

The Development of Regression Discontinuity Design and its Application in the Financial Field

Danxue Luo

School of Finance
Yunnan University of Finance and Economics
Kunming, China
1034159312@qq.com

Wei Zhou*

School of Finance
Yunnan University of Finance and Economics
Kunming, China
zw453@163.com

Abstract—Regression discontinuity design was proposed by Thistlethwaite and Campbell in 1960, and it was not applied to economic research until the end of the 20th century. With the gradual deepening of scholars' research on regression discontinuity design, its related theory is gradually perfected, and its application norms are becoming more mature. This makes regression discontinuity design gradually become a mainstream empirical research method. But compared to its application in theoretical economics, the literature on finance is still scarce. Therefore, this article intends to briefly introduce the basic regression discontinuity design to lay a certain theoretical foundation for later scholars to apply it to the financial field. Then sort out the literature and present to readers the current application of regression discontinuity design in the financial field so that scholars can grasp the current research hotspot.

Keywords: regression discontinuity design, financial sector, literature review

I. INTRODUCTION

Regression discontinuity design was proposed by Thistlethwaite and Campbell in 1960, and it was not applied to economic research until the end of the 20th century. In recent years, regression discontinuity design has become a mainstream empirical research method in the field of economics. First, compared with other non-experimental methods, the assumptions required for regression discontinuity design are more "mild"(Hahn et al., 2001; Huang et al., 20-19). Second, causal reasoning for regression discontinuity design is more credible than "natural experiments"(Lee & Lemieux, 2010; Qu et al., 2019). The current research around the design of regression discontinuity can be roughly divided into three directions. The first is to focus on the basic theory and applied research of regression discontinuity design, such as Imbens and Lemieux (2008), Van der Klaauw (2008), Lee and Lemieux (2010), Skovron and Titunik (2017), Hausman and Repson (2018). The second is to study the technical details of empirical analysis of regression discontinuity design, such as the selection of parameter estimation and non-parametric estimation, the selection of polynomials in parameter estimation, the determination of optimal bandwidth, and the reduction of error in evaluation results. The third is to modify and optimize the regression discontinuity design based on basics, such as regression discontinuity designs with assignment variable (Papay et al., 2011; Rao et al., 2019), quantile regression discontinuity design (Frandsen et al., 2012), regression kink designs(Card and Giuliano, 2016), regression designs with multiple cutoffs (Cattaneo et al., 2016), the regression discontinuity design for discrete configuration

variables(Kolesar and Rothe, 2018; Zhou Et al., 2019a,b,c,d), etc. With the gradual deepening of the research on regression discontinuity, its related theories are gradually perfected, and its application procedures are becoming more mature. This also makes this method more widely applicable to other fields. At present, some scholars have begun to apply regression discontinuity design to the financial field. But compared to its application in theoretical economics, the literature on finance is still scarce. Therefore, this article intends to briefly introduce the basic regression discontinuity design to lay a certain theoretical foundation for later scholars to apply it to the financial field. Then sort out the literature and present to readers the current application of regression discontinuity design in the financial field so that scholars can grasp the current research hotspot.

II. BASIC REGRESSION REGRESSION DISCONTINUITY

The basic idea of regression discontinuity is to use the sample smaller than the "cutoff" as the control group and the sample larger than the "cutoff" as the experimental group. The causal relationship between intervention variables and outcome variables was studied by comparing the differences between the two groups of samples.

Whether the sample is larger than the "cutoff" is random, so there seems to be random grouping near the cutoff. Because of this feature, regression discontinuity design can effectively identify and evaluate the causal effects of policies or projects (Imbens & Lemieux, 2008; Angrist & Pischke, 2009; Lee & Lemieux, 2010) The basic formula for regression discontinuity design is as follows:

$$D_i = \begin{cases} 1, & x_i \geq c \\ 0, & x_i < c \end{cases} \quad (1)$$

$$y_i = \alpha + \delta D_i + \beta x_i + \varepsilon_i \quad (2)$$

Where y_i is the outcome variable, x_i is the intervention variable (grouping variable), D_i is the processing variable, ε_i is the error term, α is the fixed effect, c is the constant, δ and β are the parameters to be estimated.

