Economics Letters 194 (2020) 109320

Contents lists available at ScienceDirect

Economics Letters

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

An alternative test for conditional unconfoundedness using auxiliary variables☆

Ying Fang[a,b], Shengfang Tang[b], Zongwu Cai[c], Ming Lin[a,b,*]

[a] Wang Yanan Institute for Studies in Economics and Fujian Key Laboratory of Statistical Sciences, Xiamen University, Xiamen, Fujian 361005, China
[b] Department of Statistics, School of Economics, Xiamen University, Xiamen, Fujian 361005, China
[c] Department of Economics, University of Kansas, Lawrence, KS 66045, USA

ARTICLE INFO

Article history:
Received 5 March 2020
Received in revised form 22 May 2020
Accepted 10 June 2020
Available online 17 June 2020

JEL classification:
C12
C13
C14

Keywords:
Conditional unconfoundedness
Moment test
Treatment effect

ABSTRACT

This paper proposes an alternative test procedure for testing the conditional unconfoundedness assumption which is an important identification condition commonly imposed in the literature of program analysis and policy evaluation. We transform the conditional unconfoundedness test to a nonparametric conditional moment test using an auxiliary variable which is independent of the treatment assignment variable conditional on potential outcomes and observable covariates. The proposed test statistic is shown to have a limiting normal distribution under the null hypothesis of conditional independence. Monte Carlo simulations are conducted to examine the finite sample performances of the proposed test statistics. Finally, the proposed test method is applied to test the conditional unconfoundedness in the real example of the return to college education.

1. Introduction

The conditional unconfoundedness refers to the assumption that conditional on observable confounders, the potential outcomes are independent of treatment status. In many applications, a weaker version, the conditional mean independence, is sufficient to identify treatment effects. Although the conditional unconfoundedness assumption plays a central role in identifying the average treatment effects, there are only a few test statistics available in the literature to test the conditional mean independence. Using binary instrumental variables (IV), Donald et al. (2014) proposed a Durbin–Wu–Hausman type test statistic to test the conditional mean independence by comparing two estimators: the one based on the local average treatment effect on the treated, which is related to the choice of the IV, and the other one based on the average treatment effect for the treated, which does not rely on the choice of the IV. Recently, by assuming the error terms in both the outcome equation and the selection equation to be symmetrically distributed, Chen et al. (2017) proposed a Kolmogorov–Smirnov type test statistic to test conditional mean independence by comparing two estimators, the one which is only valid with conditional mean independence assumption and the other without it.

This paper proposes an alternative method to test the conditional unconfoundedness complementary to the aforementioned tests in the literature. Instead of the availability of a binary instrumental variable in Donald et al. (2014) and the requirement of symmetrically distributed error terms in Chen et al. (2017), our method relies on the existence of an auxiliary variable which is correlated to potential outcomes but is independent of the treatment status given on potential outcomes and observable covariates. In other words, this auxiliary variable is possible to have an effect on the treatment choice. However, the linkage from the auxiliary variable to the treatment status can be fully captured by potential outcomes and observable covariates. When such auxiliary variables are available, the conditional independence test can be simply implemented by a conditional moment test using nonparametric method. Moreover, compared to Donald et al. (2014) and Chen et al. (2017), our method can be applied to testing not only the conditional mean independence but also the conditional independence, the stronger version. The auxiliary variable assumption has been widely used in the literature of dealing with missing data problems as in Zhao and Shao (2015) and Breunig (2019), measurement error problems in Hu and Schennach (2008), and other scenarios.

☆ The authors thank the Associate Editor and the anonymous referees for their helpful and constructive comments and they also acknowledge the financial supports, in part, from the National Science Fund of China (NSFC) for Distinguished Scholars (71625001), the NSFC key projects with grant numbers 71631004 and 71131008, the NSFC Basic Scientific Center project (71988101), and the Science Foundation of Ministry of Education of China (19YJA910003).

