

The effect of parents' schooling on child's schooling: A nonparametric bounds analysis

Monique de Haan, *University of Amsterdam*

Abstract

A positive relation between parents' schooling and child's schooling does not necessarily reflect a causal relation. This paper uses a new approach to identify intergenerational schooling effects: a nonparametric bounds analysis. By relying on relatively weak and in part testable assumptions this paper obtains informative bounds on the average causal impact of parents' schooling. The tightest bounds, using monotone instrumental variables, show that increasing mother's or father's schooling to a college degree has a positive effect on child's schooling which is significantly different from zero, but substantially lower than the OLS estimates.

1 Introduction

Is there an effect of parents' schooling on the schooling of their child? This question has received much attention in the empirical literature. Haveman and Wolfe (1995) give an overview of the literature and

Contact information: Monique de Haan, TIER, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands, Email: moniquedehaan@uva.nl.

I thank Erik Plug, Hessel Oosterbeek, Edwin Leuven and participants at the Tinbergen Institute seminar (Amsterdam, 2008), the COST conference (Paris, 2008), the RTN Network EEEPE conference (Amsterdam, 2008), the EALE conference (Tallin, 2009) and Statistics Norway seminar (Oslo, 2010) for their helpful comments. Support for collection and dissemination of data from the Wisconsin Longitudinal Study has been provided by the National Institute on Aging (AG-9775), the National Science Foundation (SBR-9320660), the Spencer Foundation, and the Center for Demography and Ecology and the Vilas Estate Trust at the University of Wisconsin–Madison. Data and documentation from the Wisconsin Longitudinal Study are available at <http://www.ssc.wisc.edu/wlsresearch/documentation>.

show that most, if not all, studies find a strong positive association between parental and child's schooling. It is not clear though whether a positive association is informative about whether increasing the education of one generation will affect the schooling of the next generation. If the schooling choices of parents are related to unobservables that affect their children's education, such as genetic endowments or child rearing talents, a positive association does not necessarily reflect a causal relation.

To estimate the causal impact of parents' schooling, a couple of innovative identification strategies have been applied in the recent empirical literature. One of the identification methods relies on identical twin parents and their children to eliminate the correlation between parental schooling and child's schooling attributable to genetics (Behrman and Rosenzweig (2002, 2005), Antonovics and Goldberger (2005)). A second approach uses a sample of adoptees, exploiting the fact that there is no genetic link between adoptive parents and their adopted child (Björklund, Lindahl and Plug (2006), Sacerdote (2002, 2007), Plug (2004)). A third identification strategy is an instrumental variable approach. The instruments that are used vary from changes in compulsory schooling laws to college openings/closings, variation in tuition fees, proximity to college and variation in the quality of entry exams. See for example, Black, Devereux and Salvanes (2005), Chevalier (2004), Oreopoulos, Page and Stevens (2006), Currie and Moretti (2003), Carneiro, Meghir, and Patey (2007) and Maurin and McNally (2008).

These recent studies that focus on estimating the causal impact of parents' education give contradictory findings. Some studies find a positive and significant impact of mother's schooling but no significant impact of father's schooling. Other studies find a significant positive effect of father's schooling but not of mother's schooling, and there are studies that find small but significant effects of both mother's and father's schooling on the schooling of their child. Holmlund, Lindahl and Plug (2010) give an overview of this recent empirical literature, and apply the three different identification methods to Swedish register data. They find that the estimates of the effect of parents' education differ systematically across the identification strategies. Since they use a single data set, they conclude that differences between studies are not due to idiosyncratic differences between the data sets that are used.

There are two possible explanations for the diverging findings in the literature. The first is that the studies use different identification methods, that are based on different sets of assumptions, to estimate the causal impact of parents' schooling. The twin-method assumes that all differences in schooling levels between monozygotic twin parents are exogenous. The adoption-approach has to assume that adoptees are randomly assigned to their adoptive parents and that parents' child-rearing talents are uncorrelated

to their level of schooling. Finally, whether an instrumental variables approach can be used to identify the causal impact of parents' schooling depends on the strength and validity of the instruments. If the assumptions on which an identification approach is based do not hold this will result in an invalid and inconsistent estimate of the effect of parent's education. The diverging findings across studies can therefore be the result of violations of (some of) these assumptions.

The second potential explanation is that the methods estimate different local average treatment effects, and not the average treatment effect. The different studies rely on sub-populations – twins, adoptees, or those affected by the instrument – that differ in their characteristics from the overall population. The estimates based on these sub-populations will only be estimates of the average treatment effect if the effect of parents' schooling is linear and homogeneous. Failure of this assumption could therefore also explain the different findings in the literature.

The contribution of the present paper is that it uses a different approach to estimate the causal impact of parents' schooling on child's schooling. The method used in this paper is a nonparametric bounds analysis based on Manski and Pepper (2000). This method has two important advantages over the other methods that have been used in the literature. First, it relies on relatively weak and in part testable assumptions. Second, it obtains bounds on the average treatment effect without having to rely on the assumption of a linear and homogeneous effect of parents' schooling.

The number of studies applying a nonparametric bounding method has been growing in recent years (Gundersen and Kreider (2009), Hill and Kreider (2009), Kreider and Pepper (2008), Blundell et al. (2007), Gerfin and Schellhorn (2006), González (2005), Pepper (2000)). There has been no study though, that applies this method to identify the causal impact of parents' schooling on child's schooling. In order to know more about external benefits of education and mechanisms behind (in)equality of opportunity it is important to have consistent information about the size of intergenerational schooling effects. Considering the controversies in the existing literature, a nonparametric bounding approach is therefore especially useful, because it can serve as an effective and credible middle ground.

The results in this paper show that even though the nonparametric bounds method is conservative and produces a range instead of a point estimate, it gives informative upper and lower bounds on the causal impact of parents' schooling. The tightest bounds show that the effect of increasing mother's or father's schooling to a college degree has an effect on child's schooling which is significantly different from zero, but substantially lower than the OLS estimates. In addition, 2SLS estimates using college

proximity as instrument for parents' schooling as well as a number of point estimates reported in the recent literature fall outside the estimated nonparametric bounds.

The remainder of the paper is organized as follows. Section 2 gives the empirical specification. Section 3 gives a description of the data. Section 4 gives the results of the nonparametric bounds analysis. The analysis starts by investigating what the data can tell us about the effect of parents' schooling without adding any assumptions and subsequently some weak nonparametric assumptions will be added to tighten the bounds. Section 5 elaborates on how the nonparametric bounds compare to findings using other identification approaches and Section 6 summarizes and conclude.

2 Empirical specification

For each child we have a response function $y_i(\cdot) : T \rightarrow Y$ which maps treatments $t \in T$ into outcomes $y_i(t) \in Y$. Where the treatment t is the level of schooling of the parent and y is years of schooling of the child. For each child we observe the realized level of parental schooling z_i and his realized years of schooling $y_i \equiv y_i(z_i)$, but we do not observe the potential outcomes $y_i(t)$ for $t \neq z_i$. To simplify notation the subscript i will be dropped in the following.

We are interested in the average treatment effect of increasing parental schooling from s to t on child's schooling, that is

$$\Delta(s, t) = E[y(t)] - E[y(s)] \quad (1)$$

This average treatment effect consists of two parts; the mean schooling we would observe if all children had a parent with schooling level t ($E[y(t)]$), and the mean schooling we would observe if all children had a parent with schooling level s ($E[y(s)]$). We will first focus on the mean schooling we would observe if all children had a parent with the same schooling level. The next step will be to look at the average treatment effect of increasing parental schooling from s to t .

By using the law of iterated expectations and the fact that $E[y(t)|z = t] = E[y|z = t]$ we can write

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

With a data set where we observe the schooling of children and their parents we can observe the mean schooling of children who have a parent with schooling level t and the proportion of children that have a parent with schooling level t . However, for the children who have a parent with a schooling level different

than t we cannot observe what their mean schooling would have been if their parents did have schooling level t . That is, we cannot observe $E[y(t)|z \neq t]$. It is only possible to say more about the effect of interest by augmenting the things that are observed with assumptions.

Manski (1989) shows though that it is possible to identify bounds on $E[y(t)]$ without adding any assumptions if the support of the dependent variable is bounded, which is the case with child's schooling. By substituting $E[y(t)|z \neq t]$ with the lowest possible level of education y_{min} we obtain a lower bound on $E[y(t)]$ and by replacing it with the highest possible level of schooling y^{max} we obtain the upper bound. This gives Manski's (1989) no-assumption bounds.

$$\begin{aligned}
 & \text{No-assumption bounds} \\
 & E[y|z = t] \cdot P(z = t) + y_{min} \cdot P(z \neq t) \\
 & \leq E[y(t)] \leq \\
 & E[y|z = t] \cdot P(z = t) + y_{max} \cdot P(z \neq t)
 \end{aligned} \tag{3}$$

These no-assumption bounds are interesting because all results which are based on different assumptions about $E[y(t)|z \neq t]$ will lie within these bounds. On the other hand these bounds can be very wide and in that sense they are not so informative. In the analysis we will therefore add some nonparametric assumptions to tighten the no-assumption bounds. We will subsequently add the monotone treatment response assumption (MTR) and the monotone treatment selection assumption (MTS) which are introduced and derived in Manski (1997) and Manski and Pepper (2000).

The monotone treatment response assumption states that the outcome is a weakly increasing function of the treatment:

$$t_2 \geq t_1 \Rightarrow y(t_2) \geq y(t_1) \tag{4}$$

Equation (4) assumes that increasing parents' schooling weakly increases child's schooling, which is also suggested by human capital theory (Becker and Tomes (1979), Solon (1999)). There are many reasons to expect a positive impact of increasing parents' schooling on children's schooling; more income, better help with homework, role model effects, etc. It is very hard though, to come up with reasons why increasing parents' schooling would have a negative impact on child's schooling.¹ It could however be the case that

¹One possible reason is brought forward by Behrman and Rosenzweig (2002). They find, using a within-twin approach, that increasing mother's schooling has a (marginally significant) negative impact on child's schooling and argue that more education can induce a mother to spend more time on the labor market and less with her children. A recent paper by Guryan, Hurts and Kearny (2008) shows, however, that higher educated mothers spend on average more time with their children than lower educated mothers (instead of less) and that this holds for working and non-working mothers. This goes

parents' schooling has no causal effect on child's schooling and that the positive associations which have been found by many studies are due to selection and not causation. It is therefore important that a zero effect is not ruled out by the MTR assumption.