Equation (1) is a grouping rule, and equation (2) is an estimation equation for processing effects. The processing variable D_i is a function of x_i . c represents a cutoff. If the treatment effect of the experimental group is greater than zero,

*Corresponding author

that is $\delta > 0$, the linear relationship between y_i and x_i will have a break point of jumping up at $x = c$; If the treatment effect of the experimental group is less than zero, that is $\delta < 0$, the linear relationship between y_i and x_i will have a break point of jumping down at $x = c$. As shown in figure 1. Because the sample is not affected by other than x_i in the vicinity of $x = c$, the reason why $E(y_i | x = c)$ will jump is only the processing effect of D_i . Hence, this jump can be seen as D_i causal effects at cutoff c .

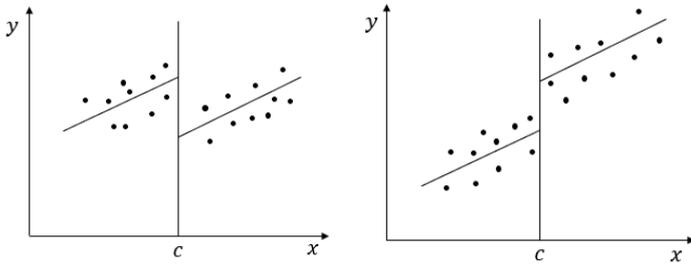


Figure 1. Schematic of basic regression discontinuity

There are two types of basic regression discontinuity designs. One is sharp regression discontinuity; the other is fuzzy regression discontinuity (Trochim, 1984). The sharp regression discontinuity design is applied to the probability that the research object gets processed from 0 to 1. For example, Cerqua and Pellegrini (2014) estimated the causal effect of capital subsidies on companies in backward regions; Falato and Liang (2016) studied the causal relationship between claims and employment risk. Fuzzy regression discontinuity design is applied to the probability that the research object will be processed from a to b ($0 < a < b < 1$). For example, Farooq and Aktaruzzaman (2016) used data from 69 villages in Bangladesh to estimate the impact of participation in microfinance programs on household spending; Zhang et al. (2017) estimated the impact of increasing regional electricity price policies on urban residents' electricity consumption in Guangdong Province, China.

Both sharp regression discontinuity and fuzzy regression discontinuity must meet certain pre-determined conditions. In the specific application, it is mainly summarized as three steps. The first step is to check if the configuration variables are manipulated. The second step checks whether the configuration variable jumps at the cutoff. The third step is to report empirical results in different forms at the same time to ensure robustness. The article needs to report the results of the local linear regression of the triangular kernel and the rectangular kernel separately; the results under different bandwidths; including and without covariates

III. APPLICATION OF RDD IN THE FINANCIAL FIELD

Regression discontinuity design is mainly used for the study of cross-section data. The regression discontinuity design with time as the running variable has been used in empirical research, which is similar to event research (Hausman & Rapson, 2018). Examples of time series as running variables are Busse et al., 2006; Davis & Kahn, 2010; Auffhammer & Kellogg, 2011; Chen & Whalley, 2012; Gallego et al., 2013; Anderson, 2014. This provides a reference for the application

of regression discontinuity design to the research of time series in the financial field. This paper reviews the related research in recent years and finds that the regression discontinuity design is mainly applied to the following aspects of the financial field.

The first is around the internal activities of the enterprise. Examples include corporate performance, corporate tax avoidance, corporate finance, and corporate innovation incentives. Blasio et al. (2011) used regression discontinuity design to study whether a subsidy scheme for the development of innovative applications in Italy is effective in stimulating innovation investment in companies. They compare companies before and after the announcement of an unexpected shortage of funds. It was found that this solution could not effectively stimulate the innovation investment of enterprises. Cuñat et al. (2015) studied the voting of SOP proposals supported by shareholders using regression discontinuity design and found SoPCan greatly increase the company's market value and improve its long-term profitability. Krishnan et al. (2014) analyzed the impact of the expansion of financing channels on the total factor productivity (TFP) of enterprises by using regression discontinuity design. They found that as companies' TFP increased, their countries implemented these deregulations. For companies with limited funds, the TFP growth after deregulation is significantly greater. At the same time, greater financing channels allow companies with limited funds to invest in productive projects. Bronzini and Piselli (2016) evaluated the impact of research and development subsidy programs implemented in northern Italy on the innovation of beneficiary companies based on regression discontinuity design theory. They found that the R & D subsidy program increased the possibility of patent applications, especially for smaller companies. Focke et al. (2016) analyzed the relationship between corporate reputation and executive compensation based on three methods: matched sample analysis, DID, and regression discontinuity design. Research showed that CEOs of well-known companies will earn less. Bird and Karolyi (2017) used regression discontinuity design to examine the effect of institutional ownership on tax avoidance. Studies have found that the positive impact of institutional ownership can effectively reduce the tax rate (ETR). This is because they can lead to increased use of international tax programs and the use of tax subsidiaries. In addition, they found that the highest ETR companies experienced the largest reductions, while the lower ETR companies increased, consistent with institutional ownership driving companies toward a common level of tax avoidance. However, Khan et al. (2017) added a Russell index based on the regression discontinuity design to isolate the provincial impact of institutional ownership. This can more clearly identify the processing effect. The research results show the impact of increased ownership concentration on tax avoidance.