* Corresponding author at: Wang Yanan Institute for Studies in Economics, Xiamen, Fujian 361005, China.
E-mail address: linming50@xmu.edu.cn (M. Lin).

https://doi.org/10.1016/j.econlet.2020.109320
0165-1765/

2. Testing conditional unconfoundedness

The model is developed within the conventional framework of the Rubin causal model, where $Y(1)$ and $Y(0)$ denote the potential outcomes for a unit receiving or not receiving treatment, respectively. Let D denote whether the treatment of interest is received, with $D = 1$ if the unit receives the treatment, otherwise $D = 0$. In addition, each unit is also characterized by a vector of covariates denoted by $X \in \mathbb{R}^d$. The fundamental problem in the treatment effect literature is that exactly one (never both) of the two potential outcomes $Y(1)$ and $Y(0)$ are observed for a particular individual. So using the notation above, for each individual, the observed data are only (Y, D, X) where $Y = D \cdot Y(1) + (1 - D) \cdot Y(0)$. A commonly used assumption to identify the treatment effect parameters of interest is the conditional unconfoundedness assumption, that is,

$$\big(Y(0), Y(1)\big) \perp\!\!\!\perp D \mid X,$$

where $\perp\!\!\!\perp$ indicates statistical independence. However, this assumption may be violated in practice if there exist unobserved confounders which affect both potential outcomes and the treatment assignment variable. Thus, it is desirable to formally propose a procedure to test whether the conditional unconfoundedness assumption holds or not.

We propose a novel method to test the conditional independence assumption by assuming the availability of a vector of auxiliary variables $Z \in \mathbb{R}^r$ to satisfy the following assumption.

Assumption 2.1. (i) Assume that there exists a vector of continuously distributed variables $Z \in \mathbb{R}^r$ which are correlated with both $Y(0)$ and $Y(1)$ and satisfy the following condition

$$Z \perp\!\!\!\perp D \mid \big(Y(0), Y(1), X\big).$$

(ii) (Bounded completeness) For each bounded function $\psi_m(\cdot)$, $E[\psi_m(Y(0), Y(1), X)|X, Z] = 0$ implies that $\psi_m(Y(0), Y(1), X) = 0$ almost surely (a.e.).

Assumption 2.1(i) requires whether receiving treatment for a unit is primarily determined by the potential outcomes $\big(Y(0), Y(1)\big)$ and the covariates X. The statement is thus that, given $Y(0)$, $Y(1)$ and X, the treatment assignment variable D and the auxiliary variable Z are mutually independent and the information about D from Z can then be completely captured by $\big(Y(0), Y(1), X\big)$. This assumption is appropriate when receiving treatment is driven by potential outcomes $Y(0)$ and $Y(1)$ and once given the information of $Y(0)$, $Y(1)$, and X, Z does not include any additional information on the assignment mechanism. Furthermore, Assumption 2.1(ii) is normally referred to as the bounded completeness in Z of the conditional distribution of $\big(Y(0), Y(1)\big)$ conditional on (X, Z). There are many families of distributions that are bounded complete and sufficient conditions for bounded completeness can be found in Mattner (1993), D'Haultfoeuille (2011) and among others. Different from Donald et al. (2014) for requiring the instrumental variable to be binary, our Z in the above assumption can be continuous. Finally, we remark that similar to finding an instrumental variable in a model with endogeneity and/or missing at random as in Breunig (2019), it might not be an easy task to find an auxiliary variable Z in real applications, so that there are no general guidelines on how to find Z.

Similar to Breunig (2019), we can show that under Assumption 2.1(i), the conditional unconfoundedness assumption implies that $E\big(D - E(D|X)|X, Z\big) = 0$. If both Assumption 2.1(i) and (ii) are satisfied, the conditional unconfoundedness assumption is equivalent to $E(D|X, Z) = E(D|X)$. In other words, the conditional unconfoundedness assumption can be tested by examining whether the auxiliary variable Z has explanatory power for the mean of the treatment assignment variable D given covariates X. Thus, one can test the conditional independence using the following conditional moment test:

$$H_0 : E(D|X, Z) = E(D|X) \;\; a.e. \;\text{ versus }\; H_1 : E(D|X, Z) \neq E(D|X) \tag{2.1}$$

on a set with positive measure.

