

Discussion Paper No. 1011

**INTERGENERATIONAL TRANSMISSION
OF EDUCATION IN JAPAN:
NONPARAMETRIC BOUNDS ANALYSIS
WITH MULTIPLE TREATMENTS**

Nobuyoshi Kikuchi

September 2017

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

Intergenerational Transmission of Education in Japan: Nonparametric Bounds Analysis with Multiple Treatments*

Nobuyoshi Kikuchi[†]

September 19, 2017

Abstract

This paper investigates the intergenerational effects of education in Japan using a nonparametric bounds approach. The educational levels of parents are considered key factors in explaining children's educational success. Nevertheless, the literature has not reached consensus on the causal effects of parents' education on their child's schooling. This is because both parents' and the child's schooling depend on unobserved heterogeneity. Moreover, the strong positive correlation of the mother's and father's schooling makes it difficult to separate the effects of each parent's schooling, making it unclear how to control spousal schooling in the analysis. Therefore, this paper estimates a set of semi-ordered vectors of both parents' schooling as an application of the nonparametric bounds method with multiple treatments. It thus derives bounds depending on relatively weak semi-monotonicity assumptions on treatment response, selection, and instrumental variables. A combination of these assumptions provides informative bounds on the average treatment effect of both parents' education on their child's schooling. The main results show that the tightest lower bounds suggest the positive causal effects of parents' schooling, but the tightest upper bounds on the effects are lower than the point estimates that rely on the assumptions of an exogenous selection for parents' schooling. These results suggest that simple regressions overestimate the true causal effect of parents' education.

JEL Classification: I24, I26, J62, C14

Keywords: Intergenerational transmission, returns to education, nonparametric bounds, multiple treatments

*This paper is based on part of my dissertation at the University of Tokyo and on an awarded paper at the 16th Labor Economics Conference at Nihon University. I would like to thank my advisor, Hidehiko Ichimura, as well as committee members Hideo Akabayashi, Yuji Genda, Hideo Owan, and Yasuyuki Sawada for their valuable advice and comments. I am also grateful to Takeshi Aida, Daiji Kawaguchi, Hirokazu Matsuyama, Tsunao Okumura, as well as the seminar participants at the University of Tokyo, and the participants at the 16th Labor Economics Conference at Nihon University, and the 2013 Japanese Economic Association Autumn Meeting at Kanagawa University for their helpful comments and suggestions on earlier versions of this paper. Any remaining errors are my own. This research was financially supported by the Grant-in-Aid for Japan Society for the Promotion of Science Fellows (DC2), Grant Number: 12J10408, and the Grant-in-Aid for Research Activity Start-up No. 16H06925 from the Japan Society for the Promotion of Science.

[†]Institute of Social and Economic Research, Osaka University. Address: 6-1, Mihogaoka, Ibaraki, Osaka 567-0047, Japan. Tel: +81-6-6879-8582, E-mail: kikuchi@iser.osaka-u.ac.jp.

1 Introduction

Parent's educational level is key to explaining children's educational success. If more educated parents provide better nurturing environments and invest in their children (e.g., Guryan, Hurst, and Kearney, 2008; Carneiro, Meghir, and Parey, 2013; Del Bono et al., 2016), an improvement of the education levels in one generation might also affect subsequent generations. This causal chain implies potential policy interventions to address inequality transmitted across generations by encouraging education. Therefore, numerous studies in social science, including economics, have paid attention to the significance of the intergenerational transmission of education and empirically found a positive correlation between the parents' and their child's schooling.¹

Unfortunately, the literature faces at least two important difficulties in quantifying the causal impacts of parents' education on their child's schooling. First, the schooling choice depends on unobserved heterogeneity, such as preferences, genetic endowments, or other personal attributes acquired during earlier life stages. If such unobservables determine both parents' and their child's schooling, the positive intergenerational correlation does not necessarily reflect a causal relation. To estimate the causal impact, the literature depends on identification strategies that use identical twins, adopted children, and instrumental variables (IV). However, as Holmlund, Lindahl, and Plug (2011) point out, different strategies lead to different results, and thus the literature does not reach a consensus on the magnitude of the causal effect.

Second, a strong positive correlation between the father's and mother's schooling makes it difficult to interpret the effect of each parent separately. Therefore, it is unclear whether spouses' education should be controlled in the analysis, and if so, how. Most previous studies focus on the effect of a single treatment of each parent. The estimated effect might not be the causal effect of interest, because the effect of parental schooling includes both a direct transfer from the given parent and an indirect transfer from the other parent due to assortative mating effects. The regression approach allows controlling for a variable of spousal education as an additional regressor, but usually ignores the endogeneity of this variable because it is difficult to find different IVs for each parent. Additionally, the regression model relies on an implicit linear separability assumption of both parents' years of schooling.²

The objective of this paper is to estimate the causal effect of a combination of parents' schooling on their child's schooling in Japan using a nonparametric bounds approach. To do so, I consider a set of treatment vectors of mother's and father's education that takes semi-ordered multiple values. I rely on monotonicity and semi-monotonicity assumptions to acquire informative bounds, such as the monotone treatment response (MTR) of Manski (1997), monotone treatment selection (MTS), and monotone instrumental variables (MIV) of Manski and Pepper (1998, 2000), which are weaker than the assumptions required to obtain point estimates. Due to the nonparametric bounds with multiple treatments framework, this paper identifies the causal effect of parents' educational attainment on the child's schooling, allowing for the separate effects of maternal and paternal education.

¹See Black and Devereux (2011) for a review.

²See Holmlund, Lindahl, and Plug (2011) for a review.

The main results show that the tightest bounds on the average effects of parents' schooling are positive, but lower than the point estimates that rely on an exogenous selection of parents' schooling. These results suggest that the regression estimates, mainly provided in previous Japanese studies, have an upward bias from the true causal effects of parents' schooling. This paper shows that the combination of weak monotone assumptions is too conservative to identify a point estimate, but is still credible and produces informative bounds on the effects of the parents' education on the child's schooling even when considering spousal schooling.

This paper provides new insights into the intergenerational effects of education in controlling for spousal education. In previous studies that utilize the regression approach, Chevalier et al. (2013), for instance, estimate the separate causal effects of parental schooling using a raising of the minimum school leaving age in the U.K. as an instrument. Their identification depends on changes in parental schooling induced by the compulsory schooling reform, only analyzing a local treatment effect of the parents at the margin.³ Oreopoulos, Page, and Stevens (2006) propose an alternative way to control for spousal education. They use the sum of mother's and father's completed education as an explanatory variable, instead of including each parent's education separately. This allows to directly estimate the effect of a one-year increase in either parent's schooling. This specification implicitly assumes that the effects of father's and mother's education are similar, although it is unclear whether this assumption is satisfied in general. This paper's specification of multiple treatments avoids the discussion on the most adequate way to control for spousal effects, but directly evaluates the effects of a combination of both parents' schooling. Furthermore, the nonparametric bounding approach is general enough to identify the average treatment effect (ATE), thus allowing for the heterogeneous effects of each parent's schooling.

This paper relates to the literature using a nonparametric bounding approach for the returns to education and skills (e.g., Ginther, 2000; Manski and Pepper, 2000; Gonzalez, 2005; Giustinelli, 2011; Huang, van den Brink, and Groot, 2012; Okumura and Usui, 2014; Mariotti and Meinecke, 2015). For example, De Haan (2011) first applies a nonparametric bounding method to examine the effects of parent's schooling on the child's schooling using U.S. data. She considers ordered multivalued treatments of each parent's educational level and shows that the mother's or father's college degree has a positive effect on the child's schooling. Moreover, the estimated bounds are significantly different from zero, but substantially lower than the ordinary least squares (OLS) estimates. This paper confirms these findings with semi-ordered multivalued treatments of a pair of both parents' educational levels using Japanese data.

This paper also contributes to the intergenerational transmission of inequality literature by analyzing the causal impacts of parental schooling in Japan. Previous empirical studies have found a significant positive correlation between parents' and child's schooling in Japan, but have paid limited attention to the causal relationship.⁴ Researchers require micro data that contain twins, adopted chil-

³Dickson, Gregg, and Robinson (2016) also use an exogenous change in the minimum school leaving age in the U.K. to identify the effects on the child's educational attainment, but mainly focus on the direct impacts of the leaving age change.

⁴For example, see Tachibanaki (1988) and Yamada (2011). See also Kariya (2001) as an example of sociological

dren, or plenty of information about individual characteristics to estimate the causal effects. However, it is difficult to access such data sets in Japan.⁵ This limited accessibility to informative micro data prevents researchers from controlling individual heterogeneity and magnifies the inevitable concern for point identification under strong assumptions on treatment selection. For this reason, I derive sets of nonparametric bounds under relatively weak and testable assumptions that share broad consensus in the literature.

