UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

WORKING PAPERS SES

Instrument-based estimation with binarized treatments: Issues and tests for the exclusion restriction

Martin Eckhoff Andresen
and
Martin Huber

N. 492
II.2018

Faculté des sciences économiques et sociales
Wirtschafts- und sozialwissenschaftliche Fakultät

Instrument-based estimation with binarized treatments: Issues and tests for the exclusion restriction

Martin Eckhoff Andresen* and Martin Huber**

*Statistics Norway, **University of Fribourg, Dept. of Economics

Abstract: When estimating local average and marginal treatment effects using instrumental variables (IV), multivalued endogenous treatments are frequently binarized based on a specific threshold in treatment support. However, such binarization introduces a violation of the IV exclusion if (i) the IV affects the multivalued treatment within support areas below and/or above the threshold and (ii) such IV-induced changes in the multivalued treatment affect the outcome. We discuss assumptions that satisfy the IV exclusion restriction with the binarized treatment and permit identifying the average effect of (i) the binarized treatment and (ii) unit-level increases in the original multivalued treatment among specific compliers. We derive testable implications of these assumptions and propose tests, which we apply to the estimation of the returns to (binary) college graduation instrumented by college proximity.

Keywords: Instrumental variable, LATE, binarized treatment, test, exclusion restriction, MTE.

JEL classification: C12, C21, C26.

We have benefited from comments by seminar participants at Statistics Norway and the internal research seminar of the Department of Economics of the University of Fribourg in Saas-Fee. Addresses for correspondence: Martin Eckhoff Andresen, Statistics Norway, Akersveien 26, 0177 Oslo, Norway; martin.andresen@ssb.no. Martin Huber, University of Fribourg, Bd. de Pérolles 90, 1700 Fribourg, Switzerland; martin.huber@unifr.ch.

1 Introduction

Instrumental variables (IV) strategies are frequently applied in empirical economics to overcome the potential endogeneity of a treatment variable, whose causal effect on some outcome variable is of interest to researchers and policy makers. In general (i.e. nonparametric) treatment effect models, an IV needs to satisfy a relevance condition, meaning that it monotonically shifts the treatment, as well as validity: The IV must not be associated with treatment-outcome confounders and not directly affect the outcome other than through the treatment, which is known as the IV exclusion restriction. For binary treatment variables, the IV assumptions allow identifying the local average treatment effect (LATE) on the compliers, whose treatment switches as a function of the instrument, see Imbens and Angrist (1994), or the marginal treatment effect (MTE), see Heckman and Vytlacil (2001) and Heckman and Vytlacil (2005). However, for multivalued treatments like years of schooling, one merely identifies a weighted average of per-unit treatment effects on several complier groups defined in terms of treatment-instrument reactions across the support of the treatment.

Unfortunately, the weights of the per unit treatment effects are unidentified and the complier groups might be overlapping, see Angrist and Imbens (1995). This complicates the interpretation of the effect, unless one is willing to assume homogeneous treatment effects across compliers and treatment levels. For LATE evaluation, multivalued treatments are therefore often binarized based on a specific threshold in the support of the multivalued treatment that appears interesting from a policy perspective. For instance, rather than considering years of schooling and aiming at evaluating a weighted average effect of a one year increase in schooling, one might prefer analyzing a binary indicator for college education for those whose college state reacts to the instrument. A further motivation is the raising interest in the evaluation of the MTE, which is defined as the average effect on those who are indifferent between taking and not taking a binary treatment for a specific value of the instrument. This is also reflected by the fact that the MTE framework makes use of the conditional probability to receive a binary treatment (given the instrument and possibly further control variables) for identification. Accordingly, studies estimating MTEs commonly make use of binarized versions of originally multivalued treatments. For example, Carneiro, Lokshin, and Umapathi (2017) use a distance to school instrument evaluate the effects of upper secondary schooling. For other examples of MTE analyses binarizing potentially multivalued treatments, see for instance Carneiro, Heckman, and Vytlacil (2011); Cornelissen, Dustmann, Raute, and Schönberg (2016); Felfe and Lalive (2017).

This paper demonstrates that binarizing multivalued treatments generally entails a violation of the IV exclusion restriction. Specifically, the violation occurs if (i) the IV affects the multivalued treatment within support areas below and/or above the threshold for binarization and (ii) such IV-induced changes in the multivalued treatment affect the outcome. As a methodological contribution, we show that part (i) has testable implications when the original treatment variable prior to binarization is observed. A necessary (but not sufficient) condition for ruling out 'off-threshold' compliance, i.e. IV effects on the multivalued treatment within support below or above the threshold, is a particular first stage condition. When binarizing the treatment at alternative values across its support, the first stage effect of the instrument must weakly increase up to the threshold actually chosen by the researcher, and weakly decrease thereafter. This can be tested in a moment inequality framework, see for instance Andrews and Shi (2013).

Furthermore, we consider two special cases of this first stage condition, firstly, that any compliers are situated at the threshold and secondly, that any compliers are situated at extreme values of the multivalued treatment. We show that both conditions allow identifying average per unit treatment effects for a well defined complier group (rather than an average of several complier groups) with unknown weights and that the conditions can be tested by means of standard F-tests. We apply our tests to wage market data from the National Longitudinal Survey of Young Males (NLSYM) as analysed in Card (1995). We consider an indicator for graduating from a 4 year college as binarized schooling treatment, where a dummy for proximity to college serves as instrument, and show that the exclusion restriction might be violated for the binarized treatment in some specifications.

Our paper relates to a growing literature on testing the assumptions for the nonparametric identification of the LATE and MTE with binary treatments. Balke and Pearl (1997) derive testable constraints whose violation would imply a negative density of compliers for some value of a binary outcome, even thought the lower theoretical bound of densities is zero. Heckman and Vytlacil (2005) generalize these constraints to the continuous outcome case. Kitagawa (2015) proposes a test of the constraints in a moment inequality framework based on resampling variance-weighted Kolmogorov-Smirnov-type statistics on the supremum of violations. Mourifié and Wan (2017) suggest an alternative test that allows controlling for covariates in a user-friendly way.

Huber and Mellace (2015) show that the LATE assumptions imply an alternative set of constraints related to the mean outcomes of non-compliers whose treatment does not react to the instrument. Their mean potential outcomes can be both point identified and bounded in specific treatment states. It can therefore be tested in a moment inequality framework whether the point identified means fall inside the respective bounds. Any such moment inequality-based tests check for necessary, albeit not sufficient conditions. That is, the tests are inconsistent in the sense that there may exist data generating processes which satisfy the constraints, but nevertheless violate the LATE assumptions. Sharma (2016) offers an extension by determining the likelihood that the LATE assumptions hold when the testable constraints are satisfied. Specifically, the test defines classes of valid causal models satisfying the LATE assumptions as well as as invalid models and compares their marginal likelihood in the observed data.