Following Fan and Li (1996) and Li (1999), we adopt the kernel estimation method to construct the test statistic under the null hypothesis. To this end, we first introduce some notations. Let $W = (X', Z')' \in \mathbb{R}^p$, where X is of dimension d and Z is of dimension r, $d + r = p$, and $\{Y_i, D_i, X_i, Z_i\}_{i=1}^n$ be a set of n independent and identically distributional (iid) observations on (Y, D, X, Z). Define $\varepsilon = D - m(X)$, where $m(X) = E(D|X)$. Note that $T = E[\varepsilon E(\varepsilon|W)] = E\big\{[E(\varepsilon|W)]^2\big\} \geq 0$ and the equality holds if and only if H_0 is true. Hence, T can serve as a proper candidate for consistent testing H_0 and we may use the sample analogue of T to form a test as

$$T_n^* = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i E(\varepsilon_i|W_i).$$

However, this test statistic is infeasible because ε_i and $E(\varepsilon_i|W_i)$ are not observed directly but they can be estimated by some standard nonparametric techniques. To be specific, to obtain a feasible test statistic, we first estimate ε_i and $E(\varepsilon_i|W_i)$ nonparametrically and then plug the corresponding estimates into the test statistic T_n^* to obtain a feasible version. In order to avoid the random denominator problem, we follow the standard procedure to adopt a density weighted version of T, which is $T_n^{**} = \frac{1}{n}\sum_{i=1}^{n}\big[\varepsilon_i f(X_i)\big]E\big[\varepsilon_i f(X_i)|W_i\big]f_W(W_i)$, where $f(\cdot)$ is the density function of X and $f_W(\cdot)$ is the density function of W.

Define a leave-one-out kernel estimator of $E(D_i|X_i)$ as

$$\widehat{D}_i = \frac{1}{(n-1)h_1^d}\sum_{j\neq i, j=1}^{n} K_1\Big(\frac{X_j - X_i}{h_1}\Big)D_j \Big/ \widehat{f}(X_i),$$

where

$$\widehat{f}(X_i) = \frac{1}{(n-1)h_1^d}\sum_{j\neq i, j=1}^{n} K_1\Big(\frac{X_j - X_i}{h_1}\Big),$$

is the leave-one-out kernel estimator of $f(X_i)$ with $K_1(\cdot)$ being a kernel function and h_1 denoting the bandwidth to estimate $m(\cdot)$. Then, a kernel-based sample analogue of T_n^{**} is given by

$$T_n = \frac{1}{n(n-1)h^p}\sum_{i=1}^{n}\sum_{j\neq i, j=1}^{n} \big(\widehat{\varepsilon}_i \widehat{f}(X_i)\big)\big(\widehat{\varepsilon}_j \widehat{f}(X_j)\big)K_{ij},$$

where $\widehat{\varepsilon}_i = D_i - \widehat{D}_i$ is the nonparametric residual estimator and $K_{ij} = K((W_j - W_i)/h)$ with $K(\cdot)$ being a kernel function and h denoting the bandwidth to estimate $E(\varepsilon|W)$.

Before establishing the asymptotic distribution of the test statistic T_n under H_0, the following assumptions are provided, where the definitions of Robinson (1988), Fan and Li (1996) and Li (1999) for the class of kernel functions $\Im_\lambda$ and the class of functions $\wp_\vartheta^\alpha$ are used.

Assumption 2.2. (i) $f(\cdot) \in \wp_\nu^\infty$, $m(x) = E(D \mid X = x) \in \wp_\nu^{4+\iota}$ and $f_W(w) \in \wp_\nu^\infty$ for some $\nu \geq 2$ and $\iota > 0$.
(ii) Let $K_1(\cdot)$ be a ν-th order kernel and let $K(\cdot)$ be a nonnegative second order kernel.
(iii) The conditional variance function $\sigma^2(w) = E(\varepsilon^2|W = w)$ and $\mu_4(w) = E(\varepsilon^4|W = w)$ are continuous. In addition, $f_W(w)\sigma^2(w)$ and $f_W(w)\mu_4(w)$ are bounded on $\mathbb{R}^p$.

Assumption 2.3. As $n \to \infty$, $h_1 \to 0$, $h \to 0$, $nh_1^d \to \infty$, $nh^p \to \infty$, $nh^{p/2}h_1^{2\nu} \to 0$ and $h^p/h_1^{2d} \to 0$.

These assumptions are quite standard and can be seen in many nonparametric test literatures. With Assumptions 2.2 and 2.3, the asymptotic distribution of the test statistic T_n under H_0 can be derived, which is formally summarized in the following theorem with its detailed proof available upon request.

