Journal of Economics and Public Finance
ISSN 2377-1038 (Print) ISSN 2377-1046 (Online)
Vol. 10, No. 4, 2024
www.scholink.org/ojs/index.php/jepf

Original Paper

Causal Inference in Transport Policy Evaluation: Principle, Method and Application of Regression Discontinuity Design

Siyi Xiao[1], Anli Leng[2], Xiaoli Lin[1], Zeming Cheng[1] & Bin Tang[1*]

[1] Center for Experimental Economics in Education, Shaanxi Normal University, Xi'an, China

[2] School of Political Science and Public Administration, Institute of Governance, Shandong University, Qingdao, China

[*] Corresponding Author: Bin Tang, Center for Experimental Economics in Education. Shaanxi Normal University, Xi'an, Shaanxi, China. E-mail: tangbin198994@163.com

Received: October 29, 2024 Accepted: November 24, 2024 Online Published: December 2, 2024
doi:10.22158/jepf.v10n4p249 URL: http://dx.doi.org/10.22158/jepf.v10n4p249

Abstract

The methods of causal inference have swept the social sciences in recent years, and regression discontinuity design is one of the typical methods in causal inference. In the field of estimation and impact of traffic policy in developing countries, the research on cause-and-effect relationships is extremely limited. This paper introduces the principles, methods of regression discontinuity design, research application in the field of transport policy, and compares regression discontinuity design with methods of causal inference on the advantages and disadvantages in the field of transport policy. The purpose of this paper is to introduce the regression discontinuity design method into the field of transport policy evaluation, to look forward to its application prospects and to provide basic support for the application of regression discontinuity design in the field of developing countries' transport policy.

Keywords

causal inference, regression discontinuity design, transport policy evaluation, impact evaluation

1. Introduction

In the current academic research of social science, the formulation and implementation of policies rely more on evidence based on causal inference. The trend and methods of causal inference have swept through almost all fields of social science research in the past 20 years and gradually replaced traditional correlation analysis methods (such as OLS regression, etc.). In the field of policy formulation, implementation, and evaluation, causality inference is highly respected, mainly due to the following reasons: First of all, under the situation of global economic slowdown and limited financial resource allocation, how to optimize the allocation of resources and choose the most efficient allocation of resources is essential, and the formulation and implementation of different projects or policies are more likely to be guided by objective empirical evidence; secondly, with the development of the democratic process, citizens are paying increasingly attention to the reliability, validity and effect of public policies. If there is no definite evidence to support or evaluate the impact of policies, it is difficult to win the trust of the people, thus public policies are questioned. In addition, with rigorous and in-depth academic research, economists and sociologists are committed to finding and trying out the most reliable and effective impact assessment methods. The change in the research paradigm is called "credibility revolution" by Angrist and Pischke (2010), and people's vision is also introduced into the research of causality inference.

Causal inference method has been widely used in policy making and evaluation in various fields all over the world. For example, in the field of education, scholars have been studying for many years whether a voucher can help to increase the enrollment rate (Angrist, Bettinger, Bloom, et al., 2002; Hsieh & Urquiola, 2006) and how class size affects students' academic performance (Hanushek, 1999); in the field of health economics, causal inference methods are widely used, such as whether vaccination helps to reduce the risk of disease (Bor, Fox, Rosen, et al., 2017), whether retirement policies have positive or negative effects on the health of the elderly (Müller & Shaikh, 2018). In the field of management, a series of old topics and new topics, such as whether the performance evaluation system significantly improves the company's performance and employee performance (Moers, 2005), whether vocational training effectively promotes the company's development (Sels, De Winne, Delmotte, et al., 2006; Aragón, Jiménez, & Valle, 2014) are all increasingly inclined to use causal inference. Identification methods of causal inference mainly include experimental methods and quasi-experimental methods. Specifically, common causal inference methods include Randomized control trial (RCT), Fixed effect (FE), double Difference in difference (DID), Matching, Instrumental variable (IV), Regression Discontinuity Design (RDD), Synthetic control methods, etc.