The remainder of this paper is organized as follows. Section 2 defines the parameter of interest and describes the identification assumptions. Section 3 explains the data set. Section 4 presents and discusses the results. Finally, Section 5 draws the conclusions.

2 Method

This paper employs a nonparametric technique to obtain the bounds on the causal effect of parents' education on their child's education, following the setup of Manski and Pepper (2000). There is a probability space (J, Ω, P) of individuals. Each member j of population J has observable covariates $x_j \in X$ and a response function $y_j(\cdot) : T \rightarrow Y$, mapping the mutually exclusive and exhaustive treatments $t \in T$ into real-valued outcomes $y_j(t) \in Y$, where treatment T is the level of schooling the parent completed, thus assumed to be an ordered set. Y represents the years of schooling or university graduation of child j , and $x = (w, v)$, $X = W \times V$. Each value of (w, v) defines an observable sub-population of children. I assume that the outcome space Y has the greatest lower bound $Y_{min} \equiv \min Y$ and the least upper bound $Y_{max} \equiv \max Y$.⁶

Child j has a realized treatment $z_j \in T$ and a realized outcome $y_j \equiv y_j(z_j)$, both observable. The latent outcomes $y_j(t), t \neq z_j$ are not observable. A researcher learns the distribution $P(x, z, y)$ of covariates, realized treatments, and realized outcomes by observing a random sample of the population. To simplify the notation, subscript j will be dropped except when it required.

The parameter of interest is the ATE between the parent's education level $t \in T$ and $t' \in T$, *s.t.* $t > t'$:

$$ATE(t, t') \equiv E[y(t) - y(t')] = E[y(t)] - E[y(t')]. \quad (1)$$

The linearity of expectations provides the second equality. The empirical problem comes down to learning $E[y(\cdot)]$ from the empirical evidence of the distribution $P(x, z, y)$ of covariates, realized treatments, and realized outcomes with assumptions. The fact that $E[y|z = t] = E[y(t)|z = t]$ and the

research on the relationship between educational attainment and family background.

⁵See Oshio and Seno (2007) for a review of studies on the economics of education in Japan.

⁶When I analyze the effects on the child's years of schooling, I assume $Y_{min} = 9$ and $Y_{max} = 18$ years of schooling from the construction of the data set.

law of iterated expectations gives

$$E[y(t)] = E[y|z = t] \cdot Pr(z = t) + E[y(t)|z \neq t] \cdot Pr(z \neq t). \quad (2)$$

With a data set, researchers can estimate $E[y|z = t]$, $Pr(z = t)$, $Pr(z \neq t)$, but cannot estimate $E[y(t)|z \neq t]$ without adding any assumptions. Manski (1989, 1990) shows that worst-case bounds on $E[y(t)]$ can be identified if the support of the dependent variable is bounded:

$$\begin{aligned} E[y|z = t] \cdot Pr(z = t) + Y_{min} \cdot Pr(z \neq t) \\ \leq E[y(t)] \leq \\ E[y|z = t] \cdot Pr(z = t) + Y_{max} \cdot Pr(z \neq t). \end{aligned} \quad (3)$$

The width of bounds is $(Y_{max} - Y_{min}) \cdot Pr(z \neq t)$, thus the bounds are informative if $Pr(z \neq t) < 1$. These bounds on $E[y(\cdot)]$ provide the lower (upper) bound on the $ATE(t, t')$ by subtracting the upper (lower) bound on $E[y(t')]$ from the lower (upper) bound on $E[y(t)]$. Hence, the identification region $H\{\cdot\}$ for the ATE is the interval:

$$\begin{aligned} H\{E[y(t)] - E[y(t')]\} \\ = \{E[y|z = t] \cdot Pr(z = t) + Y_{min} \cdot Pr(z \neq t) - E[y|z = t'] \cdot Pr(z = t') + Y_{max} \cdot Pr(z \neq t'), \\ E[y|z = t] \cdot Pr(z = t) + Y_{max} \cdot Pr(z \neq t) - E[y|z = t'] \cdot Pr(z = t') + Y_{min} \cdot Pr(z \neq t')\}. \end{aligned}$$

This interval contains the value zero. Its width is $(Y_{max} - Y_{min}) \cdot [Pr(z \neq t) + Pr(z \neq t')]$. The worst-case bounds are important at first, because they include all results that depend on different assumptions about $E[y(t)|z \neq t]$. However, the worst-case bounds are too wide to obtain an informative result in this paper. Therefore, some additional assumptions are necessary to tighten the bounds. In the remainder of this section, I explain how assumptions tighten these bounds. I consider the exogenous treatment selection (ETS) assumption and the following three assumptions: (1) the MTR, (2) MTS and (3) MIV. Additionally, I apply the semi-monotonicity version of these three assumptions.

2.1 Exogenous Treatment Selection

The ETS assumes that the realized treatment z is statistically independent of the latent outcomes $y(\cdot)$, and can be expressed as:

For each $t \in T$, and all $(u_1, u_2) \in T \times T$

$$E[y(t)|z = u_1] = E[y(t)|z = u_2], \quad (4)$$

or

$$E[y(t)] = E[y|z = t]. \quad (5)$$

Under the ETS assumption, the parameter of interest becomes: $E[y|z = t] - E[y|z = t']$. The ETS assumption or covariates x conditional version of it, $E[y(t)|x] = E[y|z = t, x]$, is implicitly imposed in most empirical literature. This assumption implies that treatments are randomly assigned to the population and excludes the existence of heterogeneous treatment effects with essential heterogeneity (Heckman, Urzua, and Vytlačil, 2006). However, this does not seem to be appropriate in the intergenerational transmission of education. For example, one could consider a situation where parents received different educational levels because they differ substantially from one another due to attributes or skills acquired in their earlier life stages, but this heterogeneity is unobservable by researchers. If the heterogeneity also determines the child's schooling, the realized treatment is endogenous in the analysis. Moreover, it is difficult to use micro data with sufficient information on individual characteristics in Japan. In this case, the researchers cannot fully control for heterogeneity, and thus the endogenous selection of parents' education violates the ETS assumption.

However, one advantage of assuming the ETS is that it yields point identification. The ETS estimates are equivalent to the coefficients obtained by regression of the child's schooling on dummy variables for each parental educational level, with no other covariates. This paper shows the point estimates under the ETS assumption for comparison with bounds estimates.

2.2 Monotone Treatment Response

Manski (1997) proposes an MTR assumption that states response functions are weakly increasing. That is,

For each $j \in J$ and all $(t_1, t_2) \in T \times T$ s.t. $t_1 \leq t_2$,

$$t_1 \leq t_2 \Rightarrow y_j(t_1) \leq y_j(t_2). \quad (6)$$

The MTR assumes that an increase in parents' schooling does not decrease their child's schooling. The validity of this assumption is suggested in the literature. For instance, the human capital theory (Becker, 1975; Becker and Tomes, 1979; Solon, 1999) suggests a positive impact of increasing parents' schooling on the child's education. The literature on the returns to schooling shows that an additional year of schooling has positive effects on both pecuniary and non-pecuniary outcomes (Card, 1999; Oreopoulos and Salvanes, 2011). These advantages are transmitted to children from several parental

investment channels (Carneiro, Meghir, and Parey, 2013), which include time spent by the parents with their children (Guryan, Hurst, and Kearney, 2008; Del Bono et al., 2016). It is worth noting that the MTR assumption does not exclude zero effect of parental education, but allows that the positive correlation between parents' and child's schooling has no causal relationship.