As an alternative strategy, Slichter (2014) suggests testing conditional IV validity by finding covariate values for which the instrument has no first stage and checking whether the instrument is associated with the dependent variable despite the absence of a first stage. Our tests differ from this and the previously mentioned approaches in that it exploits information in a multivalued treatment prior to binarization, rather than in conditional means or densities of the outcome. We therefore propose a further approach for testing IV validity in cases where the binary treatment was generated from a variable with richer support. In the presence of both a binary and a continuous instrument, Dzemski and Sarnetzki (2014) suggest a nonparametric overidentification tests for IV validity. In contrast, our approach does not require a second IV. Finally, if outcome variables are observed in periods prior to instrument assignment, placebo tests based on estimating the effect of the instrument on pre-instrument outcomes may be performed to check the plausibility of IV validity. Our tests do not rely on the availability of pre-instrument outcomes.

For the multivalued treatment case, Angrist and Imbens (1995) discuss the testable constraint that the cumulative distribution functions of the treatment in the groups receiving and not receiving the instrument, respectively, must not cross (stochastic dominance). As argued by the authors, this would point to a violation of monotonicity of the treatment in the instrument, conditional on IV validity. Fiorini and Stevens (2014) point out that testing this necessary (albeit not sufficient) condition can also have power against violations of IV validity, conditional on monotonicity. Our framework is different in that we assume that the IV relevance and validity assumptions hold for the original multivalued treatment, but not necessarily for the binarized treatment, for which we test the exclusion restriction.

This paper proceeds as follows. Section 2 introduces the econometric framework and presents assumptions related to binarized treatments. Section 3 discusses testable implications of the assumptions along with testing approaches. Section 4 presents an application to data from the NLSYM. Section 5 concludes.

2 Econometric framework and assumptions

We denote by D a multivalued treatment variable that is ordered discrete, $D \in \{0, 1, ..., J\}$ with $J+1$ being the number of possible treatment doses, or even be continuously distributed. An example is years of education. Y denotes the (discrete or continuous) outcome on which the effect ought to be estimated, for instance earnings in the labor market later in life. Under endogeneity, unobserved factors affect both D and Y such that treatment effects cannot be identified from simple comparisons of different observed levels of the treatment. One possible solution is the availability of a an instrumental variable (IV), denoted by Z, which is relevant in the sense that it influences D and valid in the sense that it is not associated with unobserved factors and does not directly affect the outcome.

For the formal discussion of the identifying assumptions and testable implications, we use the potential outcome framework, see for instance Rubin (1974). Denote by D_z the potential treatment state that would occur if the instrument Z was exogenously set to some value z, and by Y_d the potential outcome with the treatment exogenously set to some value d in the support of D. We will henceforth assume a binary instrument ($Z \in \{1,0\}$), which simplifies the exposition. but discuss a straightforward extension to a continuous or multivalued instrument at the end of section 3.

The starting point for our analysis is the standard IV assumptions, which will be maintained throughout the paper:

Assumption 1 (IV validity and relevance):

(a) $Z \perp (D_1, D_0, Y_0, Y_1, ..., Y_J)$ (IV independence),

(b) $\Pr(D_1 \geq D_0) = 1$ and $\Pr(D_1 > D_0) > 0$ (positive monotonicity).

where "$\perp$" denotes independence. Assumption 1(a) implies two conditions. First, the instrument must be random so that it is unrelated to factors affecting the treatment and/or outcome. Therefore, not only the potential outcomes/treatment states, but also the types, which are defined by the joint potential treatment states, are independent of the instrument. Second, Z must not have a direct effect on Y other than through D, i.e., satisfy an exclusion restriction, which can be seen from the fact that the potential outcomes are only defined in terms of d rather than z and d.[1] The first part of Assumption 1(b) implies that the treatment of any individual does not decrease in the instrument. The second part assumes the existence of individuals whose treatment state positively reacts to the treatment. Both parts together imply a positive first stage effect of the instrument on the treatment: $E(D|Z=1) - E(D|Z=0) > 0$. We note that Assumption 1(b)

[1] To make these two aspects explicit, Assumption 1(a) may be postulated as two conditions, see Angrist, Imbens, and Rubin (1996): (i) $Z \perp (D_1, D_0, Y_{1,0}, Y_{0,0}, Y_{1,1}, Y_{0,1}, ..., Y_{1,J}, Y_{0,J})$ and (ii) $Y_{1,d} = Y_{0,d} = Y_d$ for all d in the support of D (exclusion restriction), where $Y_{z,d}$ denotes a potential outcome defined in terms of both the instrument z and the treatment d.

could be replaced by negative monotonicity: $\Pr(D_1 \leq D_0) = 1$ and $\Pr(D_1 < D_0) > 0$. From an econometric perspective, both versions are equivalent, because when redefining the instrument under negative monotonicity to be $1 - Z$, Assumption 1(b) is satisfied.

If D was binary, the local average treatment effect (LATE) on the so-called compliers, who switch treatment from 0 to 1 as a response to a switch in the instrument from 0 to 1, could be identified by the probability limit of two stage least squares (TSLS) or the Wald estimator, see Angrist and Imbens (1995). That is, under Assumption 1 and $D \in \{0,1\}$, $E[Y_1 - Y_0 | D_1 - D_0 = 1] = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(D|Z=1) - E(D|Z=0)}$. For a multivalued treatment, however, the causal effect for a single complier population defined by specific potential treatment states, e.g. for those increasing treatment from 2 to 3 when the instrument is switched from 0 to 1, is not identified. Angrist and Imbens (1995) show for ordered discrete treatments that it is merely possible to identify a weighted average of causal effects of unit increases in the treatment, $Y_j - Y_{j-1}$, $j \in \{1, ..., J\}$. Specifically, the authors show in the proof of their Theorem 1 that under Assumption 1,

$$E(Y|Z=1) - E(Y|Z=0) \quad = \quad \sum_{j=1}^{J} E[(Y_j - Y_{j-1}) | D_1 \geq j > D_0] \cdot \Pr(D_1 \geq j > D_0), \qquad (1)$$

and that

$$E(D|Z=1) - E(D|Z=0) = \sum_{j=1}^{J} \Pr(D_1 \geq j > D_0). \qquad (2)$$

It follows that

$$\frac{E(Y|Z=1) - E(Y|Z=0)}{E(D|Z=1) - E(D|Z=0)} = \sum_{j=1}^{J} w_j \cdot E(Y_j - Y_{j-1} | D_1 \geq j > D_0) = \Delta^w, \qquad (3)$$

where the weights are given by

$$w_j = \frac{\Pr(D_1 \geq j > D_0)}{\sum_{j=1}^{J} \Pr(D_1 \geq j > D_0)}. \qquad (4)$$

Note that $0 \leq w_j \leq 1$ and $\sum_{j=1}^{J} w_j = 1$. Therefore, the probability limits of TSLS or the

Wald estimator equal a weighted average of per-unit treatment effects on various complier groups defined by different margins of the potential treatments. The weights of the per unit treatment effects, however, remain unidentified. Furthermore, the complier groups might be overlapping. Some individuals could, for instance, satisfy both $(D_1 \geq j > D_0)$ and $(D_1 \geq j+1 > D_0)$ for some j and therefore be accounted multiple times. This arguably compromises the interpretability of the effect.