Theorem 2.1. *Suppose Assumptions 2.2 and 2.3 are satisfied. Then, we have*

(1) Under H_0, $\widetilde{T}_n := \frac{nh^{p/2}T_n}{\sqrt{2}\widehat{\sigma}_T} \xrightarrow{d} \mathcal{N}(0,1)$, *where*

$$\widehat{\sigma}_T^2 = \frac{1}{n(n-1)h^p}\sum_{i=1}^{n}\sum_{j\neq i}\left(\widehat{\varepsilon}_i\widehat{f}(X_i)\right)^2\left(\widehat{\varepsilon}_j\widehat{f}(X_j)\right)^2 K_{ij}\cdot\left(\int K^2(v)dv\right),$$

is a consistent estimator of σ_T^2 *given by*

$$\sigma_T^2 = E\left(f^4(X)f_W(W)\sigma^4(W)\right)\left(\int K^2(v)dv\right).$$

(2) Under H_1, $P(\widetilde{T}_n > Q_n) \to 1$ *for any non-stochastic sequence* $Q_n = o(nh^{p/2})$.

Theorem 2.1(2) follows from the fact that under H_1, $T_n \xrightarrow{p} E\left[f_W(W)f^2(X)\left(E(D|W)-E(D|X)\right)^2\right] > 0$ and $\widehat{\sigma}_T^2 = O_p(1)$. The proofs of these are straightforward and are thus omitted. Based on Theorem 2.1(1), we can have the following one-sided asymptotic test for H_0: rejecting H_0 at the significance level α_0 if $\widetilde{T}_n > c$ where c is the upper α_0-percentile of the standard normal distribution.

However, Monte Carlo simulations reported in Li (1999) and Lavergne and Vuong (2000) reveal that the normal approximation has substantial finite sample bias. Instead, Lavergne and Vuong (2000) proposed a modified test which is given by

$$J_n = \frac{1}{n(n-1)(n-2)(n-3)} \times \left[n(n-1)^3T_n - n(n-1)(n-2)A_n - 2n(n-1)(n-2)B_n\right],$$

where

$$A_n = \frac{1}{n(n-1)(n-2)h^p h_1^{2d}}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{k\neq i,k\neq j}(D_i-D_k)(D_j-D_k) \times K_1\left(\frac{X_i-X_k}{h_1}\right)K_1\left(\frac{X_j-X_k}{h_1}\right)K\left(\frac{W_i-W_j}{h}\right),$$

and

$$B_n = \frac{1}{n(n-1)(n-2)h^p h_1^{2d}}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{k\neq i,k\neq j}(D_i-D_j)(D_j-D_k) \times K_1\left(\frac{X_i-X_j}{h_1}\right)K_1\left(\frac{X_j-X_k}{h_1}\right)K\left(\frac{W_i-W_j}{h}\right).$$

Indeed, J_n has the same asymptotic distribution as T_n.

3. Monte Carlo studies

To study the size and power properties of the test statistic J_n, the following data generating processes (DGP) is used:

$$Z \sim \mathcal{N}(0,1), \quad \xi \sim \mathcal{N}(0,1), \quad X = \gamma Z + \sqrt{1-\gamma^2}\xi + \eta,$$
$$Y(1) = \rho Z + \gamma_1 X + \epsilon_1, \quad Y(0) = \rho Z + \gamma_0 X + \epsilon_0,$$

and

$$D = I\left\{\frac{\mu}{2}\left(Y(0)+Y(1)\right) + \sqrt{1-\mu^2/2}X > U\right\}, \quad U \sim unif(0,1),$$

Table 1
Estimated sizes of J_n (nominal size $\alpha_0 = 5\%$).