Figure 1b shows how the MTR assumption can be used to tighten the no-assumption bounds (which are shown in Figure 1a). A sample of children and their parents can be divided into three groups; (1) children with a parent that has a schooling level lower than t ($z < t$), (2) children that have a parent with a schooling level equal to t ($z = t$), and (3) children who have a parent with a schooling level higher than t ($z > t$). For the second group we observe the effect on mean schooling of having a parent with schooling level t . For the first group we know that under the MTR assumption their observed mean schooling is less than or equal to what their mean schooling would have been if their parent did have schooling level t . So we can use the mean schooling we observe for this first group to tighten the lower bound. For the third group the MTR assumption implies that if they would have had a parent with schooling level t , their mean schooling would have been lower than or equal to their current mean schooling. We can therefore use the mean schooling we observe for this third group to tighten the upper bound. By combining the MTR assumption with the no-assumption bounds above we get the MTR bounds:

$$\begin{aligned}
 & \text{MTR bounds} \\
 & E[y|z < t] \cdot P(z < t) + E[y|z = t] \cdot P(z = t) + y_{min} \cdot P(z > t) \\
 & \leq E[y(t)] \leq \\
 & y_{max} \cdot P(z < t) + E[y|z = t] \cdot P(z = t) + E[y|z > t] \cdot P(z > t)
 \end{aligned} \tag{5}$$

A second assumption that will be used in the analysis in this paper is the monotone treatment selection assumption (MTS). Under this assumption children with higher schooled parents have weakly higher mean schooling functions than those with lower schooled parents:

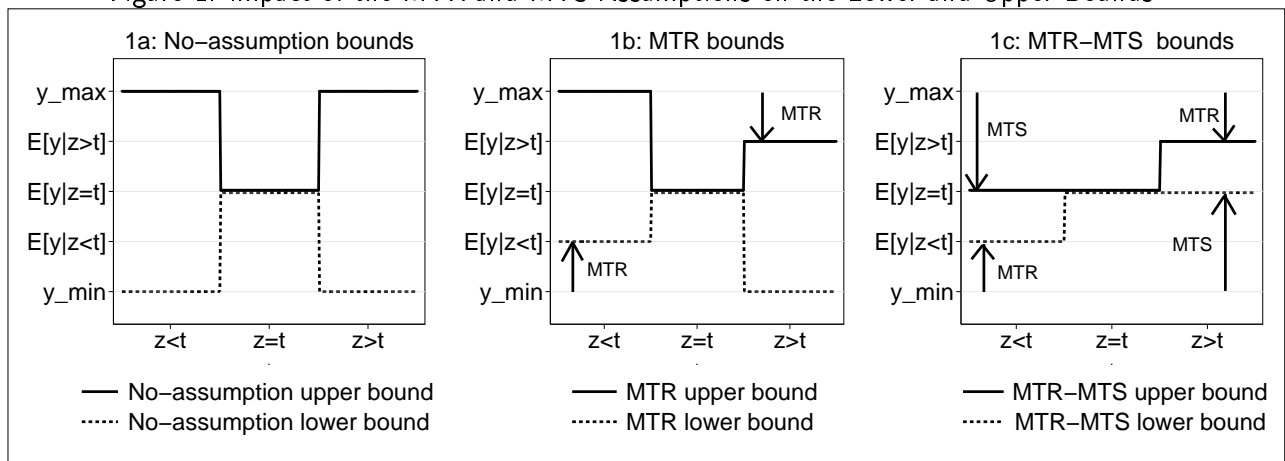
$$u_2 \geq u_1 \Rightarrow E[y(t)|z = u_2] \geq E[y(t)|z = u_1] \tag{6}$$

Many studies point out that we cannot interpret the observation that higher schooled parents have higher schooled children as a causal relation because higher educated parents are different from lower educated parents. For example, higher educated parents are likely to be of higher ability and due to genetic

against the reasoning of Berhman and Rosenzweig. In addition, a number of recent papers that also use a within-twin approach find no or small effects of mother's schooling, but never a negative effect (Antonovics and Goldberger (2005), Holmlund et al. (2010) and Bingley, Christensen and Jensen (2009)).

transmission of endowments their children will on average also have a higher ability, and therefore a higher education. In addition, higher educated parents might create a different and in general a more education-stimulating environment for their children compared to lower educated parents, regardless of their level of education. These arguments are all consistent with the monotone treatment selection assumption. The MTS assumption excludes the possibility that the characteristics that make higher educated parents different from lower educated parents have an impact on their children's schooling in such a way, that at a given level of parents' schooling children of higher educated parents would perform worse in school than children of lower educated parents.

Figure 1: Impact of the MTR and MTS Assumptions on the Lower and Upper Bounds



To use the MTS assumption we can again divide the sample into three groups, (1) children who have a parent with a schooling level lower than t ($z < t$), (2) equal to t ($z = t$), and (3) higher than t ($z > t$). If the schooling of the parents of the first group would be increased to t , we know by the MTS assumption that the mean schooling of the children would be weakly lower than the mean schooling we observe for the children who currently have a parent with schooling level t . We can therefore use the mean schooling we observe for the children who have a parent with schooling level t as an upper bound for the first group. Similarly we can use it as a lower bound for the third group. Figure 1c shows how the MTS assumption can be used to tighten the bounds.

By combining the monotone treatment response assumption and the monotone treatment selection

assumption we get the MTR-MTS bounds:²

$$\begin{aligned}
 & \text{MTR-MTS bounds} \\
 & E[y|z < t] \cdot P(z < t) + E[y|z = t] \cdot P(z = t) + E[y|z > t] \cdot P(z > t) \\
 & \leq E[y(t)] \leq \\
 & E[y|z = t] \cdot P(z < t) + E[y|z = t] \cdot P(z = t) + E[y|z > t] \cdot P(z > t)
 \end{aligned} \tag{7}$$

It is possible to test the combined MTR-MTS assumption. Under the MTR-MTS assumption the following should hold³

$$\begin{aligned}
 & \text{for } u_2 > u_1 \\
 & E[y|z = u_2] = E[y(u_2)|z = u_2] \geq E[y(u_2)|z = u_1] \geq E[y(u_1)|z = u_1] = E[y|z = u_1]
 \end{aligned}$$

So under the MTR-MTS assumption the mean schooling of a child should be weakly increasing in the realized level of schooling of the parent, if this is not the case the MTR-MTS assumption should be rejected.

Monotone instrumental variable assumption

Suppose we observe not only the schooling of the child and his parent but also a variable v . We could then divide the sample into sub-samples, one for each value of v , and obtain lower and upper bounds on the effect of parents' schooling for each sub-sample. It may well be that the bounds are relatively tight for some sub-samples but relatively wide for other sub-samples. We can exploit this variation in the bounds over the sub-samples if v satisfies the instrumental variable assumption (Manski and Pepper (2000)). A variable v satisfies the instrumental variable assumption, in the sense of mean-independence, if it holds that for all treatments $t \in T$ and all values of the instrument $m \in M$

$$E[y(t)|v = m] = E[y(t)] \tag{8}$$

This means that the schooling function of the child should be mean-independent of the variable v . If v satisfies the instrumental variable assumption, we can obtain an IV-lower bound on $E[y(t)]$ by taking the maximum lower bound over all sub-samples and an IV-upper bound by taking the minimum upper bound

²For a full derivation of the MTR and MTR-MTS bounds see Manski (1997) and Manski and Pepper (2000).

³The first inequalities follow from the MTS assumption and the second inequalities from the MTR assumption.

over all sub-samples. This gives the following IV-bounds.

$$\begin{aligned} & \text{IV-bounds} \\ & \max_{m \in M} (LB_{E[y(t)|v=m]}) \leq E[y(t)] \leq \min_{m \in M} (UB_{E[y(t)|v=m]}) \end{aligned} \tag{9}$$

Since it is difficult to find a variable which satisfies the instrumental variable assumption in equation (8) we will use a weaker version; the monotone instrumental variable assumption. A variable v is a monotone instrumental variable (MIV) in the sense of mean-monotonicity if it holds that

$$m_1 \leq m \leq m_2 \Rightarrow E[y(t)|v = m_1] \leq E[y(t)|v = m] \leq E[y(t)|v = m_2] \tag{10}$$

So instead of assuming mean-independence, the monotone instrumental variable assumption allows for a weakly monotone relation between the variable v and the mean schooling function of the child (Manski and Pepper (2000)).

To use the MIV assumption we can again divide the sample into sub-samples on the basis of v and obtain bounds for each sub-sample. From equation (10) it follows that $E[y(t)|v = m]$ is no lower than the lower bound on $E[y(t)|v = m_1]$ and it is no higher than the upper bound on $E[y(t)|v = m_2]$. For the sub-sample where v has the value m we can thus obtain a new lower bound, which is the largest lower bound over all the sub-samples where v is lower than or equal to m . Similarly we can obtain a new upper bound by taking the smallest upper bound over all sub-samples with a value of v higher than or equal to m .