Many empiricists apparently circumvent such issues of interpretability by binarizing a multivalued treatment. Examples include the effect of a binary indicator for college attendance, instrumented for instance by college proximity (Kane and Rouse, 1993; Carneiro, Heckman, and Vytlacil, 2011), or the impact of fertility measured by a dummy for having three or more children, instrumented by same-sex sibship or twin births (Angrist and Evans, 1998; Mogstad and Wiswall, 2016; Black, Devereux, and Salvanes, 2005)). Binarization is also common in the literature on the MTE, a parameter that can be regarded as the limit of the LATE for an infinitesimal change in the instrument. See Carneiro, Lokshin, and Umapathi (2017); Carneiro, Heckman, and Vytlacil (2011); Cornelissen, Dustmann, Raute, and Schönberg (2016); Felfe and Lalive (2017) for examples in the context of returns to college and child care, respectively.

Let $D_z^* = I\{D_z \geq j^*\}$ denote the potential state of the binarized treatment under $z \in \{0,1\}$, where $I\{a\}$ is the indicator function that is equal to one when a holds and zero otherwise, while j^* denotes a specific threshold value in the support of D. This allows defining the average effect among those whose treatment state passes the threshold when switching the instrument from 0 to 1:

$$\begin{aligned} \Delta^* &= E[Y_{D_1} - Y_{D_0} | D_1^* - D_0^* = 1] = E[Y_{D_1} - Y_{D_0} | D_1 \geq j^* > D_0] \qquad (5) \\ &= \sum_{j=1}^{J} E[Y_j - Y_{j-1} | D_1 \geq j > D_0, D_1 \geq j^* > D_0] \cdot \Pr(D_1 \geq j > D_0 | D_1 \geq j^* > D_0), \end{aligned}$$

The expression following the second equality in (5) shows that Δ^* is a weighted average of effects across compliers satisfying $D_1^* - D_0^* = 1$, even though they could be defined by different potential (original) treatments D_1, D_0. That is, the effect refers to all compliers satisfying $D_1 \geq j^* > D_0$, no matter how heterogeneous they are in terms of D_1 and D_0, which is important to bear in mind for interpretation. Δ^* generally differs from Δ^w identified in (3): the latter identifies an average effect of unit-level changes while the former corresponds to a total effect, i.e. the sum of effects of unit-level changes that are weighted with the probability that they occur among compliers switching the binarized treatment as a response to the instrument.

As a matter of fact frequently disregarded by empiricists, Δ^* is generally not identified by the probability limit of the Wald estimator or TSLS based on D^* rather than D,

$$\mathrm{W}^{D^*} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(D^*|Z=1) - E(D^*|Z=0)}. \tag{6}$$

This is the case despite of the supposed analogy of (6) to the results of Angrist and Imbens (1995) for a (truly) binary treatment. However, a binarization of the treatment variable generally entails a violation of the exclusion restriction such that Assumption 1a for D does not carry over to D^*.

To see this, rewrite (1) using the law of total probability and Assumption 1(b) as

$$\begin{aligned}
& E(Y|Z=1) - E(Y|Z=0) \\
= & \sum_{j=1}^{J} E[Y_j - Y_{j-1}|D_1 \geq j > D_0, D_1 \geq j^* > D_0] \cdot \Pr(D_1 \geq j > D_0, D_1 \geq j^* > D_0) \\
+ & \sum_{j=1}^{J} E[Y_j - Y_{j-1}|D_1 \geq j > D_0, D_1 > D_0 \geq j^*] \cdot \Pr(D_1 \geq j > D_0, D_1 > D_0 \geq j^*) \\
+ & \sum_{j=1}^{J} E[Y_j - Y_{j-1}|D_1 \geq j > D_0, j^* > D_1 > D_0] \cdot \Pr(D_1 \geq j > D_0, j^* > D_1 > D_0) \\
= & \sum_{j=1}^{J} E[Y_j - Y_{j-1}|D_1 \geq j > D_0, D_1 \geq j^* > D_0] \cdot \Pr(D_1 \geq j > D_0, D_1 \geq j^* > D_0) \qquad (7) \\
+ & \sum_{j=1}^{J} E[Y_j - Y_{j-1}|D_1 \geq j > D_0, I\{D_1 \geq j^* > D_0\} = 0] \cdot \Pr(D_1 \geq j > D_0, I\{D_1 \geq j^* > D_0\} = 0).
\end{aligned}$$

By summing over j, (7) simplifies to

$$
\begin{aligned}
&= E[Y_{D_1} - Y_{D_0}|D_1 \geq j^* > D_0] \cdot \Pr(D_1 \geq j^* > D_0) \\
&+ E[Y_{D_1} - Y_{D_0}|D_1 > D_0, I\{D_1 \geq j^* > D_0\} = 0] \cdot \Pr(D_1 > D_0, I\{D_1 \geq j^* > D_0\} = 0). \quad (8)
\end{aligned}
$$

Note that the condition $(D_1 > D_0, I\{D_1 \geq j^* > D_0\} = 0)$ captures complier groups whose treatment reacts to the instrument $(D_1 > D_0)$, but in a way that it does not cross the threshold j^* $(I\{D_1 \geq j^* > D_0\} = 0)$. Furthermore, consider the denominator of (6):

$$
\begin{aligned}
& E(D^*|Z = 1) - E(D^*|Z = 0) \\
&= \Pr(D \geq j^*|Z = 1) - \Pr(D \geq j^*|Z = 0) = \Pr(D_1 \geq j^*) - \Pr(D_0 \geq j^*) \\
&= \Pr(D_1 \geq j^* > D_0) + \Pr(D_0 \geq j^*) - \Pr(D_0 \geq j^*) \\
&= \Pr(D_1 \geq j^* > D_0). \quad (9)
\end{aligned}
$$

where the second equation follows from Assumption 1(a) and the third from 1(b). Division of (8) by (9) reveals that W^{D^*} does generally not identify Δ^* due to the second line in (8). The latter corresponds to the contribution of compliers whose treatment is not induce to cross j^* by the instrument. For this reason, the LATE of interest is only obtained in the special cases that either such off-threshold compliers do not exist or that their average treatment effect is zero, as formalized in Assumptions 2 and 3.