	Model	Empirical rejection probability of J_n with		
		$a = 0.25$		
μ	$\gamma(=\rho)$	$n = 100$	$n = 200$	$n = 400$
	0.0	0.058	0.038	0.051
0	0.4	0.054	0.044	0.050
	0.8	0.061	0.045	0.049
		$a = 0.5$		
μ	$\gamma(=\rho)$	$n = 100$	$n = 200$	$n = 400$
	0.0	0.037	0.041	0.055
0	0.4	0.035	0.040	0.052
	0.8	0.040	0.046	0.055
		$a = 1.0$		
μ	$\gamma(=\rho)$	$n = 100$	$n = 200$	$n = 400$
	0.0	0.031	0.033	0.041
0	0.4	0.026	0.031	0.043
	0.8	0.028	0.037	0.042
		$a = 2.0$		
μ	$\gamma(=\rho)$	$n = 100$	$n = 200$	$n = 400$
	0.0	0.018	0.020	0.038
0	0.4	0.014	0.032	0.036
	0.8	0.018	0.025	0.034

where $I\{\cdot\}$ denotes an indicator function, Z, ξ, ϵ_1, ϵ_0, U and η are mutually independent random variables and $\epsilon_1 \sim \mathcal{N}(0, 0.4^2)$, $\epsilon_0 \sim \mathcal{N}(0, 0.3^2)$, and $\eta \sim \mathcal{N}(0, 0.5^2)$, respectively. We set $\gamma_1 = 2.0$ and $\gamma_0 = 3.0$. The constants, $\gamma \in [0,1]$, $\rho \in [0,1]$ and $\mu \in [0,1]$ vary in different simulation experiments.

It is easy to see that the DGP above satisfies Assumption 2.1(i) in Section 2 no matter what values of μ take. The conditional independence assumption holds only when μ takes value of zero. We use standard normal kernel functions for both $K_1(\cdot)$ and $K(\cdot)$ with the bandwidth chosen by $h_1 = \widehat{\sigma}_X n^{-1/5}$, $h_x = a\cdot\widehat{\sigma}_X n^{-1/4}$ and $h_z = a\cdot\widehat{\sigma}_Z n^{-1/4}$, where $\widehat{\sigma}_X$ and $\widehat{\sigma}_Z$ are the sample standard deviations of $\{X_i\}_{i=1}^n$ and $\{Z_i\}_{i=1}^n$, respectively. To check the sensitivity of the test with respect to different values of the bandwidths, we set $a = 0.25$, 0.5, 1.0 and 2.0, respectively. Finally, the number of replications in each experiment is 2000 for all cases.

The actual sizes of the J_n test based on asymptotic one-sided normal critical values are reported in Table 1. We report the empirical rejection probabilities of the test J_n for different choices of the bandwidth. The test works reasonably well in finite samples in various situations. The actual sizes converge to their nominal sizes as the sample sizes n increases. Particularly, when the sample size increases to 400, the test J_n works very well in most cases. The choice of γ and ρ, which catch the correlation of the auxiliary variable with potential outcomes and covariates, has very little influence on actual sizes. It seems that the bandwidth $a = 0.25$ gives the best small sample performance. The test becomes conservative when a is too large.

Next, we plot the power curves of the J_n test with a nominal size $\alpha_0 = 5\%$ and $a = 1.0$ in Fig. 1 for various cases. Also, the power curves of the J_n test with different values of a can be obtained but the patterns are similar. The test J_n is reasonably powerful in detecting the deviations from the null in all cases even when μ is small. It is not surprising that the powers increase quickly when both the sample size and the value of μ increases. One of interesting facts is that the power performance depends heavily on γ and ρ, the correlation between the auxiliary variable and potential outcomes and covariates. In all cases, when the values of the ρ and γ increase, the powers also increase immensely.