Manski (1997) shows that the MTR assumption tightens the worst-case bound:⁷

$$\begin{aligned}
t < z_j &\Rightarrow Y_{min} \leq y_j(t) \leq y_j, \\
t = z_j &\Rightarrow y_j(t) = y_j, \\
t > z_j &\Rightarrow y_j \leq y_j(t) \leq Y_{max}.
\end{aligned} \tag{7}$$

Thus,

$$\begin{aligned}
E[y|z \leq t] \cdot Pr(z \leq t) + Y_{min} \cdot Pr(z > t) \\
\leq E[y(t)] \leq \\
Y_{max} \cdot Pr(z < t) + E[y|z \geq t] \cdot Pr(z \geq t).
\end{aligned} \tag{8}$$

2.3 Monotone Treatment Selection

Manski and Pepper (1998, 2000) weaken equation 4 to an inequality and yield the MTS assumption:

For each $t \in T$ and all $(u_1, u_2) \in T \times T$ s.t. $u_1 \leq u_2$

$$u_1 \leq u_2 \Rightarrow E[y(t)|z = u_1] \leq E[y(t)|z = u_2]. \tag{9}$$

If the MTS assumption holds, then,

$$\begin{aligned}
u < t &\Rightarrow Y_{min} \leq E[y(t)|z = u] \leq E[y|z = t], \\
u = t &\Rightarrow E[y(t)|z = u] = E[y|z = t], \\
u > t &\Rightarrow E[y|z = t] \leq E[y(t)|z = u] \leq Y_{max}.
\end{aligned} \tag{10}$$

It follows that

⁷For a full derivation of the MTR bounds, see Manski (1997).

$$\begin{aligned}
Y_{min} \cdot Pr(z < t) + E[y|z = t] \cdot Pr(z \geq t) \\
\leq E[y(t)] \leq \\
E[y|z = t] \cdot Pr(y \leq t) + Y_{max} \cdot Pr(z > t).
\end{aligned} \tag{11}$$

Manski and Pepper (1998, 2000) also show that combining the MTR and MTS assumptions can have substantial identifying power.⁸ Let the MTR and MTS assumptions 6 and 9 hold. Then, the bounds reduce to

$$\begin{aligned}
E[y|z < t] \cdot Pr(z < t) + E[y|z = t] \cdot Pr(z \geq t) \\
\leq E[y(t)] \leq \\
E[y|z = t] \cdot Pr(z \leq t) + E[y|z > t] \cdot Pr(z > t).
\end{aligned} \tag{12}$$

According to the interpretation of Manski and Pepper (1998, 2000), the MTS assumption indicates that parents who select higher levels of schooling have a weakly higher mean response function than those who select lower levels. This implies that a genetic transmission of unobserved abilities from parents to children provides the correlation of the educational levels. Researchers commonly assume that unmeasured inherent differences are a key source of the endogeneity of education. Therefore, the MTS assumption is consistent with arguments in the literature. Furthermore, Manski and Pepper (1998, 2000) suggest a simple test for the validity of the joint MTR-MTS assumption. Under this assumption:

For all $(u_1, u_2) \in T \times T$, s.t. $u_1 \leq u_2$,

$$E[y|z = u_1] = E[y(u_1)|z = u_1] \leq E[y(u_1)|z = u_2] \leq E[y(u_2)|z = u_2] = E[y|z = u_2]. \tag{13}$$

The MTS assumption implies the first inequality and the MTR assumption gives the second. Under this hypothesis, $E(y|z = u)$ must be a weakly increasing function of u . Hence the researcher should reject the hypothesis if the mean outcomes of the child's schooling are not weakly increasing in the realized parental education levels.

2.4 Monotone Instrumental Variables

The recent literature on causal intergenerational schooling effects has moved from estimating OLS under the ETS assumption, and relies on alternative identification assumptions, such as an IV assumption (Holmlund, Lindahl, and Plug, 2011). Using the notation in this paper, the familiar mean-independence

⁸For a full derivation of the MTS and the MTR-MTS bounds, see Manski and Pepper (1998, 2000).

form of the IV assumption can be written as follows,

For each $t \in T$ and all $(u_1, u_2) \in V \times V$,

$$E[y(t)|v = u_1] = E[y(t)|v = u_2]. \quad (14)$$

If the treatment effects are homogeneous, the IV identifies the parameter of interest. Even if the effects are heterogeneous, researchers can recover the parameter of interest or a related interpretable one using an IV with additional assumptions on the selection model.⁹ Unfortunately, it is difficult to find an IV that satisfies the mean-independence conditions. Instead of assuming mean-independence, Manski and Pepper (1998, 2000) introduce a weaker assumption to replace the equality in 14 by an inequality yielding a weakly monotone relation between the IV and the mean response function.

For each $t \in T$ and all $(u_1, u_2) \in V \times V$,

$$u_1 \leq u_2 \Rightarrow E[y(t)|v = u_1] \leq E[y(t)|v = u_2]. \quad (15)$$

In case the MIV assumption holds, the bounds are:

$$\begin{aligned} \sum_{u \in V} Pr(v = u) \cdot \left\{ \max_{u_1 \leq u} [E[y|z = t, v = u_1] \cdot Pr(z = t|v = u_1) + Y_{min} \cdot Pr(z \neq t|v = u_1)] \right\} \\ \leq E[y(t)] \leq \\ \sum_{u \in V} Pr(v = u) \cdot \left\{ \min_{u_2 \geq u} [E[y|z = t, v = u_2] \cdot Pr(z = t|v = u_2) + Y_{max} \cdot Pr(z \neq t|v = u_2)] \right\}. \end{aligned} \quad (16)$$

The MIV and the MTR-MTS assumptions make distinct contributions to the identification. Manski and Pepper (1998, 2000) suggest that combining these assumptions yields particularly interesting bounds.¹⁰

⁹See Imbens and Angrist (1994), and Heckman and Vytlacil (2007) for further details.

¹⁰For a full derivation of the MIV and the MTR-MTS-MIV bounds, see Manski and Pepper (1998, 2000). Following the literature, the MTS indicates a conditional MTS assumption when combining with (semi-)MIV assumptions. See Laff ers (2013) for more details.

$$\begin{aligned}
& \sum_{u \in V} Pr(v = u) \cdot \left\{ \max_{u_1 \leq u} [E[y|z < t, v = u_1] \cdot Pr(z < t|v = u_1)] \right. \\
& \qquad \qquad \qquad \left. + E[y|z = t, v = u_1] \cdot Pr(z \geq t|v = u_1) \right\} \\
& \qquad \qquad \qquad \leq E[y(t)] \leq \\
& \sum_{u \in V} Pr(v = u) \cdot \left\{ \min_{u_2 \geq u} [E[y|z = t, v = u_2] \cdot Pr(z \leq t|v = u_2)] \right. \\
& \qquad \qquad \qquad \left. + E[y|z > t, v = u_2] \cdot Pr(z > t|v = u_2) \right\}.
\end{aligned} \tag{17}$$

2.5 Semi-Monotone Instrumental Variables

Manski and Pepper (1998) also suggest ways to combine multiple scalars of IV or MIV assumptions. Here, I use two MIVs simultaneously to tighten the bounds following De Haan (2011). To do so, I assume that the pair of MIVs, (v^a, v^b) , is a two-dimensional semi-monotone instrumental variable (SMIV):

For each $t \in T$ and all $[(u_1^a, u_1^b), (u_2^a, u_2^b)] \in (V^a \times V^b) \times (V^a \times V^b)$ s.t. $u_1^a \leq u_2^a$ and $u_1^b \leq u_2^b$

$$\Rightarrow E[y(t)|(v^a, v^b) = (u_1^a, u_1^b)] \leq E[y(t)|(v^a, v^b) = (u_2^a, u_2^b)]. \tag{18}$$

In this assumption, a pair of MIVs is assumed to be semi-ordered rather than ordered, because it includes some value pairs that are not ordered.¹¹ The MTR-MTS-MIV bounds in 17 can be extended to the MTR-MTS-SMIV bounds if the maxima and minima operations are taken over ordered pairs of values.

2.6 Semi-Monotone Treatment Response

The previous studies of the effect of parents' schooling on child's schooling have focused on the identification of a single treatment of one parent and estimated the effect of mother's and father's schooling separately. By contrast, children are affected by a combination of both parents' schooling in practice. This indicates that the treatments of parents' schooling are multiple for the child. In this paper, I analyze effects of both parents' schooling simultaneously within the framework of the nonparametric bounding approach.¹²

¹¹For example, consider that $u_1^a < u_2^a$ and $u_1^b > u_2^b$. This does not predict the ordering of $E[y(t)|(v^a, v^b) = (u_k^a, u_k^b)]$, $k = 1, 2$.

¹²The matching approach is an alternative method analyzing multiple treatments. For example, see Imbens (2000) and Lechner (2001). See also Flores and Mitnik (2013) as an example of applying the difference-in-difference method for multiple treatments. Frölich (2004) is a comprehensive survey of the multiple treatments analysis.

Manski (1997) proposes the assumption of a semi-monotone treatment response (SMTR), where a semi-ordered vector of multivalued treatments \mathbf{T} exists. This suggests that the discussion based on the single ordered treatment of a parent's schooling can be extended to a semi-ordered vector of multiple treatments, which is a combination of each parent's schooling. This paper considers the two-dimensional treatment vectors of mother's and father's schooling, $(t^M, t^F) = \mathbf{t} \in \mathbf{T}$. In this case, the SMTR assumption states:

For each $j \in J$ and all $(\mathbf{t}_1, \mathbf{t}_2) \in \mathbf{T} \times \mathbf{T}$ s.t. $\mathbf{t}_1 \leq \mathbf{t}_2$,

$$\mathbf{t}_1 \leq \mathbf{t}_2 \Rightarrow y_j(\mathbf{t}_1) \leq y_j(\mathbf{t}_2), \quad (19)$$

where I define $\mathbf{t}_1 \leq \mathbf{t}_2$ if and only if $t_1^M \leq t_2^M$ and $t_1^F \leq t_2^F$.