Assumption 2 (zero average treatment effect among non-captured compliers):
$E[Y_{D_1} - Y_{D_0}|D_1 > D_0, I\{D_1 \geq j^* > D_0\} = 0] = 0.$

Assumption 3 (full capturing of compliers by threshold):
$\Pr(D_1 > D_0 \geq j^*) = \Pr(j^* > D_1 > D_0) = 0.$

Assumption 2 postulates the absence of an average causal effect for the compliers not captured by the threshold. That is, given a first stage not 'going through' j^*, the average second stage for these compliers must be zero. On the other hand, Assumption 3, which can be alternatively formalized as $\Pr(I\{D_1 \geq j^* > D_0\} = 0|D_1 > D_0) = 0$, implies that all compliers are captured by the threshold in the sense that their treatment state is shifted from some $D_0 < j^*$ to some $D_1 \geq j^*$ by the instrument. Thus, there exist no complier groups whose treatment is affected by instrument in a way that D_0, D_1 are either both below or both above the threshold. This rules out first stages not 'going through' the threshold j^*. Therefore, the IV exclusion restriction fails with binarized treatments if (i) there exist compliers not captured by the definition of D^* and (ii) the instrument-induced changes in treatment actually affects the outcome of these subjects.

In contrast, if either Assumption 2 or 3 hold,

$$E(Y|Z=1) - E(Y|Z=0) = E[Y_{D_1} - Y_{D_0}|D_1 \geq j^* > D_0] \cdot \Pr(D_1 \geq j^* > D_0), \tag{10}$$

such that $\mathrm{W}^{D^*} = \Delta^*$. Considering the expression after the first equality in (7) reveals that identification is also obtained by combinations of Assumptions 2 and 3 for different subsets of compliers not captured by D^*. For instance, Assumption 3 could hold below the threshold, securing no compliers in this region, while Assumption 2 could hold above the threshold, securing no treatment effect among these compliers. If neither Assumption 2 and 3 hold, it follows from (8) that the direction of the bias in W^{D^*} is determined by the direction of the average treatment effect among off-threshold compliers. Unfortunately, imposing the popular Monotone Treatment Response (MTR) assumption of Manski and Pepper (2000), which implies that the treatment effect goes into the same direction for both threshold and off-threshold compliers, does not permit bounding the absolute size of Δ^*. Quite on the contrary, MTR implies that W^{D^*} overstates (understates) Δ^* whenever it is positive (negative).

We subsequently discuss two special cases of Assumption 3 for the reason that they allow identifying Δ^w, the weighted average of per-unit treatment effects, based on W^{D^*}. To this end, we rewrite (1) as

$$\begin{aligned}
E(Y|Z=1)-E(Y|Z=0) &= \sum_{j=1}^{j^*-1} E[Y_j - Y_{j-1}|D_1 \geq j > D_0]\cdot \Pr(D_1 \geq j > D_0) \\
&+ E[Y_{j^*} - Y_{j^*-1}|D_1 \geq j^* > D_0]\cdot \Pr(D_1 \geq j^* > D_0) \\
&+ \sum_{j=j^*+1}^{J} E[Y_j - Y_{j-1}|D_1 \geq j > D_0]\cdot \Pr(D_1 \geq j > D_0).
\end{aligned} \tag{11}$$

and note that the probabilities in the first and third term are 0 by Assumption 3. The first special case occurs if and only if all compliers are concentrated at the threshold such that the instrument has no effect on the treatment at margins of D other than j^*, see also the discussion in Section 3.1 of Angrist and Imbens (1995).

Assumption 4 (concentration of compliers at threshold):
$\sum_{j\neq j^*}^{J} \Pr(D_1 \geq j > D_0) = 0$.

It follows from Assumption 4 that (9) and (2) are equivalent, implying $E(D|Z=1)-E(D|Z=0)=E(D^*|Z=1)-E(D^*|Z=0)$ and that $E(Y|Z=1)-E(Y|Z=0)=\Delta^*\cdot\Pr(D_1 \geq j^* > D_0)$ in (11). It follows that $\mathrm{W}^{D^*}=\Delta^*=\Delta^w$. In cases where Assumption 4 is violated, Angrist and Imbens (1995) show that W^{D^*} is larger in absolute terms than Δ^w.

As second special case, assume that all compliers in the population switch their treatment from the lowest ($D_0=0$) to the highest ($D_1=J$) possible treatment value as response to the instrument, while there exist no compliers with other treatment margins affected. This implies that the complier population remains constant across values j.

Assumption 5 (concentration of compliers at extreme treatment values):
$I\{D_1 \geq j > D_0\} = I\{D_1 \geq j^* > D_0\}$ for all $j, j^* \in \{1,...,J\}$.

Note that this assumption is stated in terms of indicator functions in contrast to Assumption 4, which is stated in terms of compliance probabilities. The reason is that while constant complier sets across j imply constant compliance probabilities, the converse is not true: There might for example exist compliers that shift D from 0 to 1 and others that shift from 1 to 2 when switching the instrument from 0 to 1. If the shares of these complier groups are the same, the complier probabilities would remain constant across $j \in \{1, 2\}$, despite the existence of compliers at intermediate treatment values.

Under Assumption 5, (11) simplifies to

$$\left\{\sum_{j=1}^{J} E[(Y_j - Y_{j-1})|D_1 \geq j^* > D_0]\right\} \cdot \Pr(D_1 \geq j^* > D_0). \tag{12}$$

Therefore, $\mathrm{W}^{D^*} = \Delta^*$ and corresponds to the sum of impacts related to unit-level changes in treatment D across the entire support of the latter. This implies $\Delta^w = \Delta^*/J$, i.e. the average per unit effect corresponds to the sum of effects across all possible unit-level changes divided by the number of possible treatment states J. The reason is that under Assumption 5, the weights in (4) become $\frac{\Pr(D_1 \geq j^* > D_0)}{J \cdot \Pr(D_1 \geq j^* > D_0)} = 1/J$, while in (3), $E(Y_j - Y_{j-1}|D_1 \geq j > D_0) = E(Y_j - Y_{j-1}|D_1 \geq j^* > D_0)$. Therefore,

$$\Delta^w = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(D|Z=1) - E(D|Z=0)} = \frac{E(Y|Z=1) - E(Y|Z=0)}{E(D^*|Z=1) - E(D^*|Z=0)} \Big/ J = \frac{\Delta^*}{J}. \tag{13}$$

3 Testing Assumptions 3, 4, and 5

This section introduces tests for necessary conditions of Assumptions 3, 4, and 5. Under the satisfaction of Assumption 3, it must hold that the share of compliers whose treatment is induced to pass j by the instrument weakly increases when gradually increasing j up to j^*, while weakly decreasing thereafter. The reason is that Assumption 3 requires that j^* captures all compliers, implying that the first stage is maximized at the threshold. Formally, the following weak moment inequality constraints need to hold:

$$\begin{aligned} &\Pr(D_1 \geq j' > D_0) \geq \Pr(D_1 \geq j'' > D_0) \text{ for all } j^* \geq j' > j'' > 0, \\ &\Pr(D_1 \geq j' > D_0) \leq \Pr(D_1 \geq j'' > D_0) \text{ for all } J \geq j' > j'' \geq j^*. \end{aligned} \tag{14}$$

Proof. Consider the first line of (14) and note that

$$\begin{aligned}\Pr(D_1 \geq j' > D_0) &= \Pr(D_1 \geq j' > j'' > D_0) + \Pr(D_1 \geq j' > D_0 \geq j'') \\ &= \Pr(D_1 \geq j'' > D_0) + \Pr(D_1 \geq j' > D_0 \geq j'') \end{aligned} \quad (15)$$