Figure 2: MIV-Bounds, an Example

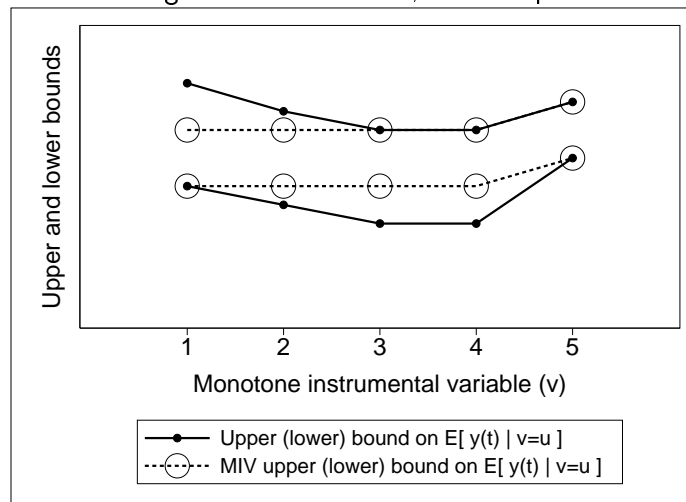


Figure 2 shows an example whereby the black dots (connected by a solid line) are lower and upper bounds for 5 values of a fictive monotone instrumental variable. If we look at the sub-sample with $v = 2$ we can take the maximum lower bound over all sub-samples with a value of v lower or equal than 2. This is the lower bound at $v = 1$, so this becomes the MIV lower bound at $v = 2$. For the upper bound we can take the lowest upper bound over all values of $v \geq 2$. This turns out to be the upper bound at $v = 3$, so this becomes the MIV upper bound at $v = 2$. By repeating this for all values of v we get the MIV bounds which are shown by the hollow circles (connected by the dashed line) in Figure 2. Equation (11) shows aggregate MIV bounds which are obtained by taking the weighted average of the MIV bounds over v .

$$\begin{aligned}
 & \text{MIV-bounds} \\
 & \sum_{m \in M} P(v = m) \cdot [\max_{m_1 \leq m} LB_{E[y(t)|v=m_1]}] \\
 & \leq E[y(t)] \leq \\
 & \sum_{m \in M} P(v = m) \cdot [\min_{m_2 \geq m} UB_{E[y(t)|v=m_2]}]
 \end{aligned} \tag{11}$$

Two monotone instrumental variables will be used in the analysis in this paper. The first is the schooling of the grandparent. By using grandparent's schooling as a MIV we assume that the mean schooling function of the child is monotonically increasing (or non-decreasing) in the schooling of the grandparent. To see what this means, suppose a grandparent can have three levels of schooling (low, middle, high) which would give the following sub-samples (1) children with a low educated grandparent, (2) children with a grandparent with a middle schooling level and (3) children with a high educated grandparent. If we would be able to give each mother (father) in each sub-sample the same schooling level t , the MIV assumption states that the average schooling level we would observe for the children in group (3) would be weakly higher than the mean schooling level we would observe for group (2) which would in turn be weakly higher than the mean schooling of the children in group (1). So in contrast to an instrumental variable assumption the monotone instrumental variable assumption allows for a direct impact of grandparent's schooling on child's schooling as long as this effect is not negative.

The second MIV that will be used is the schooling of the other parent. When we obtain bounds on the effect of mother's schooling we will use the level of schooling of the father as MIV, and when we obtain bounds on the effect of father's schooling we will use the schooling level of the mother as MIV. As with grandparent's schooling the schooling of the other parent is unlikely to satisfy the mean-independence assumption in equation (8), we will therefore use it as a MIV and assume that the mean schooling function

of the child is non-decreasing in the schooling of the other parent. This allows for direct impact of the other parent's schooling on child's schooling as long as this effect is not negative.

Using the two MIV's simultaneously

Instead of using grandparent's schooling or the schooling of the other parent separately to obtain MIV bounds it is also possible to use the two monotone instrumental variables simultaneously. A way to combine two monotone instrumental variables is to apply the following assumption (Manski and Pepper (1998)):

$$E[y(t)|(v_a, v_b) = (u'_a, u'_b)] \geq E[y(t)|(v_a, v_b) = (u_a, u_b)] \quad (12)$$

for all $[(u'_a, u'_b), (u_a, u_b)]$ s.t. $u'_a \geq u_a$ and $u'_b \geq u_b$

Whereby v_a is grandparent's schooling and v_b is the schooling of the other parent. Suppose there are two levels of grandparent's schooling (low and high) and two levels for the schooling of the other parent (low and high) which would give four sub-samples. If all mothers (or fathers) in each sub-sample would have the same schooling level t , the assumption in equation (12) implies that the children in the sub-sample with a high educated grandparent and a high educated other parent will have a weakly higher mean schooling level than children in a sub-sample in which the grandparent or the other parent (or both) have a low schooling level. This assumption says nothing about the relative magnitudes of the mean schooling levels when we compare a sub-sample with a high educated grandparent and a low educated other parent with a sub-sample with a low educated grandparent and a high educated other parent. The assumption in equation (12) implies semi-monotonicity instead of monotonicity. The set up of the bounds using assumption (12) is very similar to the MIV bounds in equation (11) except that the maxima and minima are taken over pairs of values of grandparent's and the other parent's schooling that are ordered.

Bounds on the effect of increasing parents' schooling

Equations (3), (5), (7) and (11) show bounds on $E[y(t)]$ but we are interested in the effect of increasing father's/mother's schooling from s to t on child's schooling $\Delta(s, t) = E[y(t)] - E[y(s)]$. To obtain bounds on this average treatment effect we will subtract the lower (upper) bound on $E[y(s)]$ from the upper (lower) bound on $E[y(t)]$ to get the upper (lower) bound. Under the MTR assumption an increase

in parents' education cannot be negative, the lower bound on $\Delta(s, t)$ is therefore never below zero.⁴

Estimation and inference

Estimation of the lower and upper bounds is straightforward. By plugging in the sample means and empirical probabilities in the formula's in equations (3), (5), (7) and (11) we obtain the No-assumption, MTR, MTR-MTS and MTR-MTS-MIV bounds. All bounds are consistent under the maintained assumptions, but as is noted by Manski and Pepper (2000) the MIV bounds may have finite-sample biases due to the fact that the bounds are obtained by taking maxima and minima over collections of nonparametric regression estimates.⁵

Kreider and Pepper (2007) propose a bias-correction method that uses the bootstrap distribution to estimate the finite-sample bias. Suppose $\hat{\theta}$ is the initial estimate of an MIV upper or lower bound and θ_k is the estimate of the k^{th} bootstrap replication, the bias is then estimated as $bias = \left(\frac{1}{K} \sum_{k=1}^K \theta_k \right) - \hat{\theta}$. The bias-corrected MIV-bounds are obtained by subtracting the estimated biases from the estimated upper and lower bounds. In order to show the size of the finite-sample biases in the present application the tables in this paper report both the bounds with and without the bias-correction.⁶

Inference is more challenging than estimation as the current literature is inconclusive about the type of confidence interval that should be used in a partial identification analysis. There are a number of papers that propose confidence intervals that cover the identified set (interval between upper and lower bound) with a fixed probability (for example Chernozhukov Hong and Tamer (2007) and Beresteanu and Molinari (2008)). Imbens and Manski (2004) propose a confidence interval which does not cover the identified set with fixed probability but rather the parameter of interest is covered with fixed probability. This confidence interval can be used when the parameter is partially identified, but also when the parameter is point-identified. This paper reports nonparametric bounds as well point estimates. The tables therefore report

⁴The bounds using MIV assumption(s) do *not* use the following assumption to obtain bounds on $\Delta(s, t)$, which is stronger than the assumption in equation (10):

$$m_1 \leq m \leq m_2 \Rightarrow E[\Delta(s, t)|v = m_1] \leq E[\Delta(s, t)|v = m] \leq E[\Delta(s, t)|v = m_2] \quad (13)$$

Using assumption (13) instead of assumption (10) would mean that we obtain bounds on the effect of an increase in father's/mother's schooling ($\Delta(s, t)$) for each sub-sample and thus conditional on the monotone instrumental variable. This could be problematic when using the schooling of the spouse as MIV since part of the effect of increasing mother's (father's) schooling could be through the effect that she(he) marries a higher schooled spouse. However, since we do not use assumption (13) but instead use assumption (10), this is not an issue in the analysis in this paper.

⁵This concern does not arise for the bounds based only on the MTR and MTS assumptions, since these do not require taking maxima and minima.

⁶In addition, Efron and Tibshirani (1993) state that bias estimation is worthwhile but that bias correction can be dangerous in practice due to high variability in the estimated bias, this provides an additional reason for showing both the bounds with and without the bias-correction.

the confidence intervals proposed by Imbens and Manski (2004) such that the same type of confidence interval is used for the bounds and the point estimates. Equation (14) shows the $(1 - \alpha)$ -percent confidence interval:

$$CI_{1-\alpha} = \left(\hat{l}b - c_{IM} \cdot \hat{\sigma}_{lb} \quad , \quad \hat{u}b + c_{IM} \cdot \hat{\sigma}_{ub} \right) \quad (14)$$

whereby $\hat{l}b$ and $\hat{u}b$ are the estimated upper and lower bounds and $\hat{\sigma}_{lb}$ and $\hat{\sigma}_{ub}$ are the estimated standard errors of the estimated lower and upper bounds, obtained by 1000 bootstrap replications. Since there are multiple children from one family, the sample drawn during each replication is a bootstrap sample of clusters, with the family as cluster-id. The parameter c_{IM} is obtained by solving equation (15).⁷

$$\Phi \left(c_{IM} + \frac{(\hat{u}b - \hat{l}b)}{\max\{\hat{\sigma}_{lb}, \hat{\sigma}_{ub}\}} \right) - \Phi(-c_{IM}) = 1 - \alpha \quad (15)$$

For the bias-corrected bounds two types of confidence intervals will be reported. The first is the confidence interval reported in equation (14) whereby $\hat{l}b$ and $\hat{u}b$ are replaced with the bias-corrected upper and lower bounds. A disadvantage is that this confidence interval does not take into account that the bias is estimated and not a fixed number. The second is the bias-corrected percentile confidence interval (Efron and Tibshirani (1993)). If there is no finite-sample bias this confidence interval will consist of the $(\alpha/2)$ th bootstrap quantile of the lower bound and the $(1 - \alpha/2)$ th bootstrap quantile of the upper bound. If there is finite-sample bias the bootstrap quantiles are adjusted to take into account that in case of finite-sample bias the share of bootstrap elements lower than $\hat{\theta}$ will be more (or less) than 50%.⁸ A drawback of this confidence interval is that it is developed for bias-corrected point estimates and there is uncertainty about the exact coverage properties in the case of partially identified parameters. Since each of the two confidence intervals has its advantages and disadvantages, the tables will report both.

⁷For point identified parameters the Imbens-Manski confidence interval reduces to $CI_{1-\alpha} = (\hat{\beta} \pm c_{IM} \cdot \hat{\sigma}_{\beta})$ with c_{IM} solving $\Phi(c_{IM}) - \Phi(-c_{IM}) = 1 - \alpha$.

⁸When using the bias-corrected percentile method we take the p_1 th quantile of the bootstrap distribution of the lower bound and the p_2 th quantile of the bootstrap distribution of the upper bound. The p_1 th and p_2 th quantiles are defined as $p_1 = \Phi(2z_0 - z_{1-\alpha/2})$ and $p_2 = \Phi(2z_0 + z_{1-\alpha/2})$ whereby $z_0 = \Phi^{-1}(\#(\theta_i \leq \hat{\theta})/k)$ with $\#(\theta_i \leq \hat{\theta})$ the number of elements of the bootstrap distribution that are less than or equal to the estimated upper or lower bound and Φ is the standard cumulative normal distribution.