The analysis under the MTR assumption still holds under the assumption of SMTR, except for the case where treatments are not ordered:¹³

$$\begin{aligned} \mathbf{t} < \mathbf{z}_j &\Rightarrow Y_{min} \leq y_j(\mathbf{t}) \leq y_j, \\ \mathbf{t} = \mathbf{z}_j &\Rightarrow y_j(\mathbf{t}) = y_j, \\ \mathbf{t} > \mathbf{z}_j &\Rightarrow y_j \leq y_j(\mathbf{t}) \leq Y_{max}, \\ \mathbf{t} \not\leq \mathbf{z}_j &\Rightarrow Y_{min} \leq y_j(\mathbf{t}) \leq Y_{max}, \end{aligned} \quad (20)$$

where $\not\leq$ denotes that \mathbf{t} and \mathbf{z} are not ordered. Then, I have the following SMTR bounds:

$$\begin{aligned} E[y|\mathbf{z} \leq \mathbf{t}] \cdot Pr(\mathbf{z} \leq \mathbf{t}) + Y_{min} \cdot Pr(\mathbf{z} > \mathbf{t}) + Y_{min} \cdot Pr(\mathbf{z} \not\leq \mathbf{t}) \\ \leq E[y(\mathbf{t})] \leq \\ Y_{max} \cdot Pr(\mathbf{z} < \mathbf{t}) + E[y|\mathbf{z} \geq \mathbf{t}] \cdot Pr(\mathbf{z} \geq \mathbf{t}) + Y_{max} \cdot Pr(\mathbf{z} \not\leq \mathbf{t}). \end{aligned} \quad (21)$$

2.7 Semi-Monotone Treatment Selection

The SMIV assumption of Manski and Pepper (1998) suggests an assumption on selection with a semi-ordered vector of multivalued treatment. The semi-monotone treatment selection (SMTS) assumption is the special SMIV assumption that the MIVs, $\mathbf{v} = (v^a, v^b)$ are the realized treatments \mathbf{z} . With the SMTS assumption, the analysis under the MTS assumption is also available, except for the case where treatments are not ordered. Therefore, I can obtain the SMTR-SMTS bounds by combining the SMTR and SMTS assumptions, as well as the analysis on the single treatment. Let the SMTR and SMTS

¹³For a full derivation of the SMTR bounds, see Manski (1997).

assumptions hold.¹⁴ Then,

$$\begin{aligned}
\mathbf{u} < \mathbf{t} &\Rightarrow Y_{min} \leq E[y|\mathbf{z} = \mathbf{u}] \leq E[y(\mathbf{t})|\mathbf{z} = \mathbf{u}] \leq E[y|\mathbf{z} = \mathbf{t}], \\
\mathbf{u} = \mathbf{t} &\Rightarrow E[y(\mathbf{t})|\mathbf{z} = \mathbf{u}] = E[y|\mathbf{z} = \mathbf{t}], \\
\mathbf{u} > \mathbf{t} &\Rightarrow E[y|\mathbf{z} = \mathbf{t}] \leq E[y(\mathbf{t})|\mathbf{z} = \mathbf{u}] \leq E[y|\mathbf{z} = \mathbf{u}] \leq Y_{max}, \\
\mathbf{u} \notin \mathbf{t} &\Rightarrow Y_{min} \leq y(\mathbf{t}) \leq Y_{max}.
\end{aligned} \tag{22}$$

It follows that

$$\begin{aligned}
E[y|\mathbf{z} < \mathbf{t}] \cdot Pr(\mathbf{z} < \mathbf{t}) + E[y|\mathbf{z} = \mathbf{t}] \cdot Pr(\mathbf{z} \geq \mathbf{t}) + Y_{min} \cdot Pr(\mathbf{z} \notin \mathbf{t}) \\
\leq E[y(\mathbf{t})] \leq \\
E[y|\mathbf{z} = \mathbf{t}] \cdot Pr(\mathbf{z} \leq \mathbf{t}) + E[y|\mathbf{z} > \mathbf{t}] \cdot Pr(\mathbf{z} > \mathbf{t}) + Y_{max} \cdot Pr(\mathbf{z} \notin \mathbf{t}).
\end{aligned} \tag{23}$$

These bounds can be viewed as a natural extension of the MTR-MTS bounds. The bounds can then be combined with the MIV and SMIV assumptions, as well as the single treatment bounds since the assumptions are not mutually exclusive.

3 Data

The main analysis data are the Japanese General Social Surveys (JGSS).¹⁵ The JGSS are repeated cross-sectional data for men and women aged 20–89 on each survey date. This paper uses the surveys conducted in 2000, 2001, 2002, 2003, 2005, 2006, 2008, and 2010, and pools the respondents from all waves. From the pooled original data, I exclude observations by following four steps.

First, I restrict the sample by age, excluding the respondents below 25 years old on the survey year because they might not have completed their academic schooling. Second, I drop the individuals who answered that their father or mother was absent at the age of 15, because I am unable to control unobservable characteristics if parental structures are substantially different between those with single

¹⁴For a full derivation of MTR-MTS bounds and a further discussion on semi-monotonicity assumptions, see Manski (1997) and Manski and Pepper (1998).

¹⁵The Japanese General Social Surveys (JGSS) are designed and carried out by the JGSS Research Center at Osaka University of Commerce (Joint Usage / Research Center for Japanese General Social Surveys accredited by Minister of Education, Culture, Sports, Science and Technology), in collaboration with the Institute of Social Science at the University of Tokyo.

parents and those with couples of parents. Third, I only use the observations for those born after 1940, because the current educational system was enacted after 1947. Finally, I use the individuals whose observational characteristics are available. Therefore, the remaining sample contains 13669 individuals, born in 1940–1984 and being 25–69 on the date of the survey.

For the educational attainment of the respondents (children) and their parents, I use information about the level of the last school attended and assign the standard years of schooling in Japan. For children, completion of their last schooling is available. Following Tanaka (2008), I reduce the number of years of education by one year from the standard years for those who dropped out before graduating the last attended school. To examine the robustness of the definition of the outcome variable, I also use an indicator of university graduation.

In the analysis, I consider the following four schooling levels for parents: (1) junior high school denotes that the parent completes compulsory schooling or lower secondary education, (2) high school denotes that the parent completes high school or upper secondary education, (3) some college denotes that the parent completed a two-year college or some years of higher education, and (4) university denotes that the parent has a bachelor’s degree. I also use schooling of a parent as an MIV. When I obtain bounds on the effect of the mother’s schooling, I use the indicator of father’s university graduation as an MIV. For father’s schooling, the indicator of mother’s college graduation takes the role of an MIV.

Under the SMIV assumption, I use two additional dummy variables as MIVs. The first additional MIV is a dummy variable of father’s “regular” worker status that takes the value one if the father was an executive of a company or an employee with non-terminable contract (*joji-koyo no rodosha*) when the child was 15. A father’s more stable working status implies more stable parental income during the child’s adolescent years. It is natural to assume that a child with more parental income is less likely to face credit constraints for the decision of additional schooling. The second additional MIV is a dummy variable that takes the value one if the child was born after 1975. Taking the birth cohort as an MIV is equivalent to assuming that children born in more recent years have weakly higher mean outcome functions than others. This assumption can be attributed to the decreasing trends of number of children, non-decreasing trends in the number of schools per child, and stable non-decreasing trends of post-compulsory school or university enrollment rates in recent years in Japan. Table 1 shows some summary statistics.

4 Results

Before presenting the results of the bounds on the ATE, I test that the combined assumption of the MTR and the MTS cannot be rejected from the analysis data. Table 2 reports mean schooling outcomes of children by the educational level of parents. All variables show that the MTR-MTS assumption is not rejected, since the average values of outcome variables are weakly increasing both in the levels of

the mother's and father's schooling.

4.1 Single Treatment Effects on Child's Years of Schooling

I begin by showing the results of a single treatment of one parent to confirm the findings in De Haan (2011). Table 3 provides bounds on the ATE of the mother's and father's schooling on child's years of schooling. Following Manski and Pepper (1998, 2000), Gonzalez (2005), and Giustinelli (2011) all tables below also report 95% confidence intervals obtained by the percentile bootstrap method with 3000 replications.

The column of the ETS assumption reports point estimates of the effects. The effect of the change of mother's schooling from junior high school to high school shows an increase in the child's schooling by about 1.50 years. The corresponding effect of the father's schooling increases child's schooling by about 1.30 years. The change in mother's schooling from high school to some college increases the child's schooling by about 1.02 years. Increasing father's schooling from high school to some college, increases the child's schooling by about 0.74 years. These results suggest stronger effects of maternal over paternal education on the child's years of schooling. However, the ETS results of a bachelor's degree seem to be mixed. Compared to college degrees, mother's bachelor degree has smaller effects than the father's bachelor degree. Increasing the mother's schooling from junior high school to a bachelor's degree, increases the child's schooling by about 2.95 years. This is larger than the corresponding effect of father. Increasing father's schooling from junior high school to bachelor's degree, increases the child's schooling by about 2.65 years.