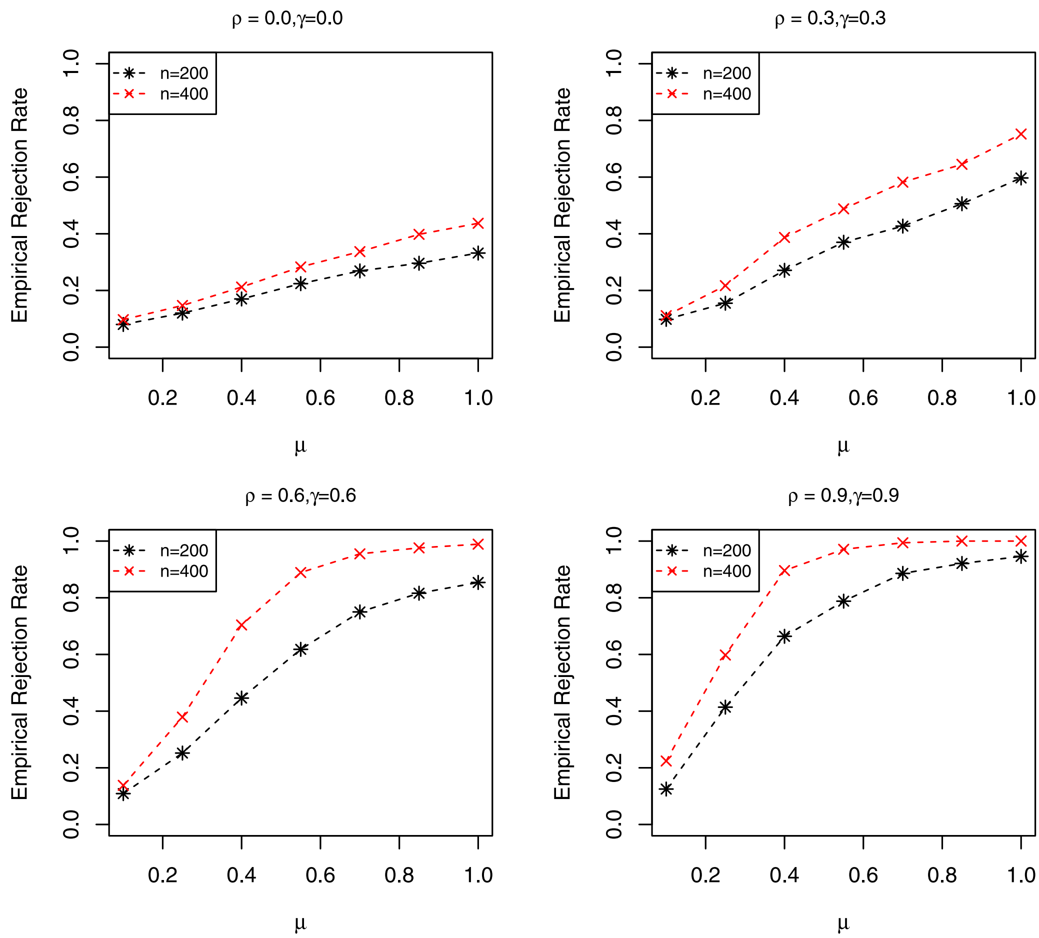

Fig. 1. Power curves for test statistic J_n with nominal size $\alpha_0 = 5\%$ and $a = 1.0$.

Table 2
Descriptive statistics (means and standard deviations).

	Entire sample	By college education	
		Receiving college education	Not receiving college education
Outcome variables:			
Income	15226.400	17923.910	14157.440
	(11678.750)	(11173.620)	(11716.48)
Covariates:			
Experience (Years of working)	10.714	8.456	11.609
	(7.456)	(7.150)	(7.394)
Age	29.153	29.396	29.057
	(7.331)	(7.123)	(7.418)
Gender	0.686	0.611	0.715
	(0.464)	(0.489)	(0.452)
Residence	0.518	0.718	0.438
	(0.500)	(0.451)	(0.496)
Mother's income	10806.760	13561.590	9715.091
	(13682.630)	(13198.820)	(13734.490)
Father's income	15963.790	20974.080	13978.330
	(16157.910)	(22606.050)	(12214.68)

4. Return to college education

Chen et al. (2017) considered the example of return to college education and tested the conditional mean independence using a Kolmogorov–Smirnov test with the assumption of symmetric distributions in error terms. Their test cannot reject the null hypothesis when some relevant covariates are controlled. We revisit the same issue using the same data in Chen et al. (2017). The data come from the China Health and Nutrition Survey (CHNS) of the year of 2004, 2006, and 2009. The data set includes various provinces in China and consists of 525 individuals aged between 18 and 65 with individual characteristics information including gender, residence type, income, education level and family background.

Table 3
Results for the unconfoundedness test.

