様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　元 年　 ６ 月 １２ 日現在

機関番号：１２１０２
研究種目：若手研究(B)
研究期間：2017～2018
課題番号：１７Ｋ１３７１３

研究課題名（和文）A study of estimation and inference of the density discontinuity

研究課題名（英文）A study of estimation and inference of the density discontinuity

研究代表者
ＹＵ　ＺＨＥＮＧＦＥＩ（YU, Zhengfei）

筑波大学・人文社会系・助教

研究者番号：４０７７４７５８

交付決定額（研究期間全体）：（直接経費）　　1,200,000円

研究成果の概要（和文）：This research proposes optimal bandwidths for estimating and testing functions of density discontinuity. When applied to the manipulation test, the optimal bandwidth has a simple analytical form. Monte Carlo simulations demonstrate advantageous finite sample performances of the proposed methods.

研究成果の学術的意義や社会的意義
The proposed optimal bandwidth selectors can be used to improve the performance of testing for the validity of regression discontinuity designs. They can also be used to improve the estimation of structural parameters in economic models related to discontinuous policy changes.

研究成果の概要（英文）：This research proposes optimal bandwidths for estimating and testing functions of density discontinuity, which can be used to check the validity of the regression discontinuity design, or to estimate agents' response to the policy change that is related to certain functions of densities around a threshold. For the estimation problem, the proposed method allows different bandwidths on both sides of the threshold, and thus generates more precise estimate for the parameter of interest. For the testing problem, the optimal bandwidth minimizes the coverage error. When applied to the manipulation test for regression discontinuity designs, the proposed method yields the optimal bandwidth with an analytical form, which cannot be obtained by existing methods of density discontinuity inference. Monte Carlo simulations demonstrate advantageous finite sample size and power performances of the proposed methods.

研究分野：計量経済学

キーワード：Density discontinuity　Optimal bandwidth　Empirical likelihood　Manipulation test　Regression discontinuity

１．研究開始当初の背景

The estimation and inference of discontinuity in density play important roles in economic researches. To check the validity of regression discontinuity designs (RDD), researchers are advised to conduct the manipulation test for the continuity of the density function of the running variable at a threshold. When discontinuity is detected and thus the validity of the standard RDD approach is in doubt, density ratio at the threshold can be used to identify the degree of manipulation and then to develop bounds for the local average treatment effect (Gerard et al, 2015). Moreover, the magnitude of density discontinuity or some functions of densities at some threshold(s) also reveal agents' response to policy changes and are closely related to some structural parameters. For example, the ratio of the wage density at the left and right hand sides of the minimum wage corresponds to the effect of a minimum wage policy on the wage distribution and unemployment (Jales, 2018). This project focuses on the optimal bandwidth choice for the estimation and inference of discontinuity in density.

２．研究の目的

This project provides explicit formulas of the asymptotic first-order mean square error optimal bandwidths for estimating non-linear functions (for example, the ratio) of the discontinuity in the density. It allows different bandwidths at the left and right hand sides of the threshold. For the testing problem where the parameter of interest takes the form of a possibly nonlinear function of densities at both sides of the threshold, this project proposes a coverage error optimal bandwidth selector, which is defined to be the minimizer of this leading coverage error term. This project also describes how to implement the optimal bandwidths in practice for estimation and testing problems.

３．研究の方法

For the estimation problem, the optimal simultaneous bandwidths (allowed to differ between both sides of the threshold) are obtained by minimizing the mean square error of parameter of interest which is a function of densities at both sides. As comparison, the project also computes the optimal bandwidth restricted to be the same at both sides of the threshold (single bandwidth), and the bandwidths optimal for separately estimating the densities at both sides, ignoring the form of the parameter of interest (independent bandwidths). For the inference problem, the densities at both sides of the threshold are estimated by empirical likelihood approach where the moment conditions follow from the minimum contrast problem instead of local linear regression or local likelihood estimation. Then the optimal bandwidths are obtained by minimizing the second order coverage error of the empirical likelihood ratio test. The finite sample performances of the optimal bandwidth selectors are evaluated through Monte Carlo simulations.

４．研究成果

(1). For the estimation problem, the optimal simultaneous bandwidths proposed by this project perform better than the single bandwidth and the independent bandwidths, in terms of smaller biases and mean square errors, see Table I as an example.

Table I. Bias and root mean square error (RMSE) for estimating density ratio at the threshold using three bandwidth selectors, Simultaneous optimal bandwidth is the recommended selector.