The second is to focus on the external behavior of enterprises. Including acquisitions between companies and corporate social responsibility. Becht et al. (2015) studied the relationship between mandatory shareholder voting and acquisition in UK based on DID and regression discontinuity design. The results of the study indicate that mandatory shareholder voting imposes constraints on the acquirer's CEO. Flammer (2015) uses the method of regression discontinuity design to explore the impact of shareholder suggestions on corporate social responsibility financial performance. He used research on corporate social responsibility proposals that passed or failed with only a slight advantage. It was found the adoption of corporate social responsibility recommendations

resulted in positive announcement returns and excellent accounting performance. Cao et al. (2016) also used regression discontinuous design to investigate the response of companies to their corporate social responsibility (CSR) adoption. They found that the peers of the voting company that passed the CSR proposal had lower announcement returns and a higher CSR score in the second year, especially in industries with more competitive pressure and more transparent information.

The third is the influence of other external factors on the company, such as social media, government measures or systems, and industry rules. Boone et al. (2015) used an important difference in US regulatory treatment to investigate competition between traditional stock exchanges and new dark trading venues. Through the method of regression discontinuity design, they found that propagation constraints can significantly weaken the competition between the two stock exchanges. Decramer and Vanormelingen (2016) used regression discontinuity design to analyze the effect of Flemish SME investment subsidy program from 2004 to 2009. Because this investment subsidy program only invests in the top rankings, that is, companies with higher scores provide subsidies. As a result, there was a sharp interruption in the distribution of subsidies. This provides the basic conditions for the article to use regression discontinuity design. The results of the study show that subsidized companies have positive effects on investment, employment, output and productivity at the enterprise level, but only for small enterprises. Bird and Karolyi (2016) used a regression discontinuity design to examine the impact of institutional ownership on corporate information disclosure policies. At the same time, they used a new dataset containing data from 1996 to 2006. Finally, they found that the positive impact of institutional ownership rebuilding around the Russell Index increased the quantity, form, and quality of disclosures. Specifically, compared with the companies at the bottom of the Russell 1000 index, the corporate ownership at the top of the Russell 2000 index increased by 9.8%, 8-K files increased by 4.7%, and embedded graphics increased by 21.3%. Kaniel and Parham (2017) used regression discontinuity design to establish a causal relationship between media attention and consumer investment behavior. Their findings indicate a 31% local average increase in quarterly capital flows into mutual funds mentioned in a prominent Wall Street Journal "Category Kings" ranking list, compared to those funds which just missed making the list. Because of the well-documented performance-to-flow relationship, this increase in traffic is approximately seven times the amount of additional traffic.

By combing the above literature. We can find that regression discontinuity design is mainly used to evaluate the effect of a cutoff in a policy or measure. There are also a few scholars who apply regression discontinuity design to multiple discontinuity measures of the same project (Xi & Liang, 2015) and spatial regression discontinuity (David et al., 2020). However, few scholars have applied regression discontinuity design to the impact of financial events on the market. In fact, random events in the financial markets (referred to as sudden outbursts or outbreaks), after being fermented by public opinion, will cause the "butterfly effect" if they are not careful, which will cause major stock market fluctuations and social impacts.

IV. CONCLUSION

Regression discontinuity design is the quasi-random experiment method that is closest to the random experiment

method. Causal inference in policy evaluation has great advantages. Causal inference is clear and easy to test. And with the deepening of research, the application of regression discontinuity design is gradually improved, and the details of its operation in the research process are continuously refined and specific. This provided a good theoretical basis and practice norms for later scholars to study regression discontinuity design. In addition, with the skillful use of regression discontinuity design with time as the running variable. This provides a reference value for scholars to regression discontinuity design and time series data in the field of finance, while broadening the scope of research. At present, the application of regression discontinuity design in the financial field mostly revolves around the evaluation of policies or projects, and there is less research on financial events.