3 Data

A data set which can be used to estimate the causal effect of parents' schooling on child's schooling using a nonparametric bounds analysis has to have (among others) two important properties. The first is that it should contain information on completed schooling outcomes of three successive generations; children, parents and grandparents. The second important property is that the data set should contain enough observations. When using the nonparametric bounds analysis to estimate intergenerational schooling effects it is necessary to have sufficient observations for each level of father's /mother's schooling within each sub-sample defined by the value(s) of the monotone instrumental variable(s). Although there are a number of U.S. data sets that have one of the two properties there are very few data sets that have many observations and contain information on completed schooling outcomes of three successive generations.⁹ An exception is the Wisconsin Longitudinal Study (WLS), which is the data set used in the analysis in this paper.

The data collection of the WLS started in 1957 with a random sample of 10,317 men and women who graduated from Wisconsin high schools. The original respondents were contacted again in 1975, 1992, and 2004. Next to information about the graduates the sample contains comparable data for a randomly selected sibling of most of the respondents. These siblings were contacted in 1977, 1994, and 2005. We will mainly use the data from the most recent waves (2004, 2005) since these contain updated information about completed schooling of the graduates and their spouses, the selected siblings and their spouses and about the children of both the graduates and the selected siblings. In 2004/2005 information is collected from 7,265 graduates and 4,271 siblings.

The sample that is used in this paper includes graduates and siblings who were married at least once and who have at least one child. It is not possible to link children to spouses, but only possible to link children to respondents and spouses to respondents. The sample is therefore further restricted to respondents who only have children from their first marriage to be sure that both the spouse and the respondent are the child's biological parents. This gives a final sample of 21,545 children of 5,167 graduates and 2,524 selected siblings.¹⁰

⁹For example the NLSY and the PSID contain completed schooling outcomes only for part of the children in the data set. As is shown in De Haan and Plug (2010) using a sample in which part of the children has not completed their schooling yet complicates the analysis and can create a bias in the estimates.

¹⁰There are some children below the age of 23 and these children might still be in school. In the analysis we eliminate these observations. De Haan and Plug (2010) show that when 23% of the sample is censored, eliminating children who are still in school can cause a small positive bias. In the sample in this paper only 1.5% is below the age of 23. It is unlikely that eliminating these observations can cause a significant bias in the estimates.

Table 1: Summary Statistics

	Mean	SD
Years of schooling child	14.50	2.322
College degree child	0.458	0.498
College degree mother	0.180	0.384
College degree father	0.280	0.449
Gender child (female=1)	0.495	0.500
Age child	38.34	5.501
Schooling level mother:		
1. Less than high school	0.035	0.183
2. High school	0.627	0.484
3. Some college	0.158	0.365
4. Bachelor's degree	0.131	0.338
5. Master's degree or more	0.048	0.215
Schooling level father:		
1. Less than high school	0.082	0.275
2. High school	0.495	0.500
3. Some college	0.142	0.349
4. Bachelor's degree	0.141	0.348
5. Master's degree or more	0.140	0.346
Schooling level grandparent (mother's parent):		
1. Elementary school	0.147	0.354
2. Middle school	0.343	0.475
3. Some high school	0.101	0.301
4. High school degree	0.274	0.446
5. More than a high school degree	0.136	0.343
Schooling level grandparent (father's parent):		
1. Elementary school	0.143	0.350
2. Middle school	0.359	0.480
3. Some high school	0.092	0.289
4. High school degree	0.268	0.443
5. More than a high school degree	0.138	0.345
N	21,545	

NOTE.—Number of observations on schooling grandparent is smaller, N=16912 for mother's parent and N=14614 for father's parent.

The analysis in this paper will look at the effect of increasing parent's schooling from level s to level t on child's schooling. We will consider the following schooling levels for parents; [1] Less than high school (<12 years), [2] High school (12 years), [3] Some college (13-15 years), [4] Bachelor's degree (16 years) and [5] Master's degree or more (>16 years). Next to considering these five schooling levels the analysis will also look at the effect of having a parent with a college degree, whereby parents with less than college have schooling level 1, 2 or 3 and parents with a college degree have schooling levels 4 or 5.¹¹

¹¹Parents' schooling is available in years, but for some observed years of schooling there are very few observations. It

For the schooling of the grandparent we will use the schooling of the head of the household when the parent was 16 (in 80-90% of the cases the father is the head of the household).¹² Since the average schooling level has increased over time we will use different schooling levels for grandparents; [1] Elementary school (≤ 6 years), [2] Middle school (7-8 years), [3] Some high school (9-11 years), [4] Graduated from high school (12 years) and [5] More than high school (≥ 13 years). Table 1 gives some descriptive statistics.

4 Results

Many studies show that there is a strong positive association between parents' schooling and child's schooling. To investigate whether this positive association reflects a causal relation, the analysis below will compare the results of the nonparametric bounds analysis with the results of using an exogenous treatment selection assumption (ETS). The exogenous treatment selection assumption implies that $E[y(t)|z \neq t] = E[y|z = t]$ and yields point identification. It assumes that the schooling level of fathers and mothers is unrelated to unobserved factors affecting child's schooling (like child rearing talents or heritable endowments). Exogenous treatment selection is also assumed when regressing child's years of schooling on years of schooling of his parents. We will however not assume a linear effect of the years of schooling of the parent but instead estimate the effect of moving from one level of parental schooling to the next. Therefore we will compare the results of the bounds analysis with the results of an ETS assumption, which is the same as running OLS on child's schooling with one dummy variable for each level of mother's (father's) schooling.

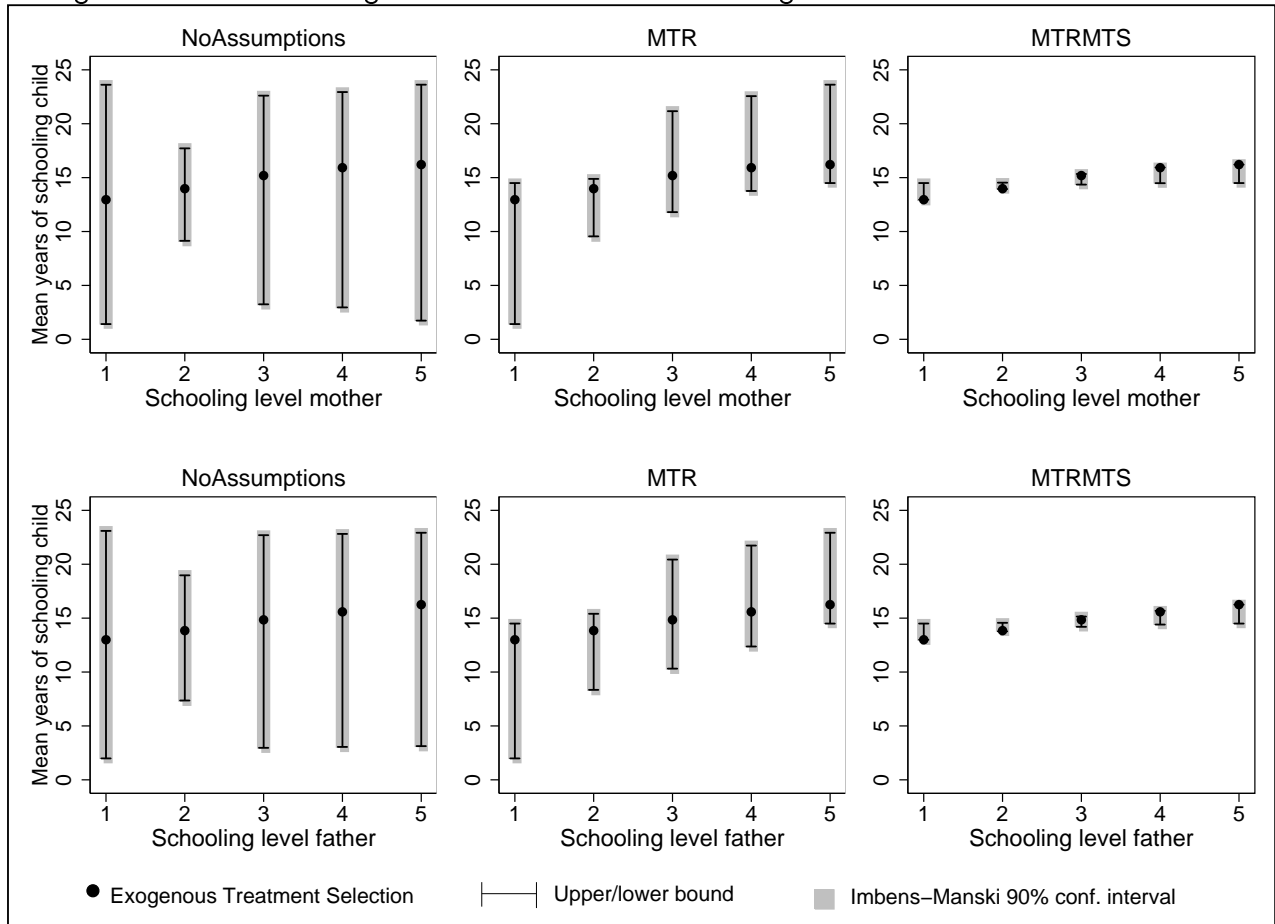
Figure 3 shows nonparametric bounds on mean years of schooling as a function of mother's and father's level of schooling, as well as the exogenous treatment selection point estimates. The nonparametric bounds are rather precisely estimated as is shown by the confidence intervals depicted by the gray areas in Figure 3. The two panels on the left show the no-assumption bounds.¹³ Both for mother's schooling as for father's schooling these bounds are very wide. We need to impose some assumptions in order to say something informative about the impact of parents' schooling on child's schooling.

is possible to obtain estimates of the bounds on the effect of increasing parents' schooling from t to $t + 1$ years, but it is complicated to obtain bootstrapped confidence intervals. If the sample contains very few observations with t or $t + 1$ years of parents' schooling, there will be many bootstrap replications with no observations for t or $t + 1$ years of parents' schooling.

¹²Unfortunately this variable is not available for the spouse of the selected sibling.

¹³For the no-assumption bounds and the MTR-bounds we take the lowest years of schooling of the child observed in the data (1 year) as y_{min} and the highest observed years (24 years) as y_{max} .

