Economics Letters 159 (2017) 23–27

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Testing the validity of the compulsory schooling law instrument

Benjamin Bolzern [a], Martin Huber [b,*]

[a] University of Bern, Switzerland
[b] University of Fribourg, Switzerland

HIGHLIGHTS

- We test the validity of compulsory schooling laws as instrument for education.
- We apply the tests of Huber and Mellace (2015) and Kitagawa (2015) to SHARE data.
- The results do not point to violations of instrument validity.

ARTICLE INFO

Article history:
Received 6 March 2017
Received in revised form 5 July 2017
Accepted 6 July 2017
Available online 15 July 2017

JEL classification:
C26
I12
I21
I28
J24

Keywords:
Instrumental variable
Schooling laws
Schooling reforms
Treatment effects
LATE
Tests

ABSTRACT

Based on Huber and Mellace (2015) and Kitagawa (2015), we test the validity of compulsory schooling laws as instrument for endogenous schooling choices in the SHARE data. We do not refute the instrument, but acknowledge that its validity might be violated even if testable implications hold.

1. Introduction

To evaluate the effects of education on later life outcomes, a vast literature relies on instrumental variables (IV) approaches to tackle the endogeneity of schooling decisions. A valid IV exogenously affects education while a direct effect on the outcome (other than through education) is excluded. Harmon and Walker (1995), Oreopoulos (2006), and many others use changes in compulsory schooling laws (CSL), which aim at shifting minimum education, as IV. Such analyses typically compare the schooling decisions and outcomes of adjacent cohorts just affected and not affected by changes in CSL. However, IV validity appears debatable. For instance, IV exogeneity does not hold if cohort specific economic shocks affect the outcome, see the discussion in Card (2001). Furthermore, the exclusion restriction likely fails if changes in CSL come with changes in the quality of education, as pointed out in Brunello et al. (2013).

To provide statistical insights on the IV validity of CSL, we jointly test random instrument assignment, weakly positive monotonicity of education in the instrument, IV exclusion restriction by applying the methods of Huber and Mellace (2015) and Kitagawa (2015) to data from the Survey of Health, Ageing and Retirement in Europe (SHARE). These assumptions allow identifying the local average treatment effect (LATE) on those whose education reacts to a change in CSL, the so-called compliers, see Angrist et al. (1996). Considering health- and income-related outcomes as well as various numbers of cohorts, testing does not provide evidence for IV invalidity. As a word of caution, we acknowledge that even asymptotically, the tests cannot detect all possible violations of IV validity when the complier share is larger than zero.

* Correspondence to: University of Fribourg, Bd. de Pérolles 90, 1700 Fribourg, Switzerland.
E-mail addresses: bolzern@gmail.com (B. Bolzern), martin.huber@unifr.ch (M. Huber).

http://dx.doi.org/10.1016/j.econlet.2017.07.004

The remainder is organized as follows. Section 2 briefly reviews the tests of Huber and Mellace (2015) and Kitagawa (2015). Section 3 describes the data. Section 4 presents the testing results.

2. Testing approach

We assume a binary treatment D ($D = 1$: attaining or exceeding a particular level of education vs. $D = 0$: not attaining this level), whose impact on outcome Y (health or income) is of policy interest. Z is the supposed instrument ($Z = 1$: exposed to CSL reform vs. $Z = 0$: not exposed). Denote by $D(z)$ the potential treatment state for instrument $Z = z$, and by $Y(d)$ the potential outcome for treatment $D = d$. The population can be categorized into four types (denoted by $T \in \{a, c, d, n\}$) as function of the potential treatments, see Angrist et al. (1996): compliers (c: $D(1) = 1, D(0) = 0$), always-takers (a: $D(1) = 1, D(0) = 1$), never-takers (n: $D(1) = 0, D(0) = 0$), and defiers (d: $D(1) = 0, D(0) = 1$). In the context of our application, compliers increase education in reaction to a rise in minimum compulsory schooling, while defiers pick a higher level of schooling in the absence than in the presence of this rise. Under the following assumptions the LATE $E[Y(1) - Y(0)|T = c]$ is obtained by the probability limit of 2 stage least squares: (i) $E(Y(d)|T = t, Z = z) = E(Y(d)|T = t)$ for $d, z \in \{0, 1\}$ and $t \in \{a, c, n\}$ (mean exclusion restriction); (ii) $\Pr(T = t|Z = 1) = \Pr(T = t|Z = 0)$ for $t \in \{a, c, n\}$ (unconfounded type); (iii) $\Pr(T = d) = 0$ (monotonicity); (iv) $\Pr(T = c) > 0$ (existence of compliers).