	Simultaneous		Independent		Single	
Sample size	Bias	RMSE	Bias	RMSE	Bias	RMSE
500	0.041	0.145	0.045	0.151	0.040	0.138
2000	0.018	0.074	0.024	0.083	0.023	0.080
5000	0.010	0.047	0.017	0.071	0.014	0.057

(2). For the inference problem, the projects provides a coverage error optimal bandwidth for the general setting of testing a possibly nonlinear function of densities at both sides of the threshold. The optimal bandwidth converges to zero at a cubic rate, which is faster than the optimal rate for the estimation problem. The empirical likelihood test proposed by the project is also Bartlett correctable.

(3). When applied to the manipulation test for RDD, where the parameter of interest is the difference between densities around the threshold, the optimal bandwidth for the proposed testing approach has an explicit form and is easy to compute in practice. These nice properties are not shared by existing methods of density discontinuity inference such as Otsu et al (2013) and Cattaneo et al (2017), and thus are an advantages of the proposed testing approach.

(4). Monte Carlo studies demonstrate that the proposed testing approach using the coverage error optimal bandwidth enjoys better size and power performance than existing methods, see Table II. In particular, the proposed approach outperforms Cattaneo's local linear regression approach in both size and power.

(5). Proposed bandwidth selectors for the estimation and inference work well when applied to the "PROGRESA" data, a conditional cash transfer program to the poor households.

Table II. The rejection frequencies for different manipulating tests for six experiments, nominal size=0.05

Experiment	Sample size	ELCE	OXM	CJM
Size I	1000	0.0504	0.0516	0.0324
	2000	0.0498	0.0498	0.0344
Size II	1000	0.0562	0.0850	0.0480
	2000	0.0604	0.1106	0.0494
Size III	1000	0.0634	0.0616	0.0440
	2000	0.0598	0.0602	0.0526
Power I	1000	0.1504	0.0070	0.1330
(diff=0.05)	2000	0.2220	0.2060	0.1912
Power II	1000	0.3132	0.3340	0.2534
(diff=.075)	2000	0.4880	0.6900	0.3918
Power III	1000	0.5364	0.6480	0.4266
(diff=.100)	2000	0.7660	0.9210	0.6380

Note: ELCE is the proposed testing approach, OXM denotes the existing empirical likelihood based test by Ostu, Xu and Matsushita (2013), and CJM denotes the existing local linear regression based test by Cattaneo, Jansson and Ma (2017).

＜引用文献＞

① Cattaneo, M. D., M. Jansson, and X. Ma (2017). Simple local polynomial density estimators. Working paper, University of Michigan.

② Gerard, F., Rokkanen, M., Rothe, C., (2015). Partial identification in regression discontinuity designs with manipulated running variables, working paper, Columbia University.

③ Jales, H. (2018). Estimating the effects of the minimum wage in a developing country: A density discontinuity design approach. *Journal of Applied Econometrics* 33, 29–51.

④ Otsu, T., K.-L. Xu, and Y. Matsushita (2013). Estimation and inference of discontinuity in density. *Journal of Business & Economic Statistics* 31, 507–524.

５．主な発表論文等

〔雑誌論文〕（計　2　件）

① Hugo Jales, Jun Ma and Zhengfei Yu (2017). "Optimal Bandwidth Selection for Local Linear Estimation of Discontinuity in Density,'' *Economics Letters,* 153, 23-27. 査読有

DOI: https://doi.org/10.1016/j.econlet.2017.01.024

② Jun Ma, Hugo Jales, and Zhengfei Yu (2019) "Minimum Contrast Empirical Likelihood Inference of Discontinuity in Density", accepted for publication by *Journal of Business & Economics Statistics*. 査読有

DOI: https://doi.org/10.1080/07350015.2019.1617155

〔学会発表〕（計　4　件）

① Zhengfei Yu " Minimum Contrast Empirical Likelihood Manipulation Testing for Regression Discontinuity Design'', International Applied Econometrics Association (IAEA) 2017 Annual Conference

② Zhengfei Yu " Minimum Contrast Empirical Likelihood Manipulation Testing for Regression Discontinuity Design'', Asian Meeting of the Econometric Society 2017

③ Zhengfei Yu "Simple Semiparametric Estimation of Ordered Response Models", Workshop on Advances in Econometrics 2018.

④ Zhengfei Yu "Simple Semiparametric Estimation of Ordered Response Models", Invited talk at Institute of Economics, Academia Sinica 2018.

６．研究組織