The first equality follows from the law of total probability and the second from Assumption 3. To see this, note that $\Pr(D_1 \geq j'' > D_0) = \Pr(D_1 \geq j' > j'' > D_0) + \Pr(j' > D_1 \geq j'' > D_0)$. However, by Assumption 3, $\Pr(j' > D_1 \geq j'' > D_0) = 0$ for any $j' \leq j^*$, such that $\Pr(D_1 \geq j'' > D_0) = \Pr(D_1 \geq j' > j'' > D_0)$. Therefore, it follows from $\Pr(D_1 \geq j' > D_0 \geq j'') \geq 0$ that $\Pr(D_1 \geq j' > D_0) \geq \Pr(D_1 \geq j'' > D_0)$. The proof of the second line of (14) is analogous and is therefore omitted. ■

By Assumption 1(a) and (b), (14) implies (in analogy to the discussion in (9) for $\Pr(D_1 \geq j^* > D_0)$) that

$$\begin{aligned}&\beta_j \geq \beta_{j'} \text{ for all } j^* \geq j > j' > 0, \\ &\beta_j \leq \beta_{j'} \text{ for all } J \geq j > j' \geq j^*, \end{aligned} \quad (16)$$

where $\beta_j = \Pr(D \geq j|Z = 1) - \Pr(D \geq j|Z = 0)$ denotes the first stage effect of Z on the probability that D is larger or equal to some value j. This allows formulating the following null hypothesis for testing Assumption 3, conditional on the satisfaction of Assumption 1:

$$H_0 : \begin{array}{l} \beta_{j+1} - \beta_j \geq 0, \text{ for all } j^* \geq j > 0, \\ \beta_j - \beta_{j+1} \geq 0, \text{ for all } J > j \geq j^* \end{array}$$

It is important to see that the satisfaction of the null hypothesis in (17) is necessary, albeit not sufficient for Assumption 3. One can easily construct cases in which the weak inequalities hold, even though a subset of individuals complies off threshold. Concerning the practical implementation, it suffices to implement the test for adjacent β_j parameters because of their nested nature: $\beta_2 \geq \beta_0$ provide no additional restrictions on the data when $\beta_2 \geq \beta_1$ and $\beta_1 \geq \beta_0$. These conditions can be verified using testing procedures for moment inequality constraints, see for instance Andrews and Shi (2013).

An implementation is available in the 'cmi_test' command for the statistical software 'Stata', see Andrews, Kim, and Shi (2017), which we use in our application presented in Section 4. We to this end reconsider the first line of (17) and note that

$$\begin{aligned}
\beta_{j+1} - \beta_j = & \quad \Pr(D \geq j+1 \mid Z=1) - \Pr(D \geq j+1 \mid Z=0) \\
- & \quad \Pr(D \geq j \mid Z=1) - \Pr(D \geq j \mid > Z=0) \\
= & \quad \Pr(D = j \mid Z=0) - \Pr(D = j \mid Z=1)
\end{aligned} \tag{17}$$

A symmetric argument follows for the second line. Therefore, (17) can be rewritten in the following way based on inverse probability weighting by $E(Z)$ and $1 - E(Z)$:

$$\begin{aligned}
E(m_j(D,Z) \mid X) & \geq 0 \\
\text{where } m_j(D,Z) & = I\{D=j\}\tfrac{E(Z)-Z}{(1-E(Z))E(Z)} \qquad \text{for } j^* > j \geq 0 \\
\text{and } m_j(D,Z) & = I\{D=j\}\tfrac{Z-E(Z}{(1-E(Z))E(Z)} \qquad \text{for } J > j \geq j^*.
\end{aligned} \tag{18}$$

These constraints match the structure of the 'cmi_test' command of Andrews, Kim, and Shi (2017), which verifies the sample analog of (18). Testing may be implemented both based on Cramer-von-Mises and Kolmogorov-Smirnov-type statistics on average or maximum violations across j, respectively, and both is considered in our empirical application. Control variables can be included in this testing approach simply by replacing all instances of $E(Z)$ with $E(Z|X)$, see example 6 in Andrews and Shi (2014).

Concerning Assumption 4, both a necessary and sufficient condition for its satisfaction is that any first stage effect of Z on on the probability that $D \geq j$ must be zero unless $j = j^*$, because all compliers must be located at the threshold. Formally,

$$H_0 : \beta_j = 0 \quad \text{for all } j \neq j^*. \tag{19}$$

Finally, a necessary condition for Assumption 5 is that the first stages or complier probabilities remain constant across j. As highlighted in the discussion of Assumption 5 in Section 2, this is implied by an concentration of compliers at extreme treatment values, albeit not sufficient for ruling out other complier groups. Formally, the hypothesis to be tested is

$$H_0 : \beta_j = \beta_{j+1} \quad \text{for all } j < J. \tag{20}$$

Both (19) and (20) can be tested by means of an F-test in a system of equations in which treatment indicator functions $I\{D \geq j\}$ at different values j are regressed on a constant and Z.

We note that the testing approaches can be extended to multivalued discrete as well as continuous instruments. For multivalued discrete instruments, the conditions given in (17), (19), and (20) must hold when defining $\beta_j = \Pr(D \geq j|Z = z') - \Pr(D \geq j|Z = z'')$ for any values $z' > z''$ in the support of Z. For continuous instruments, the conditions given in (17), (19), and (20) must hold for infinitesimal increases in Z across the entire support of Z. In this case $\beta_j = \frac{\partial \Pr(D \geq j|Z=z)}{\partial z}$ for any z in the support of Z. Finally, if the IV assumptions are not assumed to hold unconditionally but given control variables, the latter might be included as conditioning set in the moment inequality- and regression-based tests. In (18), for instance, control variables can be considered by replacing $E(Z)$ everywhere with the conditional expectation of Z given the controls, also known as instrument propensity score.

4 Empirical application

We apply our method to labor market data previously analysed by Card (1995) that comprise 3,010 observations and come from the 1966 and 1976 waves of the U.S. National Longitudinal Survey of Young Men (NLSYM). Card (1995) considers a dummy for proximity to a four-year college in 1966 as an instrument for the likely endogenous schooling decision to estimate returns to schooling in 1976. The intuition is that proximity should affect the schooling decision of some individuals, for instance, due to costs associated with going to college when not living at home.

The original data contain years of schooling as measure of education, but similar to Carneiro, Heckman, and Vytlacil (2011), we binarize the treatment to represent having at least 16 years of education, roughly comparable to a four-year college degree.