Huber and Mellace (2015) show that under Assumptions (i)–(iii), it must hold that[1]

$$\begin{aligned} E(Y|D = 1, Z = 1, Y \le y_q) &\le E(Y|D = 1, Z = 0) \\ &\le E(Y|D = 1, Z = 1, Y \ge y_{1-q}), \\ E(Y|D = 0, Z = 0, Y \le y_r) &\le E(Y|D = 0, Z = 1) \\ &\le E(Y|D = 0, Z = 0, Y \ge y_{1-r}). \end{aligned} \quad (1)$$

$q = \Pr(D = 1|Z = 0)/\Pr(D = 1|Z = 1)$ corresponds to the share of always-takers conditional on $D = 1$ and $Z = 1$, and y_q is the qth quantile of Y given $D = 1$ and $Z = 1$. $r = \Pr(D = 0|Z = 1)/\Pr(D = 0|Z = 0)$ corresponds to the share of never-takers conditional on $D = 0$ and $Z = 0$, and y_r is the rth quantile of Y given $D = 0$ and $Z = 0$. Considering the first line of (1), the intuition of the testable constraint is as follows: $E(Y|D = 1, Z = 0)$ point identifies the mean potential outcome of the always-takers under treatment, as any subject with $D = 1, Z = 0$ must be an always-taker in the absence of defiers. Furthermore, the mean potential outcomes of the always-takers are bounded by the averages in the upper and lower outcome proportions with $D = 1$ and $Z = 1$ that correspond to the share of the always-takers in the mixed population with compliers: $E(Y|D = 1, Z = 1, Y \le y_q)$, $E(Y|D = 1, Z = 1, Y \ge y_{1-q})$. $E(Y|D = 1, Z = 0)$ must lie within the latter bounds, otherwise the assumptions are necessarily violated. An analogous result applies to the mean potential outcome of never-takers under non-treatment. Any procedure suitable for testing multiple moment inequalities could be used for verifying (1), for instance the method by Chen and Szroeter (2014).

For the case that (i) and (ii) is strengthened to full statistical independence of Z and potential treatments/outcomes, i.e. Z being statistically independent of $\{Y(d), D(z)\}$ for $d, z \in \{0, 1\}$, Kitagawa (2015) proposes a test based on verifying the following constraints in the data:

$$\begin{aligned} &\Pr(Y \in A, D = 1|Z = 1) \ge \Pr(Y \in A, D = 1|Z = 0), \\ &\Pr(Y \in A, D = 0|Z = 0) \ge \Pr(Y \in A, D = 0|Z = 1), \end{aligned} \quad (2)$$

where A denotes a subset of the support of Y. A violation of (2) would imply a negative complier share in subset A given Z under the imposed IV assumptions, which is impossible. The method of Kitagawa (2015) is based on resampling a variance-weighted two sample Kolmogorov–Smirnov-type statistic using the supremum of $\Pr(Y \in A, D = 1|Z = 0) - \Pr(Y \in A, D = 1|Z = 1)$ and $\Pr(Y \in A, D = 0|Z = 1) - \Pr(Y \in A, D = 0|Z = 0)$, respectively, across multiple subsets A.[2]

3. Data

Our treatment and outcome variables come from the first wave of the Survey of Health, Ageing and Retirement in Europe (SHARE) conducted in 2004, which targets individuals who are 50 years or older and reside in Europe or Israel. Our binary CSL instrument is based on schooling reforms in seven countries in the 1960s or 1970s: Austria, Denmark, France, Greece, Italy, the Netherlands, and Spain. In our evaluation sample, we pool these seven countries.[3] Table 1 is based on Brunello et al. (2009) and provides the year of the reform, the pivotal cohort first potentially affected by it, and the expected change in minimum educational attainment as a consequence of the reform, measured according to the International Standard Classification of Education (ISCED). The ISCED levels are the following: 1—primary education or first stage of basic education; 2—lower secondary education or second stage of basic education; 3—upper secondary education; 4—post-secondary non-tertiary education; 5—first stage of tertiary education; 6—second stage of tertiary education.