Figure 3: Child's Schooling as Function of Parents' Schooling: Bounds and ETS Point Estimates



NOTE.—Levels of schooling 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree, 5: Master's degree or more.

In the middle panels the monotone treatment response assumption is added. Although this reduces the width of the bounds, the MTR-bounds are still too wide to be really informative. In the two panels on the right the monotone treatment selection assumption is added to get the MTR-MTS bounds. As was already stated in Section 2 this combined MTR-MTS assumption can be tested as $E[y|z = u]$ must be weakly increasing in u . Table 2 shows that the MTR-MTS assumption is not rejected as average years of child's schooling is indeed weakly increasing both in the level of mother's schooling as in the level of father's schooling.

Adding the monotone treatment selection assumption strongly reduces the width of the bounds. By combining two relatively weak nonparametric assumptions we obtain bounds which are a lot tighter than the no-assumptions bounds. The two panels on the right in Figure 3 also show that for the lowest levels of mother's and father's schooling the exogenous treatment selection point estimates almost coincide

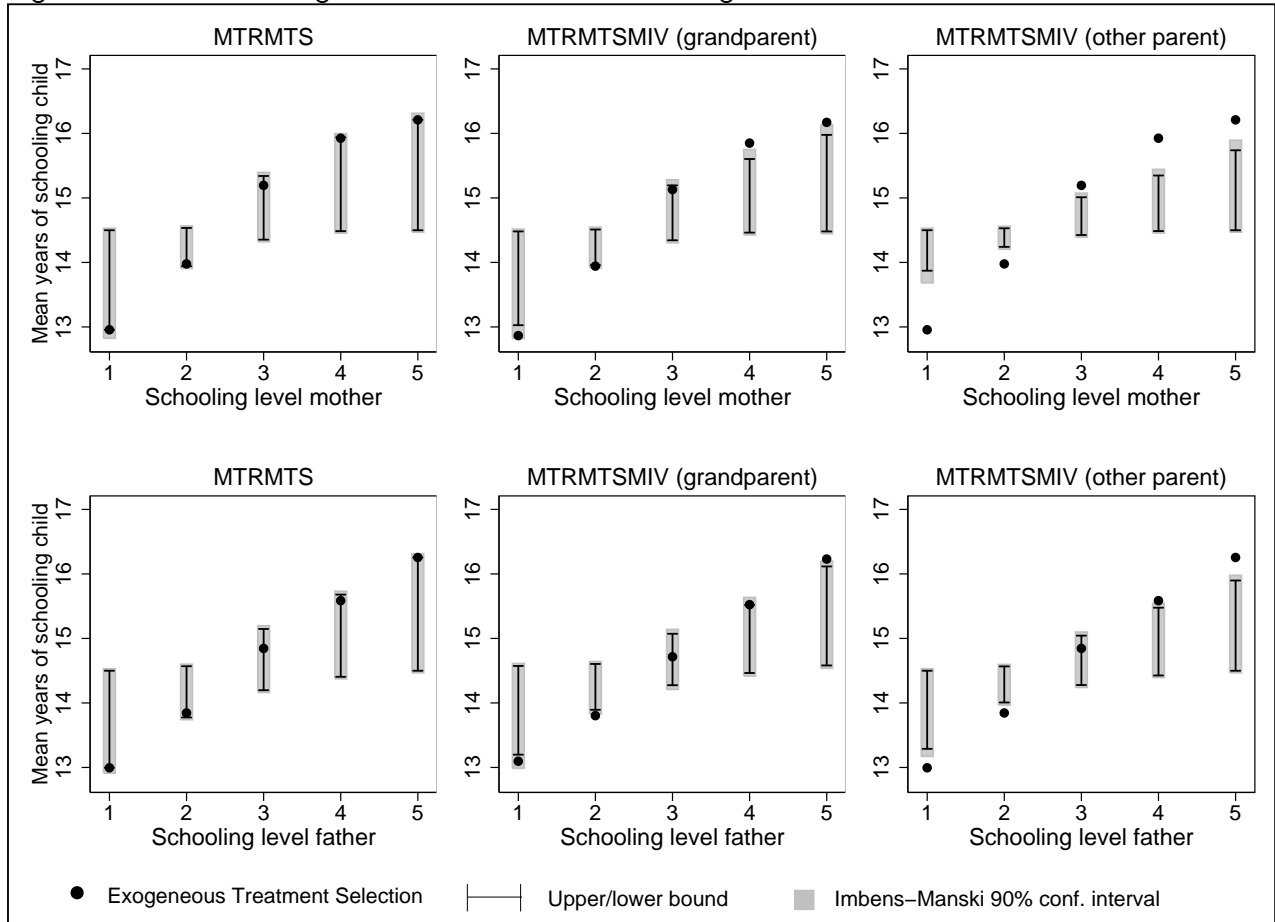
Table 2: Mean Schooling Child by Schooling Level Parent

Schooling level parent	Mothers	Fathers
	$E[y z = u]$	$E[y z = u]$
1: Less than high school	12.96	13.00
2: High school	13.98	13.85
3: Some college	15.19	14.85
4: Bachelor's degree	15.93	15.59
5: Master's degree or more	16.21	16.26
N	21,545	21,545

NOTE.—MTR-MTS assumption not rejected as average years of child's schooling is weakly increasing both in the level of mother's schooling as in the level of father's schooling.

with the lower bounds, while for the highest levels they almost coincide with the upper bounds. This is even more clear in Figure 4, which also shows the MTR-MTS bounds (in the panels on the left) but with a different scale.

Figure 4: Child's Schooling as Function of Parents' Schooling: MTRMTS+MIV Bounds & ETS Estimates



NOTE.—Levels of schooling 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree, 5: Master's degree or more. Sample using schooling grandparent as MIV is smaller N=16912 for mothers and N=14614 for fathers.

Figure 4 also shows bounds whereby the MIV assumption is added. The middle two panels show the bounds using grandparent's schooling as a monotone instrumental variable and the two panels on the right use the schooling of the other parent as a MIV.¹⁴ Using the schooling of the grandparent as MIV gives bounds which are tighter than the MTR-MTS bounds. If we compare the bounds with the ETS point estimates we see that both for mothers as for fathers the ETS results fall outside the bounds for the highest and lowest levels of parents' schooling. For fathers the point estimates are all within the confidence intervals, but for mothers two point estimates fall outside the confidence intervals around the bounds. The identifying power of the schooling of the other parent is even stronger. Using the other parent's schooling as a MIV again reduces the width of the bounds compared to the MTR-MTS bounds and now the point estimates fall outside the confidence intervals around the bounds for all levels of mother's schooling and for fathers this is true for the lowest and highest levels of schooling. ETS underestimates for low levels and overestimates for high levels of parents' schooling.

Bounds on the average treatment effect

So far we have only looked at bounds on $E[y(t)]$ while we are interested in the effect of increasing parents' schooling from one level to the next. Table 3 shows bounds on $\Delta(s, t) = E[y(t)] - E[y(s)]$ for mother's and father's level of schooling. To account for sampling variability Table 3 also shows the confidence intervals which were described in Section 2. Since Figure 3 shows that the no-assumption bounds and the MTR bounds are rather wide and not very informative these are not shown in Table 3.¹⁵

The OLS results (based on the ETS assumption) range from an increase of 0.28 years of schooling when increasing mother's schooling from a bachelor's degree to a master's degree ($\Delta(4, 5)$), to an increase of 1.22 years when increasing mother's schooling from high school to some college ($\Delta(2, 3)$). The ETS results on father's schooling seem to be more constant, although the results also vary, from 0.67 years for $\Delta(4, 5)$ to 1 year for $\Delta(2, 3)$.

¹⁴When using schooling of the other parent as a MIV, MTRMTS-bounds have to be calculated for each sub-sample defined by the schooling level of the other parent. Since the lowest category (less than high school) contains only a small amount of observations level 1 and 2 are combined when using the schooling of the other parent as a MIV.

¹⁵Results are available upon request.

Table 3: ETS Point Estimates and Nonparametric Bounds on Intergenerational Schooling Effects

	ETS		MTRMTS		MTRMTSMIV Grandparent				MTRMTSMIV Other parent			
	β		LB	UB	LB	UB	biascorrected		LB	UB	biascorrected	
							LB	UB			LB	UB
EFFECT OF MOTHER'S SCHOOLING												
$\Delta(1,2)$	1.021		0	1.578	0	1.483	0	1.535	0	0.658	0	0.655
	(0.847	1.196)	(0	1.714)	(0	1.690)	{0	1.741}	(0	0.844)	{0	0.841}
							[0	1.776]			[0	0.877]
$\Delta(2,3)$	1.218		0	1.399	0	1.236	0	1.273	0	0.769	0	0.768
	(1.123	1.313)	(0	1.463)	(0	1.329)	{0	1.366}	(0	0.840)	{0	0.839}
							[0	1.381]			[0	0.860]
$\Delta(3,4)$	0.731		0	1.586	0	1.259	0	1.262	0	0.922	0	0.922
	(0.622	0.841)	(0	1.650)	(0	1.410)	{0	1.412}	(0	1.019)	{0	1.019}
							[0	1.423]			[0	1.041]
$\Delta(4,5)$	0.284		0	1.725	0	1.515	0	1.569	0	1.252	0	1.270
	(0.136	0.432)	(0	1.831)	(0	1.674)	{0	1.728}	(0	1.408)	{0	1.427}
							[0	1.739]			[0	1.475]
$\Delta(1,5)$	3.254		0	3.254	0	2.951	0	3.053	0	1.867	0	1.882
	(3.039	3.469)	(0	3.422)	(0	3.204)	{0	3.306}	(0	2.111)	{0	2.126}
							[0	3.323]			[0	2.185]
EFFECT OF FATHER'S SCHOOLING												
$\Delta(1,2)$	0.850		0	1.573	0	1.406	0	1.516	0	1.278	0	1.291
	(0.735	0.964)	(0	1.661)	(0	1.618)	{0	1.728}	(0	1.397)	{0	1.410}
							[0	1.727]			[0	1.449]
$\Delta(2,3)$	1.001		0	1.372	0	1.179	0	1.190	0	1.036	0	1.034
	(0.902	1.099)	(0	1.433)	(0	1.267)	{0	1.279}	(0	1.102)	{0	1.100}
							[0	1.300]			[0	1.118]
$\Delta(3,4)$	0.740		0	1.481	0	1.243	0	1.260	0	1.199	0	1.201
	(0.629	0.852)	(0	1.542)	(0	1.374)	{0	1.391}	(0	1.273)	{0	1.276}
							[0	1.396]			[0	1.289]
$\Delta(4,5)$	0.669		0	1.850	0	1.652	0	1.684	0	1.472	0	1.473
	(0.568	0.770)	(0	1.912)	(0	1.743)	{0	1.775}	(0	1.555)	{0	1.556}
							[0	1.787]			[0	1.582]
$\Delta(1,5)$	3.260		0	3.260	0	2.918	0	3.048	0	2.610	0	2.624
	(3.132	3.387)	(0	3.359)	(0	3.144)	{0	3.275}	(0	2.751)	{0	2.765}
							[0	3.258]			[0	2.810]

NOTE.—Dependent variable is child's years of schooling. Numbers between parentheses are Imbens-Manski 90% confidence intervals. Numbers between braces are Imbens-Manski 90% confidence intervals using bias-corrected MIV bounds. Numbers between brackets are the biased-corrected percentile confidence intervals. To adjust for the fact that the sample contains multiple children from one family the sample drawn during each replication is a bootstrap sample of clusters. Levels of schooling: 1: Less than high school, 2: High school, 3: Some college, 4: Bachelor's degree, 5: Master's degree or more. Number of observations is 21545. Sample using schooling grandparent is smaller; N=16912 (mothers) & N=14614 (fathers).