From columns (2) and (7), the worst-case bounds are extremely wide. Imposing the MTR assumption increases significantly the lower bounds by the definition of the assumption that implies non-negative effects (columns (3) and (8)). The MTS assumption allows tightening upper and lower bounds compared to the worst-case ones, but is not powerful enough to acquire information on the treatment effect (columns (4) and (9)). The combination of the MTR and MTS assumptions leads to substantially tighter bounds than imposing each assumption alone. Unfortunately, the MTR-MTS assumption cannot provide informative results itself (columns (5) and (10)).

Table 3 also shows the bounds under the MIV assumption and its combinations. The MIV assumption tightens upper and lower bounds compared to the worst-case bounds, but it is not powerful enough to obtain informative bounds as well as the MTS assumption (columns (11) and (15)). These results do not depend on the definition of the MIV (columns (12) and (16)). Columns (13), (14), (17), and (18) show the bounds under the MTR-MTS-MIV assumption. The combination of these three assumptions leads to informative bounds. I obtain upper bounds that are lower than the ETS results in some cases. The ETS point estimates fall outside the confidence intervals when increasing parent's schooling from junior high school to university.

4.2 Single Treatment Effects on Child's University Graduation

To examine the robustness of the results in Table 3, I estimate bounds on an alternative definition of outcomes. Table 4 displays the results of the effects on a child's university graduation. The MTR and MTS bounds are rather wide and not very informative; thus these are not shown in Table 4.

At all levels of parental education, the ETS estimates show that additional years of schooling for a parent have significant impact on the probability of the child's university graduation (columns (1) and (2)). However, the MTR-MTS-MIV upper bounds suggest that the ETS estimates overstate the ATE for some levels of parental education (columns (4), (5), (9), and (10)). The upper bounds under the MTR-MTS-MIV assumption suggest much smaller effects of a parent's bachelor's degree on the child's university graduation when compared with a parent who completed only compulsory or lower secondary education. Unfortunately, the ETS estimates of the lower categories of a parent's schooling report small effects. Even when imposing the MTR-MTS-MIV assumption, the upper bounds on the effects cannot exclude the smaller effects of the point estimates for these levels of parental education.

4.3 Single Treatment Effects on Child's Education under the SMIV assumption

The results under the MTR-MTS-MIV assumption provide informative upper bounds since they are substantially smaller than the point estimates obtained under the ETS assumption. However, lower bounds do not exclude the zero effect of a parent's schooling on both years of schooling and the probability of university graduation of the child. Here, I use the two MIVs simultaneously under the SMIV assumption instead of using the schooling of the other parent and father's regular job variable separately.

Panel A of Table 5 shows the effects on a child's years of schooling. As such, the SMIV assumption gives informative bounds (columns (5)). The upper bounds are lower than the MTR-MTS-MIV ones, and indicate that both mother's and father's college degrees increases a child's schooling by about 1.41 years. The upper bounds result from a change of the parent's schooling from junior high school to bachelor's degree, showing the mother's university graduation increases a child's schooling by about 2.09 years and the father's university graduation by about 2.12 years. These upper bounds are substantially lower than the estimates under the ETS assumption.

Combining the two MIVs also plays an important role to tightening the lower bounds. For both mother's and father's schooling, the results show that the effect of an increase in parent's schooling to college degree or more impacts the child's schooling significantly different from zero. For example, a change in mother's schooling from junior high school to bachelor's degree results in an increase in child's schooling by about 0.39 years. The father's bachelor's degree increases the child's schooling by about 0.43 years.

Panel B of Table 5 shows the effects on the probability of a child's university graduation. The results using this outcome variable are similar to those using the child's years of schooling as the outcome variable. The MTR-MTS-SMIV bounds give not only more informative upper bounds but also more

informative lower ones (column (10)). The results of increasing father's or mother's schooling to a college degree or more has a positive effect on the child's university graduation, being statistically significantly different from zero, but substantially lower than the point estimates under the assumption of ETS.

To summarize, the combination of the MTR-MTS-MIV assumptions provides informative bounds on the ATE of a parent's schooling on the child's schooling. The main results show that the tightest upper bounds on the effects are lower than the point estimates under the ETS assumption. These results suggest that the regression estimates have an upward bias from the true causal effect of a parent's schooling. The combination of the MTR-MTS-SMIV assumptions also gives informative lower bounds, which suggest positive effects of parental education on the child's schooling.

4.4 Multiple Treatment Effects on Child's Years of Schooling and University Graduation

Here, I provide the results of analyses on multiple treatments of both parents' schooling. The partial identification approach discussed in the method section provides bounds on the ATE of multiple treatments under different semi-monotonicity assumptions, as well as the single treatment case. Table 6 shows the multiple treatments effects of parents' schooling on the child's schooling. The SMTR and the SMTS bounds are too wide to provide informative results, and not shown in the table.¹⁶

Panel A of Table 6 provides the effects on a child's years of schooling. I begin by examining the estimated ATE under the assumption of ETS for a useful benchmark. Children whose parents complete college or more have about 2.27 more years of schooling than those whose parents complete high school or less (column (1)). Compared to the point estimates under the ETS assumption, the conservative worst-case bounds are [-4.3261 5.8333], being too wide to acquire information on the ATE of parents' schooling (column (2)).

By combining different sets of assumptions, the results clearly illustrate more informative bounds on the ATE. The estimated bounds are [0, 2.2697] under the SMTR-SMTS assumption and [0, 2.1060] under the SMTR-SMTS-MIV assumption with an MIV of the father's regular job or [0, 2.2388] with an MIV for a child was born after 1975 (columns (3), (4), and (5)). These upper bounds are smaller than the estimates under the ETS assumption. While both of these MIV assumptions substantially reduce the upper bounds, there still remains uncertainty about the lower bounds on the ATE. Under the combined SMTR-SMTS and SMIV assumptions, the bounds narrow to [0.2397, 1.2077] and the confidence intervals are (0.1913, 1.3243), which excludes zero (column (6)). Therefore, a set of semi-monotonicity assumptions provides the non-negative impact of parents' education on the child's schooling.

Panel B of Table 6 shows the effects on the probability of a child's university graduation. The

¹⁶In this table, I only use the two levels of parents' schooling (college degree or more and high school or less), because the number of observations is too small to estimate using four levels of parents' schooling.

results of this outcome variable are similar to the results of the child's years of schooling. Under the ETS assumption, the point estimates indicate that children whose parents complete college or more are more likely graduate university than those whose parents complete high school or less (column (7)). The difference in the probability is 42%. Compared to the point estimates under the ETS assumption, the conservative worst-case bounds are [-30.17, 82.71]%, which includes the negative effects of parents' schooling (column (8)).

The estimated bounds are [0, 42.00]% under the SMTR-SMTS assumption and [0, 38.39]% under the SMTR-SMTS-MIV assumption with an MIV of father's regular job or [0, 41.43]% with an MIV for a child was born after 1975 (columns ((9), (10), and (11)). These upper bounds are smaller than the estimates under the ETS assumption. Under the SMTR-SMTS-SMIV assumption, the bounds narrow to [4.54, 19.84]% and the confidence intervals are (3.53, 22.69), which excludes zero (column (12)). This suggests that parents' college degrees are at least somewhat beneficial for the child's university graduation. However, this estimated bound is strictly smaller than the point estimates under the ETS assumption, which is sometimes implicitly assumed in the literature.

To summarize, the combination of the semi-monotonicity assumptions provides informative bounds on the ATE of parents' schooling on the child's schooling in the multiple treatments framework. The main results show that the tightest upper bounds on the effects are substantially lower than the point estimates that assume ETS. These results suggest that the regression estimates have an upward bias from the true causal effect of the parents' education. The combination of the SMTR-SMTS-SMIV assumptions also gives informative lower bounds, which suggests non-zero positive effects of parents' education on the child's schooling.

5 Conclusion

Quantifying the causal effects of parents' schooling on the child's schooling is a goal of recent empirical literature on intergenerational effects of education. The literature applies several strategies to identify the causal effects, such as identical twins, adopted children, and IV methods. Nevertheless, the empirical evidence on the causal relationship between parents' and their child's schooling is moderately convincing, because these methods identify the effects for a limited population, and the different methods have inconsistent results. The findings from a regression approach have shown large, positive, and statistically significant estimates, but implicitly assume that parents' schooling is randomly assigned to the child. This does not seem to be appropriate if unobserved heterogeneity determines both parents' and the child's educational levels.