We pool the data from the seven countries and define the binary treatment variable D to be one if an individual attains the expected minimum ISCED level after the respective CSL reform or a higher degree, and zero otherwise. For example, for an individual from Austria the treatment is one if she attained at least ISCED 2 and zero in the case of ISCED 1, while for Denmark, the treatment is one for individuals attaining at least ISCED 3 and zero otherwise.[4] For the construction of the instrument Z, we make use of adjacent cohorts not yet exposed to the respective CSL reform, for whom Z is zero, and exposed to the reform, for whom Z is one. In the analysis, we consider three different symmetric data windows around the point when the respective reform came into force. The smallest window contains four cohorts: the two cohorts prior to the pivotal cohort, the pivotal cohort, and the following cohort. Likewise, a medium-sized window with a total of eight cohorts and a large window with sixteen cohorts is chosen. We apply the test to six health- and income-related outcomes Y: frequency of alcoholic beverages consumed during the last 6 months measured on a scale from 0 (not at all) to 6 (almost every day), BMI, number of doctoral visits in the last 12 months, self-reported health measured on an scale from 1 (poor) to 5 (excellent), labour earnings in euros in 2003, and pooled income in euros from earnings and pensions in 2003.

[1] Under Assumptions (i)–(iii) and a violation of Assumption (iv), the weak inequalities in (1) become equalities.

[2] See Mourifie and Wan (forthcoming) for an alternative test of such assumptions.

[3] Although present in the first panel wave, the following countries were excluded: (1) Belgium, because its reform of 1983 did not affect any individual in the first wave of SHARE; (2) Germany and Switzerland, since they did not experience a single national schooling reform; (3) Sweden, because it lacks variation in the treatment variable to be defined; (4) Israel, as its schooling reform of 1968 was only partially implemented.

[4] A striking feature in the data is that there exist untreated individuals who did receive the instrument, i.e. never-takers. In our CSL setting, such individuals should not exist if those laws were strictly enforced. One reason for their occurrence may be gradual implementation of the policy over time. For example, complete implementation of Italy's reform of 1963 was, according to Brunello et al. (2009), not immediate, but took 13 years. Our data confirms gradual compliance. Based on the sample for the outcome income, we find for Italy that among the pivotal and the following cohort there is a 18% share of never-takers. The share drops to 12% in the two following cohorts and 7% in the following four cohorts.

Table 1
Schooling reforms.

Country	Year of reform	Pivotal cohort	Expected change in minimum ISCED level
Austria	1962	1947	ISCED1 to ISCED 2
Denmark	1971	1957	ISCED 2 to ISCED 3
France	1959[a]	1953	ISCED 2 to ISCED 3
Greece	1975	1963	ISCED 1 to ISCED 2
Italy	1963	1949	ISCED1 to ISCED 2
Netherlands	1975	1959	ISCED1 to ISCED 2
Spain	1970	1957	ISCED 1 to ISCED 2

Note:
[a] The Berthoin Reform of 1959 was implemented in 1967.

Table 2
Descriptive statistics.