The monotone treatment response assumption in combination with the monotone treatment selection assumption gives bounds on the treatment effects ranging from an effect between 0 and 1.40 years when increasing mother's schooling from high school to some college, to an effect between 0 and 1.73 years when increasing mother's schooling from a bachelor's degree to a master's degree. For the same increases in father's schooling the effects are respectively within $[0, 1.37]$ years and within $[0, 1.85]$ years. Since these bounds include a zero effect as well as an effect as large as the ETS point estimates they are not very instructive.

When we use grandparent's schooling as a monotone instrumental variable we get upper bounds that are lower than the MTR-MTS bounds, but they are higher than the ETS results. Using the schooling of the other parent as a MIV gives bounds which are more informative. For the effect of increasing mother's schooling from high school to some college ($\Delta(2, 3)$) the ETS result falls outside the confidence intervals around the (bias-corrected) MTR-MTS-MIV bounds.

Instead of looking at the effect of increasing parents' schooling from one level to the next we can also look at the effect of increasing parents' schooling from the lowest level (less than high school) to the highest level (a master degree or more) $\Delta(1, 5)$. The ETS result in Table 3 indicates that increasing mother's schooling from the lowest to the highest schooling level increases child's schooling on average by 3.25 years. This point estimate is higher than the MTR-MTS-MIV upper bounds when we use grandparent's schooling or the schooling of the other parent as MIV. This is the case for the bounds with or without the bias-correction, in both cases the confidence interval(s) exclude the ETS point estimate. Using grandparent's schooling as a monotone instrumental variable gives an upper bound of 3 years and using the schooling of the other parent as MIV gives an even lower upper bound of 1.9 years which is almost half the ETS estimate.

The results are very similar for the effect of father's schooling. The ETS estimate of increasing father's schooling from less than high school to a master's degree or more is 3.26 which is almost identical to the ETS estimate of the effect of mother's schooling. The (bias-corrected) MTR-MTS-MIV estimates show that OLS (ETS) also overestimates the effect of father's schooling, since both upper bounds using either grandparent's schooling or the schooling of the other parent as a MIV are significantly lower than the OLS results.

Table 4: ETS Point Estimates and Nonparametric Bounds on Effect of Increasing Parents' Schooling to a College Degree

	ETS		MTRMTSMIV			MTRMTSMIV			MTRMTSMIV				
			Grandparent			Other parent			Grandparent+other parent				
	β	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB		
EFFECT ON CHILD'S YEARS OF SCHOOLING													
Mother with college degree	1.809	0	1.523	0	1.525	0	1.088	0	1.086	0.048	0.873	0.025	0.997
	(1.726 1.892)	(0	1.649)	{0	1.651}	(0	1.183)	{0	1.181}	(0.014	1.018)	{0	1.141}
		[0	1.671]	[0	1.671]	[0	1.200]	[0	1.200]			[0.005	1.073]
Father with college degree	1.943	0.008	1.702	0	1.716	0	1.437	0	1.437	0.072	1.162	0.052	1.269
	(1.865 2.021)	(0	1.800)	{0	1.814}	(0	1.509)	{0	1.510}	(0.028	1.289)	{0.008	1.396}
		[0	1.840]	[0	1.840]	[0	1.527]	[0	1.527]			[0.017	1.342]
EFFECT ON PROBABILITY CHILD HAS COLLEGE DEGREE													
Mother with college degree	0.365	0	0.306	0	0.307	0	0.214	0	0.214	0.010	0.171	0.001	0.209
	(0.347 0.383)	(0	0.335)	{0	0.336}	(0	0.236)	{0	0.236}	(0.001	0.210)	{0	0.248}
		[0	0.340]	[0	0.340]	[0	0.241]	[0	0.241]			[0.001	0.214]
Father with college degree	0.393	0.001	0.347	0	0.351	0	0.300	0	0.301	0.017	0.234	0.010	0.257
	(0.376 0.409)	(0	0.367)	{0	0.372}	(0	0.316)	{0	0.317}	(0.007	0.263)	{0.0002	0.285}
		[0	0.375]	[0	0.375]	[0	0.323]	[0	0.323]			[0.003	0.273]

NOTE.—Numbers between parentheses are Imbens-Manski 90% confidence intervals. Numbers between braces are Imbens-Manski 90% confidence intervals using bias-corrected MIV bounds. Numbers between brackets are the biased-corrected percentile confidence intervals. To adjust for the fact that the sample contains multiple children from one family the sample drawn during each replication is a bootstrap sample of clusters. Results on effect mother's schooling are based on 16912 observations. Results on the effect of father's schooling are based on 14614 observations.

Increasing parents' schooling to a college degree

Increasing parents' schooling from less than high school to a master degree or more is a very big change and probably not the most interesting treatment effect to look at. Most parents finished high school, but only some went to college and obtained a bachelor's or master's degree. A lot of variation in schooling levels between parents is due to the fact that some parents obtained a college degree while others did not. It is therefore more interesting to look at the average treatment effect of increasing parents' schooling to a college degree. Table 4 shows the effect of increasing parents' schooling from less than college to a college degree on child's years of schooling.

Table 4 shows that under the exogenous treatment selection assumption increasing mother's schooling to a college degree increases child's schooling on average by 1.8 years. If we, however, use grandparent's schooling or the schooling of the other parent as MIV we obtain upper bounds which are significantly lower than 1.8. When we use grandparent's schooling as MIV we obtain an upper bound of 1.5 years, and when we use the schooling of the father as a monotone instrumental variable we obtain an upper bound of only 1.1 years. For both MIV's the point estimate based on the exogenous selection assumption falls outside the bootstrapped confidence intervals around the bounds.¹⁶

A similar pattern is observed when we look at the effect of increasing father's schooling to a college degree. The ETS estimate of this treatment effect is equal to 1.9 years. This is significantly different from the upper bound using either grandparent's schooling or the schooling of the other parent as MIV. Using grandparent's schooling as MIV gives an upper bound on the effect of increasing father's schooling to a college degree of 1.7 and using mother's schooling as MIV gives an upper bound of 1.4.¹⁷

These nonparametric bounds do not exclude a zero effect of parents' schooling on years of schooling of the child, but the upper bounds are informative since they are substantially smaller than the point estimates obtained under the exogenous treatment selection assumption.

Using the two MIV's simultaneously

As was discussed in Section 2, instead of using grandparent's schooling and the schooling of the other parent separately to obtain MTR-MTS-MIV bounds it is also possible to use the two monotone instrumental variables simultaneously.¹⁸ The results in the final column of Table 4 show that using the two

¹⁶This holds both for bounds with and without the bias-correction and irrespective of the type of confidence interval.

¹⁷This also holds both for bounds with and without the bias-correction and irrespective of the type of confidence interval.

¹⁸Unfortunately it is not possible to use both MIV's simultaneously in Table 3. When considering five different schooling levels of the parent, for some combination of values of the MIV's not all schooling levels of the parent have a positive

MIV's simultaneously gives very informative bounds. The upper bounds are lowered and show that the effect of increasing mother's schooling to a college degree increases child's schooling at most with 0.9 years and for fathers the effect is at most 1.2 years. Both upper bounds are substantially lower than the OLS (ETS) estimates.

Combining the two monotone instrumental variables not only gives more informative upper bounds, but also more informative lower bounds. Both for mothers as for fathers the results show that the effect of increasing their schooling to a college degree has an impact on child's schooling which is significantly different from zero. Both the bounds with and without the bias-correction are above zero. For fathers all confidence intervals exclude zero and for mothers two of the three confidence intervals exclude a zero effect of mother's schooling.¹⁹

Probability that the child obtains a college degree

Instead of investigating the effect of increasing parents' schooling to a college degree on child's years of schooling it might be instructive to use the same schooling measure for parents and children. The bottom panel of Table 4 therefore shows the effect of increasing parents' schooling to a college degree on the probability that the child obtains a college degree. The results using this outcome variable are very similar to the results using child's years of schooling as outcome variable. Assuming that parents' schooling is unrelated to any unobservables affecting child's schooling gives point estimates which indicate that increasing mother's (father's) schooling to a college degree increases the probability that the child obtains a college degree with 36 (39) percent. These point estimates are substantially larger than the (bias-corrected) MTR-MTS-MIV upper bounds using either grandparent's schooling or the schooling of the other parent as a monotone instrumental variable. Combining the two MIV's gives the most informative bounds. The final column in Table 4 shows that the causal effect of increasing mother's schooling to a college degree on the probability that her child obtains a college degree is above zero and at most 20 percent, which is substantially lower than the ETS point estimate.²⁰ Also for fathers the lower bound is significantly different from zero. The upper bound is a bit higher than for mothers and indicates that increasing father's schooling to a college degree increases the probability that his child obtains a

number of observations. For example there are no fathers with schooling level 1 (less than high school) for the sub-sample where the grandparent has schooling level 4 (graduated from high school) and the other parent (mother) has schooling level 3 (some college). In this case it is not possible to estimate the upper and lower bounds. With two levels of parent's schooling (college degree or less than a college degree) this is not a problem.