Table 1: Summary statistics,

Variable	N	mean	s.d.	min	max	comment
Years of schooling	3,010	13.3	2.68	1	18	1976
College dummy	3,010	0.27	0.44	0	1	Dummy for 16 or more years of education
College proximity	3,010	0.68	0.47	0	1	= 1 if near 4-year college in 1966
Age	3,010	28.1	3.14	24	34	
Father's educ	2,320	10.0	3.72	0	18	
Mothers' educ	2,657	10.3	3.18	0	18	
Region	3,010	4.64	2.27	1	9	Regional dummy, 1966
SMSA	3,010	0.71	0.45	0	1	Metropolitan area of residence dummy
Black	3,010	0.23	0.42	0	1	
Family type	2,796	1.07	0.38	0	2	Single mom / both parents / step-parent
IQ	2,061	102.4	15.4	50	149	

Note: Data source: National Longitudinal Study of Young Men, 1966 and 1976 waves.

The variables used in our analysis are summarized in Table 1. The multivalued treatment is years of schooling in 1976, which varies from 1 to 18 years with a mean of 13.3. Our binarized treatment is a dummy for having 16 or more years of schooling, which has a mean of 0.27. The instrument is a dummy equal to 1 for people living close to a 4-year college in 1966, which around 68 % of the sample did. In addition, we report a range of control variables, including age, parents' education, geographic dummies, race, a dummy for family type at age 14, and IQ score.

To illustrate our tests, we first estimate the β_j parameters outlined in Section 3, which reflect increases in the probability of having j or more years of schooling when living close to a four-year college compared to living further away, for all margins of education in the Card data. To this end we estimate a system of equations in which the indicators of having at least j years of education at various j are regressed on a constant and the instrument. Figure 1 displays the β_j estimates along with pointwise 95% confidence intervals. The results in the top graph are without controls, while in the remainder graphs we control for the variable(s) indicated above the respective graph in a fully saturated way, i.e. nonparametrically.[2] The reason for this is that proximity to college

[2]We drop observations in singleton groups (observations with unique combinations of control variables) when

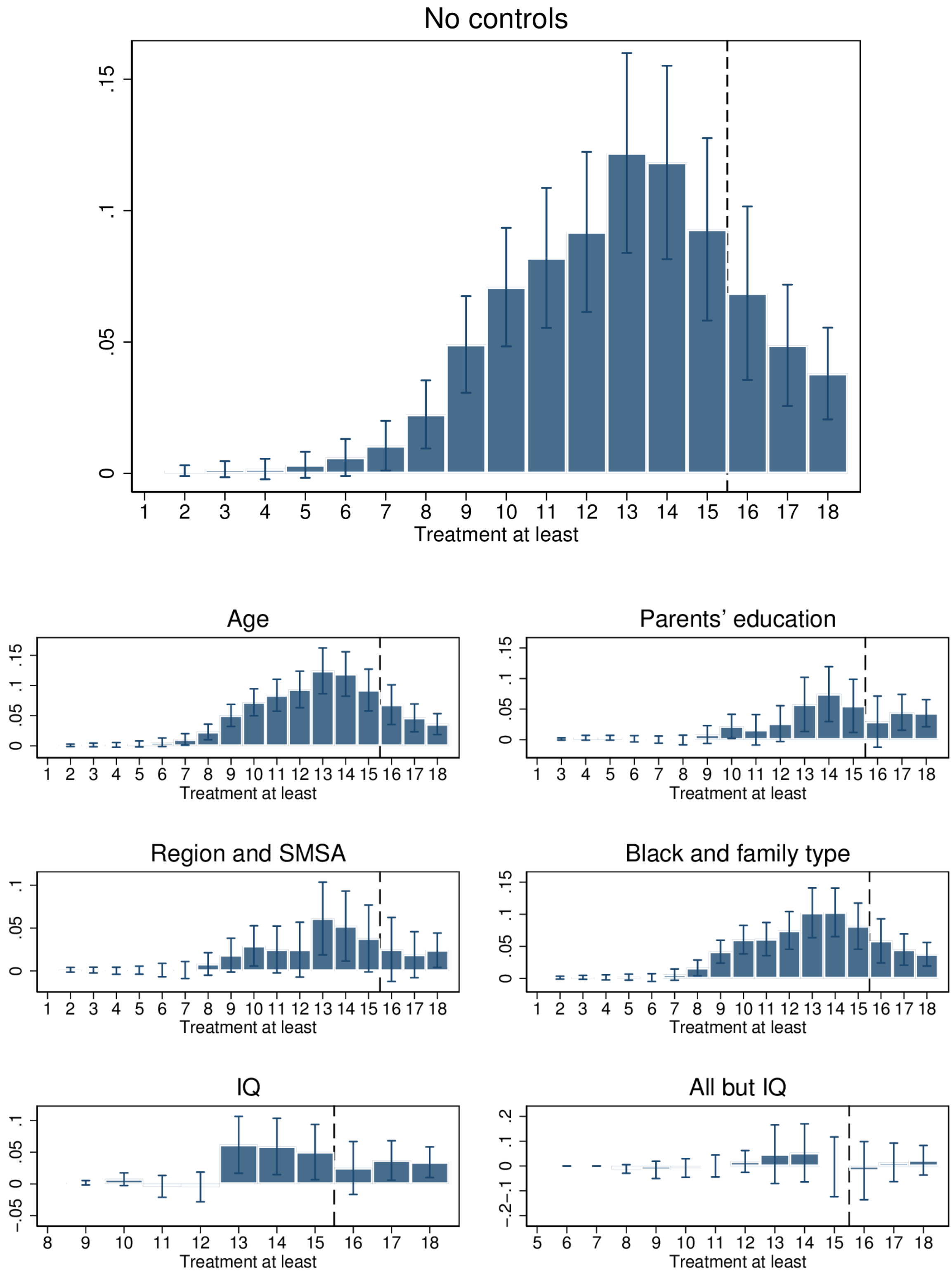

Figure 1: Effects of living close to a four-year college on years of education

Note: Data from NLSYM. Figure shows the estimated impact on binary measures of years of education above j of living close to a four-year college. The threshold for the binarized treatment is 16 or more years of education as indicated by a dashed line, corresponding roughly to a four-year college degree. Fully flexible controls as indicated above each panel.

is likely associated with factors also affecting wages, like local labor market conditions or family background, which would violate Assumption 1. As testing Assumptions 3, 4, and 5 is conditional on Assumption 1, we similarly to Card (1995) control for regional variables (SMSA and region in the US) and socio-economic factors (e.g. parents' education and ethnicity) to increase plausibility of IV exogeneity.

Inspecting the graphs of Figure 1 allows eye-balling the plausibility of our assumptions. We observe that the pattern of coefficients in the various scenarios is not consistent with Assumption 4, which requires all coefficients except β_{j^*} to be 0. Neither does it appear to support Assumption 5, which requires the coefficients to be constant across j. Concerning Assumption 3, notice that the dashed line indicating 16 years of education is consistently to the right of (rather than at) the mode of the β_j estimates, pointing to a violation of the conditions in (14).