Outcome	±2 cohorts				±4 cohorts				±8 cohorts			
	obs	range Y	av D	av Z	obs	range Y	av D	av Z	obs	range Y	av D	av Z
Alc. beverages	1124	0–6	0.46	0.67	2306	0–6	0.35	0.66	4766	0–6	0.25	0.57
BMI (rounded)	1116	15–51	0.46	0.67	2289	15–59	0.35	0.66	4711	15–60	0.25	0.57
Doctor visits	1121	0–98	0.46	0.67	2295	0–98	0.35	0.66	4745	0–98	0.25	0.57
Health	583	1–5	0.48	0.67	1157	1–5	0.37	0.65	2328	1–5	0.26	0.56
Income in million	1130	0–0.33	0.46	0.67	2317	0–1.50	0.35	0.66	4785	0–100	0.25	0.57
Pooled inc. in m.	1130	0–0.33	0.46	0.67	2317	0–1.50	0.35	0.66	4785	0–100	0.25	0.57

Table 3
Empirical results.

Outcome	±2 cohorts				±4 cohorts				±8 cohorts			
	obs	first	p HM	p K	obs	first	p HM	p K	obs	first	p HM	p K
Total sample												
Alc. beverages	1124	0.08	1.00	0.98	2306	0.11	0.99	0.99	4766	0.12	1.00	1.00
BMI	1116	0.07	0.99	0.61	2289	0.10	1.00	0.91	4711	0.11	1.00	0.95
Doctor visits	1121	0.08	1.00	0.99	2295	0.10	0.57	0.82	4745	0.11	1.00	1.00
Health	583	0.09	1.00	0.95	1157	0.11	1.00	0.97	2328	0.12	1.00	0.99
Income	1130	0.08	0.98	0.57	2317	0.11	0.88	0.86	4785	0.12	0.81	1.00
Pooled income	1130	0.08	0.99	0.63	2317	0.11	0.93	1.00	4785	0.12	0.81	0.96
Female												
Alc. beverages	665	0.07	0.77	0.72	1329	0.08	0.39	0.27	2786	0.11	1.00	0.84
BMI	660	0.07	1.00	0.79	1317	0.08	1.00	0.71	2741	0.11	1.00	0.96
Doctor visits	663	0.07	0.97	0.79	1321	0.08	1.00	0.74	2772	0.10	1.00	0.84
Health	340	0.10	1.00	0.99	671	0.09	1.00	0.77	1383	0.11	1.00	0.87
Income	1130	0.08	0.97	0.56	1333	0.08	0.76	0.86	2793	0.11	1.00	1.00
Pooled income	666	0.07	0.17	0.73	1333	0.08	0.74	0.95	2793	0.11	0.99	0.97

(*continued on next page*)

Our evaluation samples based on the various cohorts only contain individuals without missing values in the treatment, the instrument, and the respective outcome. Furthermore, individuals who had immigrated to the respective countries at the age 11 or later are dropped, as they might not have been affected by the reform. Separately by outcomes and the number of cohorts, Table 2 reports the number of observations ('obs'), the empirical range of values in the outcome ('range Y'), the share of treated ('av D'), and the share of instrumented individuals ('av Z'). An implication of the age distribution in the first wave of SHARE is that our samples are skewed towards an increasingly smaller fraction receiving the instrument as more cohorts are included. The fraction receiving the instrument is approximately 46% when using 4 cohorts. This share drops to around 36% for 8 cohorts and to approximately 25% for 16 cohorts. We also note that the contribution of the seven countries to our samples is not uniform. Observations from Austria, France and Italy account for approximately 92% of observations in our samples, with each having a share of between 23% and 37% depending on the three sample sizes.[5]

4. Empirical results

Table 3 presents the testing results for various numbers of not affected (i.e. earlier) and affected (i.e. later) cohorts included in the analysis, namely ±2 cohorts around the change in CSL, ±4 cohorts, and ±8 cohorts. For each number of cohorts, the tests are run in the respective total sample as well as in subsamples conditional on gender or various socio-economic characteristics when the interviewee was 10 years old. The latter include language competence, drinking or mental problems of parents, number of books at home, availability of hot water, and household's rooms to members ratio. The column 'outcome' contains the outcome variables for which (1) and (2) are tested. 'obs' reports the number of observations. 'first' gives the first stage effect of Z on D, which is an unbiased estimate of the complier share if the LATE assumptions hold. Column 'p HM' provides the p-value when implementing the Huber and Mellace (2015) test based on the method by Chen

[5] A further point worth noting is that even though SHARE targets individuals who are 50 plus, our samples also contain observations under the age of 50, as spouses of eligible respondents are included in the data set. However, their numbers are substantially lower and generally decrease in the year of birth. These spouses are predominantly female, causing an over-representation of women in particular in countries with relatively late reforms. For instance, 58% of our respondents in the BMI sample with 16 cohorts are female and 61% of the subsample receiving the instrument are female. However, countries with late reforms do not contribute importantly to our samples.