¹⁹The Imbens-Manski confidence intervals around the bias-corrected bounds does not exclude zero.

²⁰For mothers, the Imbens-Manski confidence interval around the bias-corrected bounds does not exclude zero.

college degree by at most 25 percent.

5 How do the nonparametric bounds compare to results using other identification approaches?

The previous section showed that combining some weak nonparametric assumptions gives informative upper and lower bounds that exclude zero as well as the OLS estimate. But how do the results using the nonparametric bounds analysis compare to the results using other identification approaches. This section aims to answer this question. First by comparing the bounds with the results of a linear IV using an instrument which has been used by previous studies, but applying it to the WLS. Secondly we will compare the nonparametric bounds to point estimates reported by recent empirical papers.

The recent literature has used several instruments for parents' schooling, some examples are changes in compulsory schooling laws, college openings/closings, variation in tuition fees and variation in the quality of entry exams. An instrument which has been used by previous studies and which is available in the WLS is proximity to college. This instrument has been used to estimate intergenerational schooling effects, for example by Carneiro et al. (2007) and it is related to the instrument used by Currie and Moretti (2003), who use college openings and closings as instruments for mother's schooling. In addition, college proximity has been used as an instrument for schooling by Card (1995), Kling (2001) and Cameron and Taber (2004).

Table 5 show results whereby college proximity is used as an instrument in a 2SLS regression.²¹ The top panel shows results for the effect of mother's schooling. The OLS estimate indicates that an additional year of mother's schooling increases child's schooling by 0.43 years. Using college proximity as an instrument for mother's schooling gives an estimate which is higher than the OLS estimate and equal to 0.75. For fathers the OLS estimate is equal to 0.39. The result of the linear IV is a bit smaller than the corresponding OLS estimate and indicates that increasing father's schooling by one year increases child's schooling by about 0.32 years.

²¹College proximity is a binary variable which is equal to one if the parent attended high school in a community 15 miles or less from state/private college or university, or attended high school in a city with a college or university (private or public) and zero otherwise. This variable is only available for the original respondent of the WLS, which gives a sample of 2727 mothers and 2240 fathers.

Table 5: The Effect of Parent's Schooling on Child's Years of Schooling Using Proximity to College as (M)IV

	OLS		Linear IV		MTRMTSMIV		MTRMTSMIV		MTRMTSMIV					
	β		nearby college	nearby college	nearby college	nearby college+grandparent	nearby college	nearby college+grandparent	nearby college	nearby college+other parent				
			LB	UB	LB	UB	LB	UB	LB	UB				
			biascorrected		biascorrected		biascorrected		biascorrected					
			LB	UB	LB	UB	LB	UB	LB	UB				
			MOTHERS											
Years of education	0.425 (0.397 0.454)	0.752 (0.411 1.093)												
1 st stage F		21.33												
College degree	1.799 (1.669 1.930)	4.619 (2.202 7.037)	0 (0 1.880)	1.777 (0 1.880)	0 {0 1.882}	1.779 {0 1.882}	0.013 (0 1.576)	1.387 (0 1.576)	0 {0 1.660}	1.470 {0 1.660}	0.023 (0 1.238)	1.084 (0 1.238)	0 {0 1.347}	1.192 {0 1.347}
1 st stage F		12.56			[0 1.915]	[0 1.647]			[0 1.886]	[0 1.360]				
			FATHERS											
Years of education	0.399 (0.379 0.420)	0.323 (0.130 0.516)												
1 st stage F		25.75												
College degree	2.015 (1.900 2.130)	1.885 (0.734 3.037)	0 (0 2.056)	1.943 (0 2.056)	0 {0 2.064}	1.950 {0 2.064}	0.052 (0 1.800)	1.619 (0 1.800)	0 {0 1.917}	1.736 {0 1.917}	0.032 (0 1.323)	1.162 (0 1.323)	0.012 {0 1.341}	1.180 {0 1.341}
1 st stage F		21.83			[0 2.076]	[0 1.886]			[0 1.886]	[0 1.360]				

NOTE.—Numbers between parentheses under the OLS and linear IV results are 90% confidence intervals based on cluster-robust standard errors to account for correlation within families. Numbers between parentheses under MTR-MTS-MIV bounds are Imbens-Manski 90% confidence intervals. Numbers between braces are Imbens-Manski 90% confidence intervals using bias-corrected MIV bounds. Numbers between brackets are the biased-corrected percentile confidence intervals. To adjust for the fact that the sample contains multiple children from one family the sample drawn during each replication is a bootstrap sample of clusters. Nearby college is a binary variable which is equal to one if the parent attended high school in a community 15 miles or less from state/private college or university, or attended high school in a city with a college or university (private or public) and zero otherwise. This variable is only available for the original respondent of the WLS, which gives a sample of 2727 mothers and 2240 fathers. Results on the effect of mother's schooling are based on 8122 observations. Results on the effect of father's schooling are based on 6162 observations.

A potential problem is that parents who live(d) nearby a college might not have higher educated children because they obtained more schooling due to the proximity of a college, but because of other unobservable characteristics related to the area in which they live(d). Since an important feature of the partial identification methodology is to assess the sensitivity of inferences to commonly used assumptions, it makes sense to relax the mean-independence assumption and to allow for a monotone (positive) relation between proximity to college and the education of the child. We will do this by using college proximity as a monotone instrumental variable in a nonparametric bounds analysis and compare the results to using college proximity as an instrument in a 2SLS regression model. Since the proximity of a college has an effect on education by affecting the decision of whether or not to go to college, we will focus on the treatment of having a parent with a college degree versus having a parent with less than college.

Table 5 shows OLS and 2SLS estimates of the effect of increasing father's and mother's schooling to a college degree as well as the MTR-MTS-MIV bounds. Comparing the 2SLS point estimates with the results of using college proximity as a MIV shows that the 2SLS estimate of having a mother with a college degree is significantly larger than the (bias-corrected) MTR-MTS-MIV upper bound. For fathers the 2SLS estimate falls within the MTR-MTS-MIV bounds when using only college proximity as MIV. If we however combine college proximity with one of the other MIV's used in the paper, grandparent's schooling or the schooling of the other parent, we see that both for mothers as for fathers the 2SLS estimates are substantially larger than the (bias-corrected) nonparametric upper bounds. Not only the OLS estimates are larger than the nonparametric upper bounds but also 2SLS results using an instrument which has been used by previous studies as instrument for education, fall outside the bounds.

As a second step to compare the nonparametric bounds with findings using other identification approaches, we can compare the results of this paper with point estimates reported by other recent papers. Before comparing the bounds with the estimates obtained using one of the other identification approaches, it is instructive to compare the OLS estimates obtained using the WLS with the OLS estimates of the other recent papers on intergenerational schooling effects. The top panel of Table 6 shows OLS estimates of the effect of an additional year of father's or mother's schooling on child's years of schooling. The OLS estimates obtained on the WLS are in line with the estimates obtained on other U.S. data sets, but they are in general larger than the OLS estimates obtained using data sets from the Northern European countries.

Table 6: Main Assumptions and Results of Different Approaches to Identify Intergenerational Schooling Effects

Main assumptions necessary for consistent estimate of ATE	Study	Country	Additional year schooling	mother	father
OLS					
* Exogenous Treatment Selection (ETS)	This paper	US	0.43		0.40
* Effect parents' schooling is linear	Sacerdote (2000)	US	0.35		0.28
	Behrman and Rosenzweig (2002)	US	0.33		0.47
	Plug (2004)	US	0.54		0.39
	Antonovics and Goldberger (2005)	US	0.28		0.49
	Sacerdote (2007)	US	0.32		-
	Bjorklund, Lindahl and Plug (2006)	Sweden	0.24		0.24
	Holmlund, Lindahl and Plug (2010)	Sweden	0.28		0.23
	Black, Devereux and Salvanes (2005)	Norway	0.24		0.22
	Pronzato (2010)	Norway	0.24		0.21
	Bingley, Christensen and Jensen (2009)	Denmark	0.18		0.18
NONPARAMETRIC BOUNDS					
* MTR, MTS, MIV	This paper	US	(0.01, 0.21)		(0.02, 0.29)
	This paper (bias-corrected)	US	(0.01, 0.25)		(0.01, 0.32)
TWINS					
* Effect parents' schooling is homogeneous or twins are random sample of population	Behrman and Rosenzweig (2002)	US	-0.25		0.36
* Effect parents' schooling is linear	Antonovics and Goldberger (2005)	US	0.03		0.48
* Schooling difference twin parents is exogenous	Pronzato (2009)	Norway	0.10		0.16
	Bingley, Christensen and Jensen (2009)	Denmark	0.05		0.08
	Holmlund, Lindahl and Plug (2010)	Sweden	0.06		0.12
ADOPTEEES					
* Effect parent's schooling is homogeneous or adoptees are random sample of population	Sacerdote (2000)	US	0.22		0.16
* Effect parents' schooling is linear	Plug (2004)	US	0.28		0.27
* Schooling adoptive parents is exogenous	Sacerdote (2007)	US	0.09		-
	Bjorklund, Lindahl and Plug (2006)	Sweden	0.07		0.11
	Holmlund, Lindahl and Plug (2010)	Sweden	0.11		0.10
INSTRUMENTAL VARIABLES					
* Effect parent's schooling is homogeneous or compliers are random sample of population	Black, Devereux and Salvanes (2005)	Norway	0.12		0.04
* Effect parents' schooling is linear	Holmlund, Lindahl and Plug (2010)	Sweden	0.07		0.02
* Instrument is exogenous					
* Relation instrument and parent's schooling is not weak					

NOTE:—The point estimates are from regressions including only one of the parents (not controlling for assortative matching) because the nonparametric bounds analysis in this paper also does not control for assortative matching. OLS point estimates of "This paper" are the linear OLS point estimates which are also shown in Table 5.

Table 6 also shows an overview of the main assumptions and results of the different approaches to identify the effect of mother's and father's schooling. The point estimates reported are estimates of the effect of one additional year of mother's or father's schooling on child's years of schooling.²² The nonparametric bounds reported in this paper are not bounds on one additional year of schooling but are bounds on the average treatment effect of increasing parents' schooling from a certain schooling level to a higher level. This complicates a direct comparison of the bounds with the point estimates.