To formally investigate Assumption 3, we test the constraints in (18) by the 'cmi_test' command of Andrews, Kim, and Shi (2017) based on Cramer-von-Mises and Kolmogorov-Smirnov statistics.[3] The results are provided in panel A of Table 2. Without including control variables, the p-value of either statistic is 0.062, pointing to a marginally significant violation of the constraint. When including various control sets, p-values increase, but remain marginally significant for several specifications. Also note the low and insignificant first stage coefficient in the more robust specifications, challenging the relevance of the instrument in these specification even when we cannot reject Assumption 3. For testing Assumptions 4 and 5, we test the null hypotheses in (19) and (20) using F-tests in our system of equations used to estimate the β_j parameters.[4] The outcomes are displayed in panels B and C of Table 2, respectively. Both assumptions are rejected at the 5% level in most specifications, the exceptions being the specification controlling flexibly

estimating the β_j coefficients (for a discussion, see e.g. Correia (2015)). In order to keep a consistent sample, we also drop singleton observations for the conditional moment inequality test discussed further below, leading to a drop in observations as we include more control variables in Table 2. Observations in singleton groups would in any case not contribute to the estimation of β_j or to the rejection of the restrictions in (18), because the empirical moment conditions for these observations for instance in (18) are 0.

[3] A small program for Stata, available upon request, estimates and plots the β_j coefficients, tests Assumption 4 and 5 using F-tests and constructs the moment inequalities and tests them using cmi_test.

[4] The system of equations is estimated in a stacked regression using the reghdfe command (Correia, 2014) to account for the covariance of the β_j estimates. Standard errors are clustered at the individual level and robust to heteroskedasticity.

Table 2: Tests of instrument validity with a binarized instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A: Conditional moment inequalities tests of Assumption 3							
Cramer-von-Mises type test statistic							
Test statistic	1.618	3.028	2.764	2.783	1.801	2.686	0.178
Critical value 1%	2.260	5.151	3.530	3.615	3.157	3.840	0.313
Critical value 5%	1.697	4.358	3.006	3.092	2.681	3.321	0.272
Critical value 10%	1.437	4.059	2.765	2.834	2.404	3.028	0.249
p-value	0.062	0.501	0.101	0.115	0.363	0.206	0.514
Kolmogorov-Smirnov type tests statistic							
Test statistic	13.570	10.725	11.187	14.983	7.800	10.984	3.980
Critical value 1%	18.953	21.858	18.626	19.238	18.693	17.005	11.409
Critical value 5%	14.233	17.651	14.955	15.733	14.360	13.958	8.792
Critical value 10%	12.050	15.593	13.136	14.140	12.392	12.302	7.689
p-value	0.062	0.492	0.202	0.072	0.467	0.181	0.656
B: F-test of Assumption 4							
F	4.532	4.654	2.303	1.208	3.711	2.062	0.763
p-value	0.000	0.000	0.004	0.253	0.000	0.030	0.664
C: F-test of Assumption 5							
F	4.639	4.755	2.449	1.173	3.794	2.301	0.723
p-value	0.000	0.000	0.002	0.282	0.000	0.014	0.717
Controls							
Age		✓					✓
Fathers' education			✓				✓
Mothers' education			✓				✓
Region				✓			✓
SMSA				✓			✓
Black					✓		✓
Family type					✓		✓
IQ						✓	
N	3,010	3,010	2,179	3,010	2,796	2,048	638

Note: Panel A shows the results from an F-test of $\beta_j = 0$ for all $j \neq j^*$, testing the special case in Assumption 4. Panel B shows tests of whether all β_j are the same, testing the special case in Assumption 5. Panel C shows shows test statistics, critical values and resulting p-values from test of the moment inequalities in (15). All tests performed with `cmi_test` for Stata (Andrews, Kim, and Shi, 2017). Controls as indicated in the bottom panel. Singleton groups are dropped.

for region and SMSA and the very parsimonious specification with all controls but IQ, which has little power. The results therefore suggest that compliers do not exclusively choose schooling levels situated at the threshold, i.e. 16 and 15 years with and without instrument, nor exclusively at the highest and lowest possible values of education. Therefore, the weighted average of per-unit treatment effects, Δ^w, cannot be recovered based on the binarized treatment.

Overall, our results point to the possibility that the exclusion restriction might be violated for the binarized education measure considered, depending on control set. Even though the graphs and estimates suggest that proximity to a four-year college indeed affects education, it may do so not by an exclusive shift towards obtaining at least a four-year college degree. Rather, the instrument seems to also affect the probability of both starting without finishing college and of obtaining a two-year college degree. However, such possibilities are ignored when defining the treatment as a four-year college degree. Judging from the graphs in Figure 1, the exclusion restriction is more likely satisfied if treatment is defined as having at least some college education versus having less education. In fact, constraints (17) and (18) cannot be rejected if the threshold is chosen at the mode of an unimodal set of β_j parameters. Yet, we need to bear in mind that even in this case Assumption 3 might be violated, namely if some compliers shift from a two-year college degree to a four-year degree. After all, we can only test necessary, albeit not sufficient conditions for the exclusion restriction after binarization.

5 Conclusion

In the context of IV-based estimation, we discuss threats to the exclusion restriction when binarizing a multivalued endogenous treatment. Such a violation occurs whenever (i) the IV affects the multivalued treatment within support areas below and/or above the threshold for binarization and (ii) such IV-induced changes in the multivalued treatment affect the outcome. As a consequence, IV estimation with a binarized treatment identifies the causal effect among individuals whose binary treatment complies with the IV only if either (i) or (ii) can be ruled out. We show that (i) has testable implications that can be tested in a moment inequality framework when the original treatment variable prior to binarization is observed. Furthermore, when ruling out (i) and restricting the support of the multivalued treatment in a particular way, not only the average complier effect of the binarized treatment, but also a weighted average of per-unit treatment effects of the multivalued treatment is recovered. We derived testable implications of these support restrictions that can be verified by standard F-tests. Finally, we provided an empirical illustration to the estimation of returns to a four year college degree, a binarized treatment generated from the multivalued measure of years of education. Our results suggested that such a coarse definition of education may violate the exclusion restriction.

As a final word of caution, we emphasize that the threats to the exclusion restriction not only arise when binarizing a treatment. The issues discussed in this paper prevail whenever the IV affects a finer measure of treatment than used by the researcher in her IV analysis, even when finer treatment measures are not available in the data. Examples include binning a truly continuous treatment into a discrete number of categories or coarsening ordered discrete treatments into a smaller number of categories (e.g. considering low vs. intermediate vs. high levels of education rather than years of schooling). The conditions in this paper highlight under which circumstances the IV validity for the underlying finer treatment measure carries over to a more coarsely defined treatment.

References

Andrews, D., and X. Shi (2013): "Inference Based on Conditional Moment Inequalities," *Econometrica*, 81(2), 609–666.

Andrews, D. W., and X. Shi (2014): "Nonparametric inference based on conditional moment inequalities," *Journal of Econometrics*, 179(1), 31 – 45.

Andrews, D. W. K., W. Kim, and X. Shi (2017): "Commands for testing conditional moment inequalities and equalities," *Stata Journal*, 17(1), 56–72.