The main contributions of this paper are twofold. First, it estimates the causal effects of parents' schooling on their child's schooling in Japan using a nonparametric bounds method, and provides the ATE bounds without relying on an invalid exogenous selection assumption. Second, it presents results of the multiple treatments of both parents' schooling as an application of the multiple treatments

framework. To obtain informative bounds, this paper imposes relatively weak and partially testable monotonicity and semi-monotonicity assumptions on treatment response, selection, and IVs.

The tightest bounds under the combination of these three assumptions show that the obtained lower bounds on the ATE are positive and significantly different from zero, but the upper bounds are lower than the point estimates that rely on the assumption of the exogenous selection of parents' schooling. This suggests that the previous studies using OLS overestimate the true causal effect of parents' education. This implies that a combination of semi-monotone assumptions is useful to acquire informative bounds on the intergenerational effects of education as semi-ordered multivalued treatments of the parents' schooling.

References

- Becker, Gary S. 1975. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: National Bureau of Economic Research: distributed by Columbia University Press.
- Becker, Gary S. and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy* 87 (6):1153–1189.
- Black, Sandra E. and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*, vol. 4, Part B, edited by David Card and Orley C. Ashenfelter, chap. 16. Amsterdam: Elsevier, 1487–1541.
- Card, David. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, vol. 3, edited by Orley C. Ashenfelter and David Card, chap. 30. Amsterdam: Elsevier, 1801–1863.
- Carneiro, Pedro, Costas Meghir, and Matthias Parey. 2013. "Maternal Education, Home Environments, and the Development of Children and Adolescents." *Journal of the European Economic Association* 11 (s1):123–160.
- Chevalier, Arnaud, Colm Harmon, Vincent O'Sullivan, and Ian Walker. 2013. "The Impact of Parental Income and Education on the Schooling of Their Children." *IZA Journal of Labor Economics* 2 (1):8.
- De Haan, Monique. 2011. "The Effect of Parents' Schooling on Child's Schooling: A Nonparametric Bounds Analysis." *Journal of Labor Economics* 29 (4):859–892.
- Del Bono, Emilia, Marco Francesconi, Yvonne Kelly, and Amanda Sacker. 2016. "Early Maternal Time Investment and Early Child Outcomes." *The Economic Journal* 126 (596):F96–F135.
- Dickson, Matt, Paul Gregg, and Harriet Robinson. 2016. "Early, Late or Never? When Does Parental Education Impact Child Outcomes?" *The Economic Journal* 126 (596):F184–F231.
- Flores, Carlos A. and Oscar A. Mitnik. 2013. "Comparing Treatments across Labor Markets: An Assessment of Nonexperimental Multiple-Treatment Strategies." *Review of Economics and Statistics* 95 (5):1691–1707.
- Frölich, Markus. 2004. "Programme Evaluation with Multiple Treatments." *Journal of Economic Surveys* 18 (2):181–224.
- Ginther, Donna K. 2000. "Alternative Estimates of the Effect of Schooling on Earnings." *The Review of Economics and Statistics* 82 (1):103–116.
- Giustinelli, Pamela. 2011. "Non-Parametric Bounds on Quantiles under Monotonicity Assumptions: With an Application to the Italian Education Returns." *Journal of Applied Econometrics* 26 (5):783–824.
- Gonzalez, Libertad. 2005. "Nonparametric Bounds on the Returns to Language Skills." *Journal of Applied Econometrics* 20 (6):771–795.
- Guryan, Jonathan, Erik Hurst, and Melissa Kearney. 2008. "Parental Education and Parental Time with Children." *Journal of Economic Perspectives* 22 (3):23–46.

- Heckman, James J., Sergio Urzua, and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *Review of Economics and Statistics* 88 (3):389–432.
- Heckman, James J. and Edward Vytlacil. 2007. "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments." In *Handbook of Econometrics*, vol. 6, Part B, chap. 71. Amsterdam: Elsevier, 4875–5143.
- Holmlund, Helena, Mikael Lindahl, and Erik Plug. 2011. "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods." *Journal of Economic Literature* 49 (3):615–651.
- Huang, Jian, Henriëtte Maassen van den Brink, and Wim Groot. 2012. "Does Education Promote Social Capital? Evidence from IV Analysis and Nonparametric-Bound Analysis." *Empirical Economics* 42 (3):1011–1034.
- Imbens, Guido W. 2000. "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika* 87 (3):706–710.
- Imbens, Guido W. and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62:467–475.
- Kariya, Takehiko. 2001. *Kaisoka Nihon to Kyoiku Kiki (Education in Crisis in Stratified Japan)*. Tokyo, Japan: Yushindo Kobusha. (in Japanese).
- Lafférs, Lukáš. 2013. "A Note on Bounding Average Treatment Effects." *Economics Letters* 120 (3):424–428.
- Lechner, Michael. 2001. "Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption." In *Econometric Evaluation of Labour Market Policies*, edited by Michael Lechner and Friedhelm Pfeiffer. Heidelberg: Physica-Verlag/Springer, 43–58.
- Manski, Charles F. 1989. "Anatomy of the Selection Problem." *Journal of Human Resources* 24 (3):343–360.
- . 1990. "Nonparametric Bounds on Treatment Effects." *American Economic Review* 80 (2):319–23.
- . 1997. "Monotone Treatment Response." *Econometrica* 65 (6):1311–1334.
- Manski, Charles F. and John V. Pepper. 1998. "Monotone Instrumental Variables: With an Application to the Returns to Schooling." Working Paper 224, National Bureau of Economic Research, Cambridge, MA.
- . 2000. "Monotone Instrumental Variables: With an Application to the Returns to Schooling." *Econometrica* 68 (4):997–1010.
- Mariotti, Martine and Juergen Meinecke. 2015. "Partial Identification and Bound Estimation of the Average Treatment Effect of Education on Earnings for South Africa." *Oxford Bulletin of Economics and Statistics* 77 (2):210–233.

- Okumura, Tsunao and Emiko Usui. 2014. "Concave-Monotone Treatment Response and Monotone Treatment Selection: With an Application to the Returns to Schooling." *Quantitative Economics* 5 (1):175–194.
- Oreopoulos, Philip, Marianne E. Page, and Ann H. Stevens. 2006. "The Intergenerational Effects of Compulsory Schooling." *Journal of Labor Economics* 24 (4):729–760.
- Oreopoulos, Philip and Kjell G. Salvanes. 2011. "Priceless: The Nonpecuniary Benefits of Schooling." *Journal of Economic Perspectives* 25 (1):159–184.
- Oshio, Takashi and Wataru Seno. 2007. "The Economics of Education in Japan: A Survey of Empirical Studies and Unresolved Issues." *Japanese Economy* 34 (1):46–81.
- Solon, Gary. 1999. "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics*, vol. 3, Part A, edited by Orley C. Ashenfelter and David Card, chap. 29. Amsterdam: Elsevier, 1761–1800.
- Tachibanaki, Toshiaki. 1988. "Education, Occupation, Hierarchy and Earnings." *Economics of Education Review* 7 (2):221–229.
- Tanaka, Ryuichi. 2008. "The Gender-Asymmetric Effect of Working Mothers on Children's Education: Evidence from Japan." *Journal of the Japanese and International Economies* 22 (4):586–604.
- Yamada, Ken. 2011. "Family Background and Economic Outcomes in Japan." Working Paper 1246, Singapore Management University.

Table 1: Summary Statistics

| | mean | sd | min | max |
|--------------------------------|---------|--------|-----|-----|
| Child's years of schooling | 13.0893 | 2.1713 | 9 | 18 |
| Child's university graduation | 0.2452 | 0.4302 | 0 | 1 |
| Mother's years of schooling | 9.7783 | 2.5953 | 6 | 18 |
| Mother's some college | 0.0750 | 0.2634 | 0 | 1 |
| Mother's university graduation | 0.0258 | 0.1584 | 0 | 1 |
| Father's years of schooling | 10.2526 | 3.1694 | 6 | 18 |
| Father's some college | 0.0681 | 0.2519 | 0 | 1 |
| Father's university graduation | 0.1231 | 0.3286 | 0 | 1 |
| Father had a regular job | 0.5872 | 0.4924 | 0 | 1 |
| Child was born after 1975 | 0.0941 | 0.2920 | 0 | 1 |
| Female | 0.5406 | 0.4984 | 0 | 1 |

Notes: This table reports summary statistics of the analysis data. Number of observations is 13,669. Father had a regular job indicates that the father was an executive of a company or was a regular employee (*joji-koyo no rodosha*) when the child was 15 years old.