The tightest bounds are obtained by using grandparent's schooling and the schooling of the other parent simultaneously as MIV's. The last columns in Table 4 show the bounds which are obtained using both MIV's simultaneously. These are bounds on the effect of increasing mother's (father's) schooling to a college degree.²³ Increasing parents' schooling to a college degree corresponds to increasing parents' schooling by about 4 years. The results in the last columns in Table 4 thus indicate that on average one extra year of mother's schooling has an effect on child's years of schooling that lies between 0.01 and 0.21 (0.05/4 and 0.87/4). One extra year of father's schooling has on average an effect on child's years of schooling that lies between 0.02 and 0.29 (0.07/4 and 1.16/4).²⁴

Comparing the nonparametric bounds with the point estimates shows that none of the estimates obtained using data from the Northern European countries fall outside the bounds. If we however look at the results using U.S. data we see that three of the five estimates for mothers are outside the bounds.²⁵ For fathers two of the four estimates are larger than the nonparametric upper bounds. Tables 5 and 6 show that the nonparametric bounds analysis gives informative bounds on intergenerational schooling effects. The OLS estimates, 2SLS estimates using college proximity as instrument as well as a number of point estimates reported in the recent literature fall outside the estimated nonparametric bounds.

²²The point estimates reported in Table 6 are from regressions including only one of the parents (not controlling for assortative matching) because nonparametric bounds analysis in this paper also does not control for assortative matching.

²³Unfortunately using both MIV's simultaneously is not possible when investigating the 5 schooling levels reported in Table 3 due to data limitations, which I describe in footnote 18. The distribution of the data shows large spikes at 12 and 16 years of schooling which shows that the main difference in schooling levels between individuals is due to the fact that some individuals stop after high school while others go to college and obtain a college degree. In addition this treatment is also used to compare the results of using college proximity as instrument in a 2SLS model with using it as a MIV in the nonparametric bounds analysis. For these reasons the bounds reported in Table 4 are used as comparison to the point estimates reported in the recent literature.

²⁴These are the results without the bias-correction, results with the bias-correction are also shown in Table 6.

²⁵When using the bias-corrected bounds as comparison two of the five estimates fall outside the bounds.

6 Conclusion

Regressing child's schooling on parents' schooling generally gives large positive and significant estimates. Since it is not clear whether these associations are informative about the causal impact of parents' schooling, different identification strategies have been used in the recent empirical literature. This recent literature has not reached consensus, especially not regarding the relative importance of mother's and father's schooling for the schooling outcomes of their child. Holmlund et al. (2010) conclude in their overview of this recent literature that none of the applied identification approaches is perfect and that in each case internal or external validity assumptions are easily violated. This motivates the use of an alternative method to identify intergenerational schooling effects.

This paper uses a nonparametric bounds analysis which give bounds on the effect of parents' schooling by relying on a set of weak and in part testable nonparametric assumptions. In addition bounds on the average treatment effect are obtained without having to rely on the assumption of a linear and homogeneous effect of parents' schooling.

The tightest bounds show that the effect of increasing mother's or father's schooling to a college degree has an effect on child's schooling which is significantly different from zero, but substantially lower than the OLS estimates. Although the upper bounds on the impact of father's schooling tend to be a bit higher than the upper bounds on the impact of mother's schooling, the results provide no evidence that fathers matter more than mothers for the schooling outcomes of their child. The bounds show that both parents matter for the schooling outcomes of their child, but not as much as we would think on the basis of the observed positive associations.

A comparison with results using other identification approaches shows that even though the nonparametric bounding method is conservative it gives informative bounds on intergenerational schooling effects. The OLS estimates, 2SLS estimates using college proximity as instrument as well as a number of point estimates reported in the recent literature fall outside the estimated nonparametric bounds. This is mainly due to the informativeness of the upper bounds, since most of these estimates are larger than the upper bounds.

References

- Antonovics, Kate, and Arthur S. Goldberger. 2005. Does increasing women's schooling raise the schooling of the next generation? Comment. *American Economic Review* 95 (December): 1738–1744.
- Becker, Gary S., and Nigel Tomes. 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87 (December): 1153–1189.
- Behrman, Jere R., and Mark R. Rosenzweig. 2002. Does increasing women's schooling raise the schooling of the next generation? *American Economic Review* 92 (March): 323–334.
- Behrman, Jere R., and Mark R. Rosenzweig. 2005. Does increasing women's schooling raise the schooling of the next generation? Reply. *American Economic Review* 95 (December): 1745–1751.
- Beresteanu, Arie and Francesca Molinari. 2008. Asymptotic properties for a class of partially identified models. *Econometrica* 76 (July): 763–814.
- Bingley, Paul, Kaare Christensen and Vibeke M. Jensen. 2009. Parental schooling and child development: Learning from twin parents. Working Paper no. 07:2009, Danish National Centre for Social Research (SFI), Denmark.
- Björklund, Anders C.T., Mikael Lindahl, and Erik J.S. Plug. 2006. The origins of intergenerational associations: Lessons from Swedish adoption data. *Quarterly Journal of Economics* 121 (August): 999–1028.
- Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes. 2005. Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review* 95 (March): 437–449.
- Blundell, Richard, Amanda Gosling, Hidehiko Ichimura, and Costas Meghir. 2007. Changes in the distribution of male and female wages accounting for employment composition using bounds. *Econometrica* 75 (March): 323–363.
- Cameron, Stephen and Christopher R. Taber. 2004. Borrowing constraints and the returns to schooling. *Journal of Political Economy* 112 (February): 132–182.

- Carneiro, Pedro, Costas Meghir, and Matthias Parey. 2007. Maternal education, home environments and the development of children and adolescents. IZA Working Paper no. 3072, Institute for the Study of Labor, Bonn.
- Chernozhukov, Victor, Han Hong and Elie Tamer. 2007. Parameter set inference in a class of econometric models. *Econometrica* 75 (September): 1243–1284.
- Chevalier, Arnaud. 2004. Parental education and child's education: A natural experiment. IZA Working Paper no. 1153, Institute for the Study of Labor, Bonn.
- Currie, Janet, and Enrico Moretti. 2003. Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of Economics* 118 (November): 1495–1532.
- De Haan, Monique, and Erik J.S. Plug. 2010. Estimating intergenerational schooling mobility on censored samples: Consequences and remedies. *Journal of Applied Econometrics* 26 (January/February): 151-166.
- Efron, Bradley and Robert J. Tibshirani. 1993. *An introduction to the bootstrap*. New York: Chapman & Hall.
- Gerfin, Michael, and Martin Schellhorn. 2006. Nonparametric bounds on the effect of deductibles in health care insurance on health care demand – Swiss evidence. *Health Economics* 15 (September): 1011–1020.
- González, Libertad. 2005. Nonparametric bounds on the returns to language skills. *Journal of Applied Econometrics* 20 (September/October): 771–795.
- Gundersen, Craig, and Brent Kreider. 2009. Bounding the effects of food insecurity on children's health outcomes. *Journal of Health Economics* 28 (September): 971-983.
- Guryan, Jonathan, Erik Hurst, and Melissa S. Kearney. 2008. Parental education and parental time with children. *Journal of Economic Perspectives* 22 (Summer): 23–46.
- Haveman, Robert H., and Barbara L. Wolfe. 1995. The determinants of children attainments: A review of methods and findings. *Journal of Economic Literature* 33 (December): 1829–78.

- Hill, Steven C., and Brent Kreider. 2009. Partially identifying treatment effects with an application to covering the uninsured. *Journal of Human Resources* 44 (Spring): 409–449.
- Holmlund, Helena, Mikael Lindahl and Erik J.S. Plug. 2010. The causal effect of parent's schooling on children's schooling: A comparison of estimation methods. *Journal of Economic Literature* (forthcoming).
- Imbens, Guido W., and Charles F. Manski. 2004. Confidence intervals for partially identified parameters. *Econometrica* 72 (November): 1845–1857.
- Kreider, Brent, and John V. Pepper. 2007. Disability and employment: Reevaluating the evidence in light of reporting errors. *Journal of the American Statistical Association* (June): 432-441.
- Kreider, Brent, and John V. Pepper. 2008. Inferring disability status from corrupt data. *Journal of Applied Econometrics* 23 (April): 329–349.
- Maurin, Eric, and Sandra McNally. 2008. Vive la révolution! Long term returns of 1968 to the angry students. *Journal of Labor Economics* 26 (January): 1–35.
- Manski, Charles F. 1989. Anatomy of the selection problem. *Journal of Human Resources* 24 (Summer): 343–360.
- Manski, Charles F. 1997. Monotone treatment response. *Econometrica* 65 (November): 1311–1334.
- Manski, Charles F., and John V. Pepper. 1998. Monotone instrumental variables: With an application to the returns to schooling. Technical Working Paper no. 0224, National Bureau of Economic Research, Cambridge, MA.
- Manski, Charles F., and John V. Pepper. 2000. Monotone instrumental variables: With an application to the returns to schooling. *Econometrica* (July): 997-1010.
- Oreopoulos, Philip, Marianne E. Page, and Ann Huff Stevens. 2006. The intergenerational effects of compulsory schooling. *Journal of Labor Economics* 24 (October): 729–760.
- Pepper, John V. 2000. The intergenerational transmission of welfare receipt: A nonparametric bounds analysis. *Review of Economics and Statistics* 82 (August): 472–478.

- Plug, Erik J.S. 2004. Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. *American Economic Review* 94 (March): 358–368.
- Pronzato, Chiara D. 2010. An examination of paternal and maternal intergenerational transmission of schooling. *Journal of Population Economics* (forthcoming).
- Sacerdote, Bruce I. 2000. The nature and nurture of economic outcomes. Working Paper no. 7949, National Bureau of Economic Research, Cambridge, MA.
- Sacerdote, Bruce I. 2002. The nature and nurture of economic outcomes. *American Economic Review* 92 (May): 344–348.
- Sacerdote, Bruce I. 2007. How large are the effects from changes in family environment? A study of korean American adoptees. *Quarterly Journal of Economics* 122 (February): 119–157.
- Solon, Gary. 1999. Intergenerational mobility in the labor market. In *Handbook of Labor Economics Vol. III*, ed. Orley C. Ashenfelter and David Card. Amsterdam: Elsevier Science.
- Wisconsin Longitudinal Study (WLS). Version 12.20 Preliminary Release (2007). Hauser, R.M; Hauser, T.S. Madison, WI: University of Wisconsin-Madison, WLS; <http://www.ssc.wisc.edu/~wls/documentation>.