Auxiliary variables (Z)	Covariates (X)	Test statistic J_n (p-value)
	Experience	0.008
	Experience, Gender	0.228
	Experience, Gender, Residence	0.282
Age	Experience, Gender, Residence, Logarithm of mother's income	0.420
	Experience, Gender, Residence, Logarithm of mother's income, Logarithm of father's income	0.472

In this example, the outcome variable of interest, denoted by Y, is the logarithm of annual income and the treatment variable D is a binary variable which takes a value of 1 for college graduates and 0 otherwise. The covariates X include experience, gender, residence type (urban or rural) and the family background which is represented by parents' income. Table 2 shows that most individual characteristics, except age, are very different for people receiving college education and not receiving. However, the average age between the treated and control group is very similar, 29.39 for the treated group and 29.05 for the control group, which motivate us to consider using age as the proper candidate for the auxiliary variable Z. Since the dimension of the auxiliary variable Z is one in this example, the product normal kernel function is adopted and the bandwidths are chosen as $h_1 = \widehat{\sigma}_X n^{-1/(2(d+2))}$ and $h = 0.5\widehat{\sigma}_X n^{-1/(d+2)}$, where $\widehat{\sigma}_X$ is the sample standard deviation and d is the dimension of X.

Table 3 reports the testing results for the conditional independence using age as the auxiliary variable, conditional on different covariates. Our results are basically similar to those in Chen et al. (2017). The test rejects the null hypothesis of conditional unconfoundedness only in the case that a single covariate experience is controlled, with a p-value 0.008. However, when more conditioning covariates, such as gender, residence type and parents' income, are added, our test cannot reject the null hypothesis any more. The results also show that the p-values increases as more covariates are included in the model, which is in line with the intuition that the conditional unconfoundedness assumption is more likely to hold when more relevant variables are included into the model.

5. Conclusion

This paper proposes an alternative method to test the conditional unconfoundedness assumption, which relies on an auxiliary variable whose potential influence on the treatment decision is fully captured by potential outcomes and observable covariates. In practice, any covariates which are insignificant in the balance check for the treated and control groups can be considered as possible candidates for the auxiliary variable. We also establish the asymptotic properties of our proposed test. The simulation experiments show that our test works very well even in a small sample size. We finally apply our testing method to the example of the return to college education. We find the conditional unconfoundedness can basically hold when the conditioning variables are appropriately chosen.

Finally, we note several possible extensions of the present study. For example, it may be of interest to apply the proposed test in testing conditional independence in studying the partially conditional quantile treatment model as in Cai et al. (2020) with possible irrelevant covariates as in Chen et al. (2019). We leave such extensions as possible future research topics.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.econlet.2020.109320.

References

Breunig, C., 2019. Testing missing at random using instrumental variables. J. Bus. Econom. Statist. 37 (2), 223–234.
Cai, Z., Fang, Y., Lin, M., Tang, S., 2020. Inferences for partially conditional quantile treatment effect model. Working Paper. Department of Economics, University of Kansas.
Chen, T., Ji, Y., Zhou, Y., Zhu, P., 2017. Testing conditional mean independence under symmetry. J. Bus. Econom. Statist. 36 (4), 615–627.
Chen, X., Li, D., Li, Q., Li, Z., 2019. Nonparametric estimation of conditional quantile functions in the presence of irrelevant covariates. J. Econometrics 212 (2), 433–450.
D'Haultfoeuille, X., 2011. On the completeness condition in nonparametric instrumental problems. Econom. Theory 27 (3), 460–471.
Donald, S.G., Hsu, Y.-C., Lieli, R.P., 2014. Testing the unconfoundedness assumption via inverse probability weighted estimators of (l) att. J. Bus. Econom. Statist. 32 (3), 395–415.
Fan, Y., Li, Q., 1996. Consistent model specification tests: omitted variables and semiparametric functional forms. Econometrica 64 (4), 865–890.
Hu, Y., Schennach, S.M., 2008. Instrumental variable treatment of nonclassical measurement error models. Econometrica 76 (1), 195–216.
Lavergne, P., Vuong, Q., 2000. Nonparametric significance testing. Econom. Theory 16 (4), 576–601.
Li, Q., 1999. Consistent model specification tests for time series econometric models. J. Econometrics 92 (1), 101–147.
Mattner, L., 1993. Some incomplete but boundedly complete location families. Ann. Statist. 21 (4), 2158–2162.
Robinson, P.M., 1988. Root-N-consistent semiparametric regression. Econometrica 56 (4), 931–954.
Zhao, J., Shao, J., 2015. Semiparametric pseudo-likelihoods in generalized linear models with nonignorable missing data. J. Amer. Statist. Assoc. 110 (512), 1577–1590.