Table 2: Mean of Child's Schooling Outcomes by Parent's Education Level

| Education level of parent | Mother | | Father | |
|---------------------------|----------------------------|--------------------|----------------------------|--------------------|
| | Child's years of schooling | Child's university | Child's years of schooling | Child's university |
| Junior high school | 12.1753 | 0.1184 | 12.1779 | 0.1179 |
| High school | 13.6801 | 0.3095 | 13.4729 | 0.2691 |
| Some college | 14.6966 | 0.5220 | 14.2095 | 0.4135 |
| University | 15.1222 | 0.6648 | 14.8247 | 0.5639 |

Notes: This table reports mean schooling outcomes of children by educational level of parents for the test of the MTR-MTS assumption.

Table 3: Bounds of ATE of Parent's Schooling on Child's Years of Schooling

| Assumption | ETS | | Worst-Case | | MTR | | MTS | | MTR MTS | |
|-------------------------------|------------------------------|----|-----------------------------|----------------------|-----------------|----|-----------------------------|----------------------|---------|-----------------|
| | LB | UB | LB | UB | LB | UB | LB | UB | LB | UB |
| | Effect of Mother's Schooling | | | | | | | | | |
| | (1) | | (2) | | (3) | | (4) | | (5) | |
| $E[y(t=HS)] - E[y(t=JH)]$ | 1.5049 (1.4365 1.5743) | | -4.2558 (-4.2673 5.6327) | 5.6509 (0 5.3291) | 0 (0 5.3626) | | -3.7843 (-3.8194 2.0043) | 1.9400 (0 1.6873) | | 0 (0 1.6182) |
| $E[y(t=Coll)] - E[y(t=HS)]$ | 1.0165 (0.8969 1.1355) | | -6.7095 (-6.7224 6.7343) | 6.7336 (0 5.1737) | 0 (0 5.2075) | | -4.5414 (-4.5855 3.4028) | 3.2915 (0 1.8479) | | 0 (0 2.0438) |
| $E[y(t=Univ)] - E[y(t=Coll)]$ | 0.4256 (0.1902 0.6548) | | -8.5946 (-8.6021 8.4932) | 8.4987 (0 5.0297) | 0 (0 5.0297) | | -5.6240 (-5.7323 5.7433) | 5.5483 (0 2.2469) | | 0 (0 2.2469) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | 2.9469 (2.7452 3.1443) | | -6.1168 (-6.1874 7.4581) | 7.4401 (0 7.4631) | 0 (0 7.4631) | | -6.1168 (-6.1402 3.1480) | 2.9469 (0 3.1562) | | 0 (0 3.1562) |
| | Effect of Father's Schooling | | | | | | | | | |
| | (6) | | (7) | | (8) | | (9) | | (10) | |
| $E[y(t=HS)] - E[y(t=JH)]$ | 1.2950 (1.2222 1.3728) | | -4.7572 (-4.7708 5.9647) | 5.9639 (0 5.3485) | 0 (0 5.3485) | | -3.9018 (-3.9380 2.2265) | 2.1607 (0 1.5806) | | 0 (0 1.5116) |
| $E[y(t=Coll)] - E[y(t=HS)]$ | 0.7366 (0.6015 0.8681) | | -7.0819 (-7.0943 7.1987) | 7.1973 (0 5.3335) | 0 (0 5.3671) | | -4.3424 (-4.3933 3.3848) | 3.2763 (0 1.5240) | | 0 (0 1.4125) |
| $E[y(t=Univ)] - E[y(t=Coll)]$ | 0.6153 (0.4633 0.7619) | | -8.0247 (-8.0270 8.2639) | 8.2542 (0 5.2369) | 0 (0 5.2702) | | -4.9590 (-5.0624 4.9196) | 4.8285 (0 1.9110) | | 0 (0 1.8111) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | 2.6468 (2.5472 2.7423) | | -5.5845 (-5.5978 7.1416) | 7.1362 (0 7.1362) | 0 (0 7.1612) | | -5.5845 (-5.6104 2.7476) | 2.6468 (0 2.7480) | | 0 (0 2.6468) |

Notes: This table reports average treatment effects of parent's education on child's years of schooling. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are (1) JH = junior high school, (2) HS = high school, (3) Coll = some college, (4) Univ = bachelor's degree or more. Number of observations is 13,669.

Table 3: Bounds of ATE of Parent's Schooling on Child's Years of Schooling (cont.)

| Assumption | MIV | | MIV | | MTR MTS MIV | | MTR MTS MIV | |
|-------------------------------|------------------------------------|----------------------|------------------------------------|----------------------|------------------------------------|----------------------|------------------------------------|----------------------|
| | Other parent's univ / college grad | Father's regular job | Other parent's univ / college grad | Father's regular job | Other parent's univ / college grad | Father's regular job | Other parent's univ / college grad | Father's regular job |
| | LB | UB | LB | UB | LB | UB | LB | UB |
| | (11) | | (12) | | (13) | | (14) | |
| $E[y(t=HS)] - E[y(t=JH)]$ | -4.2558 | 5.4734 | -4.2558 | 5.2105 | 0 | 1.4162 | 0 | 1.4958 |
| | (-4.2877 | 5.5076) | (-4.2862 | 5.2594) | (0 | 1.5011) | (0 | 1.5648) |
| $E[y(t=Coll)] - E[y(t=HS)]$ | -6.7095 | 6.1976 | -6.5188 | 6.6702 | 0 | 1.4125 | 0 | 1.6587 |
| | (-6.7327 | 6.2405) | (-6.5492 | 6.6945) | (0 | 1.5514) | (0 | 1.7822) |
| $E[y(t=Univ)] - E[y(t=Coll)]$ | -8.0586 | 8.0606 | -8.5312 | 8.4728 | 0 | 1.4178 | 0 | 1.8613 |
| | (-8.0974 | 8.0996) | (-8.5442 | 8.4839) | (0 | 1.9128) | (0 | 2.1019) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | -6.1168 | 6.8245 | -6.1168 | 7.1645 | 0 | 2.1948 | 0 | 2.6893 |
| | (-6.1405 | 6.8711) | (-6.1405 | 7.2055) | (0 | 2.7014) | (0 | 2.9373) |
| | (15) | | (16) | | (17) | | (18) | |
| $E[y(t=HS)] - E[y(t=JH)]$ | -4.6679 | 5.8243 | -4.7572 | 5.4669 | 0 | 1.3649 | 0 | 1.4038 |
| | (-4.6989 | 5.8570) | (-4.7869 | 5.5166) | (0 | 1.4448) | (0 | 1.4715) |
| $E[y(t=Coll)] - E[y(t=HS)]$ | -7.0819 | 6.4940 | -6.9310 | 7.1391 | 0 | 1.2180 | 0 | 1.2909 |
| | (-7.1026 | 6.5455) | (-6.9591 | 7.1615) | (0 | 1.3513) | (0 | 1.4148) |
| $E[y(t=Univ)] - E[y(t=Coll)]$ | -7.4106 | 7.0613 | -7.9665 | 8.1252 | 0 | 1.5775 | 0 | 1.6869 |
| | (-7.4599 | 7.1286) | (-7.9825 | 8.1426) | (0 | 1.6966) | (0 | 1.8045) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | -5.5845 | 5.8037 | -5.5845 | 6.6612 | 0 | 2.2836 | 0 | 2.4542 |
| | (-5.6098 | 5.8780) | (-5.6097 | 6.7041) | (0 | 2.4119) | (0 | 2.5719) |

Notes: This table reports average treatment effects of parent's education on child's years of schooling. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are (1) JH = junior high school, (2) HS = high school, (3) Coll = some college, (4) Univ = bachelor's degree or more. MIVs are the indicator of other parent's university / college graduation for mother's / father's education and father's regular job status. Number of observations is 13,669.

