.386
Average Effect		0.083	0.091	0.096

Table 5: Upper Bounds on Returns to Schooling:
Manski and Pepper's (2000) Bounds and Manski's (1997) Bounds

		Upper Bounds on $\Delta(t_1, t_2)$			
		Manski and Pepper's (2000) Bounds		Manski's (1997) Bounds	
		Estimate	0.95 Subsampling Quantile	Estimate	0.95 Subsampling Quantile
		(1)	(2)	(3)	(4)
t_1	t_2				
9	10	0.505	0.839	1.500	1.528
10	11	0.445	0.941	1.023	1.050
11	12	0.332	0.619	0.789	0.817
12	13	0.361	0.470	0.705	0.734
13	14	0.249	0.380	0.656	0.685
14	15	0.295	0.484	0.627	0.658
15	16	0.386	0.514	0.608	0.639
16	17	0.341	0.585	0.601	0.632
17	18	0.445	0.682	0.597	0.628
18	19	0.406	0.754	0.595	0.627
19	20	0.442	0.753	0.594	0.627
12	16	0.551	0.662	2.430	2.555
Average Effect		0.138	0.165	0.608	0.639