REFERENCES

- [1] A. Bird, A. Stephen. K. "Governance and taxes: evidence from regression discontinuity (Retracted). *The Accounting Review*, 2017, Vol. 92, pp. 29-50.
- [2] A. Bird, S. A. Karolyi, "Do institutional investors demand public disclosure?" *Review of Financial Studies*, 2016, Vol. 29(12), pp. 3245-3277.
- [3] A. Cerqua, G. Pellegrini, "Do subsidies to private capital boost firms growth? A multiple regression discontinuity design approach," *Journal of Public Economics*, 2014, Vol. 109, pp. 114-126.
- [4] A. Falato, N. Liang, "Do creditor rights increase employment risk? Evidence from loan covenants," *The Journal of Finance*, 2016, Vol. 71, pp. 2545-2590.
- [5] A. L. Boone, J. T. White, "The effect of institutional ownership on firm transparency and information production," *Journal of Financial Economics*, 2015, Vol. 115, pp. 330-348.
- [6] B. Nyhan, C. Akovron, R. Titunik, "Differential registration bias in voter file data: A sensitivity analysis approach," *American Journal of Political Science*, 2017, Vol. 61, pp. 744-760.
- [7] B. R. Frandsen, M. Frolich, B. Melly, "Quantile treatment effects in the regression discontinuity design," *Journal of Econometrics*, 2012, Vol. 168, pp. 382-395.
- [8] C. Flammer, "Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach," *SSRN Electronic Journal*, 2012, Vol. 61(11), pp. 2549-2568.
- [9] C. Hausman, S. David, Rapson, "Regression Discontinuity in time: Considerations for empirical applications," *Annual Review of Resource Economics*, 2018, Vol. 10, pp. 533-552.
- [10] C. J. Rao, H. Lin, M. Liu, "Design of comprehensive evaluation index system for P2P credit risk of 'Three Rural' borrowers," *Soft Computing*, 2019, DOI: <https://doi.org/10.1007/s00500-019-04613-z>.
- [11] D. Card, L. Giuliano, "Can tracking raise the test scores of high-ability minority students?" *The American Economic Review*, 2016, Vol. 106, pp.2783-2816.
- [12] D. L. Thistlethwaite, D. T. Campbell, "Regression-discontinuity analysis: An alternative to the ex post facto experiment," *Journal of Educational Psychology*, 1960, Vol. 51, pp. 309 -317.
- [13] D. S. Lee, T. Lemieux, "Regression discontinuity designs in economics," *Journal of Economic Literature*, 2010, Vol. 48, pp. 281-355.
- [14] Der Klaauw W V. *Regression-discontinuity analysis: A survey of recent developments in economics*[J]. *Labour*, 2008, 22(2): 219-245."
- [15] F. Focke, E. Maug, A. Niessenruezi, "The impact of firm prestige on executive compensation," *Journal of Financial Economics*, 2016, Vol. 123, pp. 313-336.
- [16] F. Gallego, J. P. Montero, C. Salas, "The effect of transport policies on car use: Evidence from Latin American cities," *Journal of Public Economics*, 2013, Vol. 107, pp. 47-62.
- [17] G. D. Blasio, D. Fantino, G. Pellegrini, "Evaluating the impact of innovation incentives: Evidence from an unexpected shortage of funds," *Industrial and Corporate Change*, 2011, Vol. 24, pp. 1285-1314.
- [18] G. Imbens, T. Lemieux, "Special issue editors' introduction: The regression discontinuity design-Theory and applications," *Journal of Econometrics*, 2008, Vol.142, pp. 611-614.