Angrist, J., and G. W. Imbens (1995): "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity," *Journal of American Statistical Association*, 90, 431–442.

Angrist, J., G. W. Imbens, and D. Rubin (1996): "Identification of Causal Effects using Instrumental Variables," *Journal of American Statistical Association*, 91, 444–472 (with discussion).

Angrist, J. D., and W. N. Evans (1998): "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *The American Economic Review*, 88(3), 450–477.

Balke, A., and J. Pearl (1997): "Bounds on Treatment Effects From Studies With Imperfect Compliance," *Journal of the American Statistical Association*, 92, 1171–1176.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005): "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education," *The Quarterly Journal of Economics*, 120(2), 669–700.

Card, D. (1995): "Using Geographic Variation in College Proximity to Estimate the Return to Schooling," in *Aspects of Labor Market Behaviour: Essays in Honour of John Vanderkamp*, ed. by L. Christofides, E. Grant, and R. Swidinsky, pp. 201–222. University of Toronto Press, Toronto.

Carneiro, P., J. J. Heckman, and E. J. Vytlacil (2011): "Estimating Marginal Returns to Education," *American Economic Review*, 101(6), 2754–81.

Carneiro, P., M. Lokshin, and N. Umapathi (2017): "Average and Marginal Returns to Upper Secondary Schooling in Indonesia," *Journal of Applied Econometrics*, 32(1), 16–36.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2016): "Who benefits from universal child care? Estimating marginal returns to early child care attendance," Forthcoming in Journal of Political Economy.

Correia, S. (2014): "REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects," Statistical Software Components, Boston College Department of Economics.

Correia, S. (2015): "Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix," Unpublished.

Dzemski, A., and F. Sarnetzki (2014): "Overidentification test in a nonparametric treatment model with unobserved heterogeneity," *mimeo, University of Mannheim.*

Felfe, C., and R. Lalive (2017): "Does Early Child Care Affect Childrens Development?," Working paper.

Fiorini, M., and K. Stevens (2014): "Monotonicity in IV and fuzzy RD designs - A Guide to Practice," *mimeo, University of Sydney.*

Heckman, J. J., and E. Vytlacil (2001): "Local Instrumental Variables," in *Nonlinear Statistical Inference: Essays in Honor of Takeshi Amemiya*, ed. by C. Hsiao, K. Morimune, and J. Powell. Cambridge University Press, Cambridge.

——— (2005): "Structural equations, treatment effects, and econometric policy evaluation 1," *Econometrica*, 73, 669–738.

Huber, M., and G. Mellace (2015): "Testing instrument validity for LATE identification based on inequality moment constraints," *Review of Economics and Statistics*, 97, 398–411.

Imbens, G. W., and J. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.

Kane, T. J., and C. E. Rouse (1993): "Labor Market Returns to Two- and Four-Year Colleges: Is a Credit a Credit and Do Degrees Matter?," Working Paper 4268, National Bureau of Economic Research.

Kitagawa, T. (2015): "A test for instrument validity," *Econometrica*, 83, 2043–2063.

Manski, C. F., and J. V. Pepper (2000): "Monotone Instrumental Variables: With an Application to the Returns to Schooling," *Econometrica*, 68(4), 997–1010.

Mogstad, M., and M. Wiswall (2016): "Testing the quantityquality model of fertility: Estimation using unrestricted family size models," *Quantitative Economics*, 7(1), 157–192.

Mourifié, I., and Y. Wan (2017): "Testing LATE assumptions," *The Review of Economics and Statistics*, 99, 305–313.

Rubin, D. B. (1974): "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 66, 688–701.

Sharma, A. (2016): "Necessary and Probably Sufficient Test for Finding Valid Instrumental Variables," *working paper, Microsoft Research, New York.*

Slichter, D. (2014): "Testing Instrument Validity and Identification with Invalid Instruments," *mimeo, University of Rochester.*

Working Paper 492 february 2018

Authors
Martin Eckhoff ANDRESEN
Statistics Norway, Akersveien 26, 0177 Oslo, Norway. Phone: +47 62 88 50 00
Email: martin.andresen@ssb.no ; Website: https://sites.google.com/site/martineckhoffandresen/

Martin HUBER
University of Fribourg, Department of Economics, Bd. de Pérolles 90, 1700 Fribourg, Switzerland. Phone: +41 26 300 8274;
Email: martin.huber@unifr.ch; Website: http://www.unifr.ch/appecon/en/team/martin-huber

Abstract
When estimating local average and marginal treatment effects using instrumental variables (IV), multivalued endogenous treatments are frequently binarized based on a specific threshold in treatment support. However, such binarization introduces a violation of the IV exclusion if (i) the IV affects the multivalued treatment within support areas below and/or above the threshold and (ii) such IV-induced changes in the multivalued treatment affect the outcome. We discuss assumptions that satisfy the IV exclusion restriction with the binarized treatment and permit identifying the average effect of (i) the binarized treatment and (ii) unit-level increases in the original multivalued treatment among specific compliers. We derive testable implications of these assumptions and propose tests, which we apply to the estimation of the returns to (binary) college graduation instrumented by college proximity.

Citation proposal
Martin Eckhoff Andresen, Martin Huber. 2018. «Instrument-based estimation with binarized treatments: Issues and tests for the exclusion restriction». Working Papers SES 492, Faculty of Economics and Social Sciences, University of Fribourg (Switzerland)

Jel Classification
C12, C21, C26

Keywords
Instrumental variable, LATE, binarized treatment, test, exclusion restriction, MTE.

Working Papers SES collection

Last published
485 Buechel B., Mechtenberg L.: The Swing Voter’s Curse in Social Networks; 2017
486 Buechel B., Klössner S., Lochmüller M., Rauhut H.: The Strength of Weak Leaders. An Experiment on Social Infuence and Social Learning in Teams; 2017
487 Denisova-Schmidt E., Huber M., Leontyeva E., Solovyeva A.: Combining Experimental Evidence with Machine Learning to Assess Anti-Corruption Educational Campaigns among Russian University Students; 2017
488 Buechel B., Mechtenberg L., Julia Petersen J.: Peer Effects on Perseverance; 2017
489 Frölich M., Huber M.: Including covariates in the regression discontinuity design; 2017
490 Eugster N., Isakov D.: Founding family ownership, stock market returns, and agency problems; 2017
491 Eugster N.: Family Firms and Financial Analyst Activity; 2017

Catalogue and download links
http://www.unifr.ch/ses/wp
http://doc.rero.ch/collection/WORKING_PAPERS_SES

Publisher
Université de Fribourg, Suisse, Faculté des sciences économiques et sociales
Universität Freiburg, Schweiz, Wirtschafts- und sozialwissenschaftliche Fakultät
University of Fribourg, Switzerland, Faculty of Economics and Social Sciences

Bd de Pérolles 90, CH-1700 Fribourg
Tél.: +41 (0) 26 300 82 00
decanat-ses@unifr.ch www.unifr.ch/ses