Table 3 (*continued*)

Outcome	±2 cohorts				±4 cohorts				±8 cohorts			
	obs	first	p HM	p K	obs	first	p HM	p K	obs	first	p HM	p K
Male												
Alc. beverages	459	0.08	0.99	0.68	977	0.14	1.00	0.98	1980	0.13	1.00	0.61
BMI	456	0.08	1.00	0.83	972	0.14	1.00	0.74	1970	0.13	1.00	1.00
Doctor visits	458	0.08	0.92	0.97	974	0.14	1.00	0.92	1973	0.13	0.99	1.00
Health	243	0.06	1.00	0.87	486	0.15	1.00	0.98	945	0.14	1.00	0.99
Income	1130	0.08	0.97	0.55	984	0.14	0.97	0.58	1992	0.13	0.51	0.86
Pooled income	464	0.08	0.98	0.70	984	0.14	0.98	0.56	1992	0.13	0.47	0.91
Bad or very bad relative position to others in terms of language at age 10												
Alc. beverages	115	0.03	0.83	0.77	213	0.09	1.00	0.96	448	0.12	1.00	0.95
BMI	113	0.03	0.95	0.68	209	0.11	0.79	0.85	440	0.13	0.83	0.96
Doctor visits	115	0.03	0.24	0.61	212	0.09	0.99	0.39	447	0.12	0.66	0.69
Health	60	0.07	0.76	0.62	114	0.21	1.00	0.96	232	0.15	1.00	0.96
Income	115	0.03	0.68	0.75	213	0.09	0.50	0.37	448	0.12	0.99	0.85
Pooled income	115	0.03	0.74	0.21	213	0.09	0.57	0.40	448	0.12	0.98	0.85
Neither bad nor very bad relative position to others in terms of language at age 10												
Alc. beverages	529	0.06	0.94	0.74	1084	0.10	1.00	0.99	2257	0.10	0.97	0.70
BMI	526	0.06	0.99	0.56	1080	0.10	1.00	0.93	2242	0.10	1.00	0.74
Doctor visits	528	0.06	1.00	0.71	1079	0.10	1.00	0.94	2247	0.10	0.97	0.89
Health	272	0.04	0.98	0.42	552	0.05	1.00	0.87	1105	0.09	1.00	0.75
Income	531	0.07	0.66	0.77	1089	0.10	0.50	0.67	2263	0.10	0.90	0.89
Pooled income	531	0.07	0.53	0.90	1089	0.10	0.53	0.73	2263	0.10	0.89	0.99
Parents drank or had mental problems at age 10												
Alc. beverages	74	−0.02	0.54	0.52	146	0.08	0.76	0.87	278	0.13	0.49	0.69
BMI	74	−0.02	0.40	0.50	146	0.08	0.84	0.80	276	0.14	1.00	0.98
Doctor visits	74	−0.02	0.69	0.35	146	0.08	0.87	0.77	277	0.13	0.95	0.60
Health	42	−0.04	0.58	0.30	78	−0.03	0.11	0.39	135	0.07	0.94	0.78
Income	75	−0.01	0.73	0.65	148	0.08	0.64	0.98	280	0.13	0.98	0.92
Pooled income	75	−0.01	0.72	0.28	148	0.08	0.65	0.84	280	0.13	0.98	0.95
Parents did not drink or have mental problems at age 10												
Alc. beverages	576	0.07	0.67	0.37	1164	0.10	0.97	0.98	2445	0.10	1.00	0.95