Table 4: Bounds of ATE of Parent's Schooling on Child's University Graduation

| Assumption MIV | ETS | | Worst-Case | | MTR MTS | | MTR MTS MIV | | MTR MTS MIV | |
|-------------------------------|------------------------------|----|-----------------------------|--------|---------|----------------------|-------------|----------------------|-------------|----------------------|
| | LB | UB | LB | UB | No | UB | LB | UB | LB | UB |
| | Effect of Mother's Schooling | | | | | | | | | |
| | (1) | | (2) | | (3) | | (4) | | (5) | |
| $E[y(t=HS)] - E[y(t=JH)]$ | 0.1912 (0.1772 0.2054) | | -0.4539 (-0.4669 0.6471) | 0.6468 | 0 | 0.2163 (0 0.2302) | 0 | 0.1815 (0 0.1987) | 0 | 0.2036 (0 0.2178) |
| $E[y(t=Coll)] - E[y(t=HS)]$ | 0.2124 (0.1801 0.2446) | | -0.6630 (-0.6652 0.8323) | 0.8306 | 0 | 0.3056 (0 0.3373) | 0 | 0.2326 (0 0.2687) | 0 | 0.2881 (0 0.3200) |
| $E[y(t=Univ)] - E[y(t=Coll)]$ | 0.1428 (0.0829 0.2024) | | -0.9470 (-0.9487 0.9536) | 0.9522 | 0 | 0.4233 (0 0.4736) | 0 | 0.2795 (0 0.4072) | 0 | 0.3896 (0 0.4491) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | 0.5464 (0.4977 0.5967) | | -0.5703 (-0.5764 0.9365) | 0.9360 | 0 | 0.5464 (0 0.5977) | 0 | 0.3838 (0 0.5128) | 0 | 0.5079 (0 0.5678) |
| | Effect of Father's Schooling | | | | | | | | | |
| | (6) | | (7) | | (8) | | (9) | | (10) | |
| $E[y(t=HS)] - E[y(t=JH)]$ | 0.1512 (0.1371 0.1663) | | -0.4983 (-0.5038 0.6987) | 0.6930 | 0 | 0.1973 (0 0.2103) | 0 | 0.1785 (0 0.1948) | 0 | 0.1889 (0 0.2028) |
| $E[y(t=Coll)] - E[y(t=HS)]$ | 0.1445 (0.1098 0.1779) | | -0.7194 (-0.7242 0.8721) | 0.8671 | 0 | 0.2330 (0 0.2611) | 0 | 0.1917 (0 0.2268) | 0 | 0.2036 (0 0.2346) |
| $E[y(t=Univ)] - E[y(t=Coll)]$ | 0.1503 (0.1100 0.1904) | | -0.8906 (-0.8943 0.9218) | 0.9181 | 0 | 0.3372 (0 0.3633) | 0 | 0.2794 (0 0.3096) | 0 | 0.3153 (0 0.3436) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | 0.4460 (0.4209 0.4702) | | -0.5218 (-0.5263 0.8965) | 0.8917 | 0 | 0.4460 (0 0.4720) | 0 | 0.3728 (0 0.4056) | 0 | 0.4218 (0 0.4498) |

Notes: This table reports average treatment effects of parent's education on child's university graduation. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are (1) JH = junior high school, (2) HS = high school, (3) Coll = some college, (4) Univ = bachelor's degree or more. MIVs are the indicator of other parent's university / college graduation for mother's / father's education and father's regular job status. Number of observations is 13,669.

Table 5: Bounds of ATE of Parent's Schooling on Child's Years of Schooling and University Graduation

| Assumption | ETS | | Worst-Case | | MTR MTS MIV | | MTR MTS MIV | | MTR MTS SMIV | |
|---|---------|---------|------------|---------|-------------|---------|-------------|---------|--------------|---------|
| | LB | UB | LB | UB | LB | UB | LB | UB | LB | UB |
| Panel A: Effect on Child's Years of Schooling | | | | | | | | | | |
| Mother's Schooling | (1) | | (2) | | (3) | | (4) | | (5) | |
| $E[y(t=Coll+)] - E[y(t=HS-)]$ | 1.9083 | | -3.8263 | 5.1737 | 0 | 1.4588 | 0 | 1.7959 | 0.0475 | 1.4133 |
| | (1.8040 | 2.0119) | (-3.8609 | 5.2100) | (0 | 1.5945) | (0 | 1.9147) | (0.0028 | 1.5528) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | 2.9469 | | -6.1168 | 7.4401 | 0 | 2.1948 | 0 | 2.6893 | 0.3890 | 2.0886 |
| | (2.7452 | 3.1443) | (-6.1874 | 7.4581) | (0 | 2.7014) | (0 | 2.9373) | (0.3175 | 2.4566) |
| Father's Schooling | | | | | | | | | | |
| $E[y(t=Coll+)] - E[y(t=HS-)]$ | 1.8748 | | -3.6665 | 5.3335 | 0 | 1.5601 | 0 | 1.7141 | 0.1308 | 1.4145 |
| | (1.7930 | 1.9562) | (-3.7019 | 5.3685) | (0 | 1.6600) | (0 | 1.8095) | (0.0867 | 1.5182) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | 2.6468 | | -5.5845 | 7.1362 | 0 | 2.2836 | 0 | 2.4542 | 0.4321 | 2.1167 |
| | (2.5472 | 2.7423) | (-5.5978 | 7.1416) | (0 | 2.4119) | (0 | 2.5719) | (0.4221 | 2.2564) |
| Panel B: Effect on Probability of Child University Graduation | | | | | | | | | | |
| Mother's Schooling | (6) | | (7) | | (8) | | (9) | | (10) | |
| $E[y(t=Coll+)] - E[y(t=HS-)]$ | 0.3484 | | -0.2334 | 0.7666 | 0 | 0.2429 | 0 | 0.3245 | 0.0092 | 0.2317 |
| | (0.3214 | 0.3755) | (-0.2402 | 0.7738) | (0 | 0.2797) | (0 | 0.3533) | (0.0007 | 0.2689) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | 0.5464 | | -0.5703 | 0.9360 | 0 | 0.3838 | 0 | 0.5079 | 0.0567 | 0.3773 |
| | (0.4977 | 0.5967) | (-0.5764 | 0.9365) | (0 | 0.5128) | (0 | 0.5678) | (0.0413 | 0.4734) |
| Father's Schooling | | | | | | | | | | |
| $E[y(t=Coll+)] - E[y(t=HS-)]$ | 0.3279 | | -0.2412 | 0.7588 | 0 | 0.2618 | 0 | 0.2981 | 0.0246 | 0.2338 |
| | (0.3072 | 0.3486) | (-0.2478 | 0.7656) | (0 | 0.2863) | (0 | 0.3200) | (0.0153 | 0.2596) |
| $E[y(t=Univ)] - E[y(t=JH)]$ | 0.4460 | | -0.5218 | 0.8917 | 0 | 0.3728 | 0 | 0.4218 | 0.0665 | 0.3499 |
| | (0.4209 | 0.4702) | (-0.5263 | 0.8965) | (0 | 0.4056) | (0 | 0.4498) | (0.0557 | 0.3849) |

Notes: This table reports average treatment effects of parent's education on child's years of schooling and child's university graduation. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are JH = junior high school, Univ = bachelor's degree or more, HS- = high school or less, Coll+ = college or more. MIVs are the indicator of father's regular job status and the indicator that denotes the child was born after 1975. Number of observations is 13,669.

Table 6: Bounds of ATE of Parents' Schooling on Child's Years of Schooling and University Graduation

| Assumption | ETS | | Worst-Case | | SMTR SMTS | | SMTR SMTS MIV | | SMTR SMTS MIV | | SMTR SMTS SMIV | |
|---|---------|---------|------------|---------|-----------|---------|----------------------|---------|-----------------|---------|--|---------|
| | No | LB | No | LB | No | LB | Father's regular job | LB | Born After 1975 | LB | Father's regular job and Born After 1975 | UB |
| MIV | LB | UB | LB | UB | LB | UB | LB | UB | LB | UB | LB | UB |
| Panel A: Effect on Child's Years of Schooling | | | | | | | | | | | | |
| | (1) | | (2) | | (3) | | (4) | | (5) | | (6) | |
| $E[y](t^M, t^F)=(\text{Coll}+, \text{Coll}+)]$ | 2.2697 | | -4.3261 | 5.8333 | 0 | 2.2697 | 0 | 2.1060 | 0 | 2.2388 | 0.2397 | 1.2077 |
| $-E[y](t^M, t^F)=(\text{HS}-, \text{HS}-)]$ | (2.1610 | 2.3856) | (-4.3591 | 5.8669) | (0 | 2.3864) | (0 | 2.2344) | (0 | 2.3499) | (0.1913 | 1.3243) |
| Panel B: Effect on Probability of Child University Graduation | | | | | | | | | | | | |
| | (7) | | (8) | | (9) | | (10) | | (11) | | (12) | |
| $E[y](t^M, t^F)=(\text{Coll}+, \text{Coll}+)]$ | 0.4200 | | -0.3017 | 0.8271 | 0 | 0.4200 | 0 | 0.3839 | 0 | 0.4143 | 0.0454 | 0.1984 |
| $-E[y](t^M, t^F)=(\text{HS}-, \text{HS}-)]$ | (0.3893 | 0.4491) | (-0.3080 | 0.8332) | (0 | 0.4510) | (0 | 0.4177) | (0 | 0.4443) | (0.0353 | 0.2269) |

Notes: This table reports average treatment effects of a set of vector of mother's and father's education on the child's years of schooling and the child's university graduation. t^M and t^F denotes the treatment level of mother's and father's education respectively. Numbers between parentheses are 95% confidence intervals calculated by the percentile bootstrap method. Levels of parental education are HS- = high school or less, Coll+ = college or more. MIVs are the indicator of father's regular job status and the indicator that denotes the child was born after 1975. Number of observations is 13,669.