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]

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

$q = \Pr(D = 1|Z = 0)/\Pr(D = 1|Z = 1)$ 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: $E(Y|D = 1, Z = 1, Y \leq y_q)$, $E(Y|D = 1, Z = 1, Y \geq y_{1-q})$. $E(Y|D = 1, Z = 0)$ 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:

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

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.