- [19] J. Cao, H. Liang, X. Zhan, "Peer effects of corporate social responsibility," Social Science Electronic Publishing, 2016.
- [20] J. D. Angrist, J. S. Pischke, "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics," *Journal of Economic Perspectives*, 2010, Vol. 24, pp. 3-30.
- [21] J. Hahn, P. Todd, V. D. K. Wilbert, "Identification and estimation of treatment effects with a regression-discontinuity design," *Econometrica*, 2001, Vol. 69, pp. 201-209.
- [22] J. P. Papay, J. B. Willett, R. J. Murnane, "Extending the regression-discontinuity approach to multiple assignment variables," *Journal of Econometrics*, 2011, Vol. 161, pp. 203-207.
- [23] K. Aktaruzzaman, O. Farooq, "Impact of microcredit on borrowers' expenditures: A fuzzy regression discontinuity design approach," *Applied Economics*, 2016, Vol. 48, pp. 3685-3694.
- [24] K. Krishnan, D. K. Nandy, M. Puri, "Does Financing Spur Small Business Productivity? Evidence from a Natural Experiment," *SSRN Electronic Journal*, 2013, Vol. 28, pp. 1768-1809.
- [25] L. W. Davis, "The effect of driving restrictions on air quality in Mexico city," *Journal of Political Economy*, 2008, Vol. 116, pp. 38-81.
- [26] M. Auffhammer, R. Kellogg, "Clearing the air? the effects of gasoline content regulation on air quality," *American Economic Review*, 2011, Vol. 101, pp. 2687-2722.
- [27] M. Becht, A. Polo, S. Rossi, "Does Mandatory Shareholder Voting Prevent Bad Acquisitions?" *Social Science Electronic Publishing*, 2016, Vol. 29, pp. 3035-3067
- [28] M. Busse, R. J. Silva, F. Zettelmeyer, "\$1,000 cash back: The pass-through of auto manufacturer promotions," *American Economic Review*, 2006, Vol. 96, pp. 1253-1270.
- [29] M. D. Cattaneo, L. Keele, R. Titiunik, "Interpreting regression discontinuity designs with multiple cutoffs," *The Journal of Politics*, 2016, Vol. 78, pp. 1229-1248.
- [30] M. Khan, S. Srinivasan, L. Tan, "Institutional ownership and corporate tax avoidance: New evidence," *Social Science Electronic Publishing*, 2017, Vol. 92, pp. 101-122.
- [31] M. Kolesar, C. Rothe, "Inference in regression discontinuity designs with a discrete running variable," *The American Economic Review*, 2018, Vol. 108, pp. 2277-2304.
- [32] M. L. Anderson, "Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion," *American Economic Review*, 2014, Vol. 104, pp. 2763-2796.
- [33] R. Bronzini, P. Piselli, "The Impact of R&D Subsidies on Firm Innovation," *SSRN Electronic Journal*, 2016.
- [34] R. Kaniel, R. Parham, "WSJ Category kings - the impact of media attention on consumer and mutual fund investment decisions," *Journal of Financial Economics*, 2017, Vol. 123, pp. 337-356.
- [35] R. P. Huang, S. J. Qu, X. G. Yang, Z. M. Liu, "Multi-stage distributionally robust optimization with risk aversion," *Journal of Industrial and Management Optimization*, 2019, DOI: 10.3934/jimo.2019109.
- [36] S. Decramer, S. Vanormeling, "The effectiveness of investment subsidies: evidence from a regression discontinuity design," *Small Business Economics*, 2016, Vol. 47, pp. 1007-1032.
- [37] S. J. Qu, Y. Y. Zhou, Y. L. Zhang, M. I. M. Wahab, G. Zhang, Y. Y. Ye, "Optimal strategy for a green supply chain considering shipping policy and default risk," *Computers and Industrial Engineering*, 2019, Vol. 131, pp. 172-186.
- [38] V. Cuñat, M. Gine, M. Guadalupe, "Say pays! Shareholder voice and firm performance." *Social Science Electronic Publishing*,
- [39] W. Zhou, R. T. Gu, L. Shuai, "Penetrating the real performance of SSE STAR enterprises: A double-market investigation," *Finance Research Letters*, 2019a, DOI: 10.1016/j.frl.2019.101346.
- [40] W. Zhou, S. Lu, J. Chen, Q. J. Chen, S. Meng, "E2E double-process efficiency analysis from the perspectives of energy consumption and environmental treatment," *Emerging Markets Finance and Trade*, 2019b, DOI: 10.1080/1540496X.2019.1697673.
- [41] W. Zhou, Z. S. Xu, "Envelopment analysis, preference fusion, and membership improvement of intuitionistic fuzzy numbers," *IEEE Transactions on Fuzzy Systems*, 2019c.
- [42] W. Zhou, W. Rao, S. Lu, "Market stability analysis after the circuit breaker for the CSI 300 energy index," *Finance Research Letters*, 2019d, DOI: 10.1016/j.frl.2019.101348.
- [43] Y. Chen, A. Whalley, "Green infrastructure: The effects of urban rail transit on air quality," *American Economic Journal: Economic Policy*, 2012, Vol. 4, pp. 58-97.
- [44] Z. Zhang, W. Cai, X. Feng, "How do urban households in China respond to increasing block pricing in electricity? Evidence from a fuzzy regression discontinuity approach," *Energy Policy*, 2017, Vol. 105, pp. 161-172.