BMI	571	0.07	1.00	0.94	1156	0.10	1.00	0.79	2424	0.10	1.00	0.71
Doctor visits	575	0.07	1.00	0.40	1158	0.09	0.99	0.99	2435	0.10	1.00	0.98
Health	293	0.05	0.83	0.74	596	0.09	1.00	0.93	1214	0.10	1.00	0.98
Income	577	0.07	0.95	0.78	1167	0.10	0.92	0.63	2450	0.10	0.64	0.92
Pooled income	577	0.07	0.97	0.90	1167	0.10	0.94	0.55	2450	0.10	0.62	0.65
0–10 books at home at age 10												
Alc. beverages	314	0.04	0.87	0.73	625	0.09	0.77	0.54	1243	0.12	0.99	0.80
BMI	310	0.04	0.90	0.74	620	0.09	1.00	0.91	1234	0.13	1.00	1.00
Doctor visits	313	0.04	0.76	0.74	620	0.08	0.72	0.96	1237	0.12	0.99	0.94
Health	169	0.09	0.99	0.89	334	0.12	1.00	0.94	644	0.13	1.00	0.99
Income	315	0.04	0.12	0.83	626	0.09	1.00	0.97	1245	0.12	0.98	0.98
Pooled income	315	0.04	0.14	0.83	626	0.09	1.00	0.76	1245	0.12	0.97	1.00
More than 10 books at home at age 10												
Alc. beverages	336	0.05	0.82	0.85	682	0.10	1.00	0.77	1472	0.10	0.93	0.39
BMI	335	0.05	1.00	0.99	679	0.09	1.00	0.57	1458	0.10	1.00	0.91
Doctor visits	336	0.05	1.00	0.59	681	0.09	0.95	0.89	1467	0.10	0.96	0.98
Health	166	−0.01	0.86	0.48	337	0.03	0.90	0.67	700	0.09	1.00	0.78
Income	337	0.05	0.74	0.91	686	0.10	0.58	0.85	1477	0.10	0.17	0.50
Pooled income	337	0.05	0.75	0.94	686	0.10	0.55	0.72	1477	0.10	0.23	0.53
Hot water at home at age 10												
Alc. beverages	297	−0.01	0.83	0.66	590	0.07	1.00	0.85	1181	0.06	0.82	0.90
BMI	297	−0.01	0.77	0.57	589	0.07	1.00	0.87	1175	0.06	1.00	0.65
Doctor visits	297	−0.01	0.74	0.67	587	0.07	1.00	0.52	1176	0.06	1.00	0.98
Health	144	−0.01	0.61	0.71	293	0.06	1.00	0.97	572	0.07	1.00	0.54
Income	297	−0.01	0.95	0.61	592	0.07	0.17	0.71	1185	0.06	1.00	0.76
Pooled income	297	−0.01	0.96	0.36	592	0.07	0.18	0.89	1185	0.06	1.00	0.67
No hot water at home at age 10												
Alc. beverages	353	0.10	1.00	0.99	719	0.11	0.99	0.97	1538	0.13	1.00	1.00
BMI	348	0.10	1.00	0.63	712	0.11	1.00	0.96	1521	0.13	1.00	0.83
Doctor visits	352	0.10	0.92	0.68	716	0.11	0.60	0.87	1532	0.13	1.00	1.00
Health	191	0.07	0.99	0.67	380	0.08	1.00	1.00	776	0.11	1.00	0.83
Income	355	0.10	1.00	0.96	722	0.11	0.92	0.89	1541	0.13	0.64	1.00
Pooled income	355	0.10	1.00	1.00	722	0.11	0.96	0.99	1541	0.13	0.73	0.99

(*continued on next page*)

Table 3 (*continued*)

Outcome	±2 cohorts				±4 cohorts				±8 cohorts			
	obs	first	p HM	p K	obs	first	p HM	p K	obs	first	p HM	p K
(Household's rooms: members) ≥ 0.35 at age 10												
Alc. beverages	533	0.04	0.66	0.28	1095	0.09	1.00	0.50	2307	0.11	1.00	1.00
BMI	529	0.04	0.98	0.78	1089	0.09	1.00	0.89	2289	0.11	1.00	0.99
Doctor visits	532	0.04	0.82	0.84	1091	0.09	1.00	0.81	2298	0.11	0.98	0.99
Health	269	0.00	0.30	0.61	552	0.06	1.00	0.79	1134	0.09	1.00	0.75
Income	535	0.04	1.00	0.46	1100	0.10	0.84	0.98	2314	0.11	0.73	0.96
Pooled income	535	0.04	1.00	0.84	1100	0.10	0.86	0.86	2314	0.11	0.73	1.00
(Household's rooms: members) < 0.35 at age 10												
Alc. beverages	113	0.03	0.80	0.61	208	0.05	0.71	0.74	399	0.09	1.00	0.63
BMI	112	0.05	0.88	0.78	206	0.06	0.99	0.83	394	0.09	1.00	0.96
Doctor visits	113	0.03	0.91	0.41	206	0.04	1.00	0.84	397	0.08	0.94	0.99
Health	63	0.06	0.95	0.42	117	0.08	0.54	0.57	207	0.17	1.00	0.90
Income	113	0.03	0.83	0.31	208	0.05	0.87	0.84	399	0.09	0.99	0.95
Pooled income	113	0.03	0.84	0.84	208	0.05	0.87	0.94	399	0.09	0.95	0.71

and Szroeter (2014) and 999 bootstrap replications. Column 'p K' contains the p-value of the Kitagawa (2015) test with 15 subsets A defined on a quantile-based grid over the support of Y.[6]

For any outcome and sample, the p-values of either test are larger than conventional levels of significance.[7] We therefore conclude that we find no statistical evidence for a violation of the CSL instrument in the SHARE data. As a word of caution, however, we point out that even asymptotically, such tests cannot detect all potential violations of Assumptions (i)–(iii) if the complier share is non-zero.[8]

Acknowledgement

We have benefited from comments by an anonymous referee.

References

Angrist, J., Imbens, G.W., Rubin, D., 1996. Identification of causal effects using instrumental variables. J. Amer. Statist. Assoc. 91, 444–472 (with discussion).

Brunello, G., Fort, M., Weber, G., 2009. Changes in compulsory schooling, education and the distribution of wages in europe. Econ. J. 119, 516–539.

Brunello, G., Fort, M., Weber, G., Weiss, C.T., 2013. Testing the internal validity of compulsory school reforms as instrument for years of schooling, IZA Discussion Paper No. 7533.

Card, D., 2001. Estimating the return to schooling: progress on some persistent econometric problems. Econometrica 69, 1127–1160.

Chen, L.-Y., Szroeter, J., 2014. Testing multiple inequality hypotheses: a smoothed indicator approach. J. Econometrics 178, 678–693.

Harmon, C., Walker, I., 1995. Estimates of the economic return to schooling for the united kingdom. Amer. Econ. Rev. 85, 1278–1286.

Huber, M., Mellace, G., 2015. Testing instrument validity for late identification based on inequality moment constraints. Rev. Econ. Stat. 97, 398–411.

Kitagawa, T., 2015. A test for instrument validity. Econometrica 83, 2043–2063.

Mourifie, I., Wan, Y., 2016. Testing local average treatment effect assumptions. Rev. Econ. Stat. (forthcoming).

Oreopoulos, P., 2006. Estimating average and local average treatment effects of education when compulsory schooling laws really matter. Amer. Econ. Rev. 96, 152–175.

[6] Choosing the number of subsets to be 3, 5, 10, 25, 50, or 100 yields similar results.

[7] Our findings do not change when applying a modified version of the Kitagawa (2015) test that allows for a multivalued treatment. We use the procedure to jointly test the IV assumptions under treatment switches from ISCED 1 to at least ISCED 2 and from ISCED 2 or less to at least ISCED 3 but do not reject the instrument.

[8] Our results are in line with Brunello et al. (2013), who do not find evidence for a violation of the exclusion restriction through educational quality based on a different testing approach based on distinct measures of students' cognitive test scores.