In the field of traffic policy evaluation in developing countries, there is a lack of inference research based on causality, such as whether road restrictions have really played a role in easing traffic congestion, and whether clean energy emission standards for motor vehicles have played a role in curbing air pollution. Scientific evaluation methods are urgently needed for diagnosis. This paper introduces the basic principle of Regression Discontinuity Design method, investigates the application of Regression Discontinuity Design method in the field of traffic policy evaluation, compares the advantages and disadvantages of RDD and other causal inference methods, and looks forward to the application prospect of RDD method in traffic policy evaluation.

2. The Principal and Method of Regression Discontinuity Design

2.1 The Principal of Regression Discontinuity Design

Regression Discontinuity Design (RDD) was first proposed by Thistlethwait and Campell (1960) when exploring the impact of scholarships on student performance. Since the scholarship is based on the students' past achievements, if the students' achievements reach a certain standard, they can obtain the scholarship. If the student's score is lower than the standard, they will not be able to obtain a scholarship. Therefore, the achievement threshold in the scholarship appraisal is used as a breakpoint to identify the impact of the scholarship on students' future studies. At first, this method was not widely noticed, then Rubin (1977) introduced a design idea into statistics, Berk and Rauma (1983) used a logistic model to expand the model to binary variables, and it was not until Hahn et al. (2001) made a theoretical and systematic discussion on RDD that it regained its vitality, and then a large number of theoretical evaluation and empirical application documents based on RDD emerged.

The basic idea of RDD is that an intervention variable (*Di*) depends entirely on an allocation variable (*Xi*) assuming that the relationship between the outcome variable (*Yi*) and the allocation variable (*Xi*) is continuous, other variables (*Z*) that may affect the outcome are also continuous at the breakpoint. As shown in Fig. 1, there is no jump of the outcome variable at the breakpoint. As shown in Figure 2, there is a jump of the outcome variable at the breakpoint, then the jump of the outcome variable (*Yi*) at the breakpoint can be interpreted as the impact of the cause variable or intervention variable (*Di*) (Zhao, 2017).

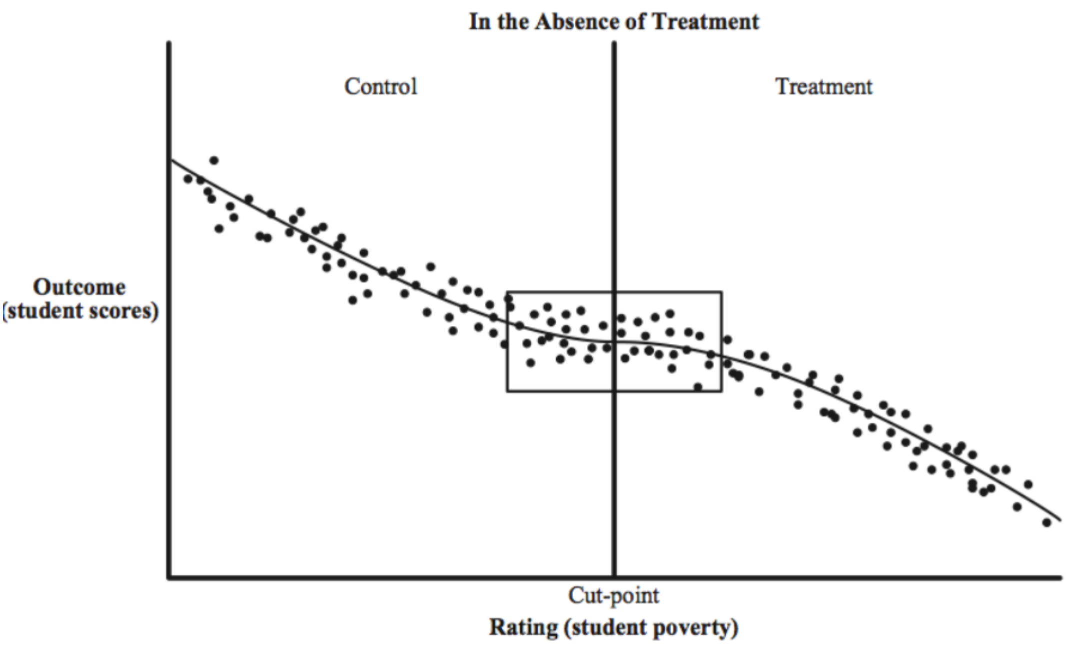

Figure 1. On Outcome Variable Jumping at Cut Point

Source: Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of economics and statistics, 86*(1), 226-244.

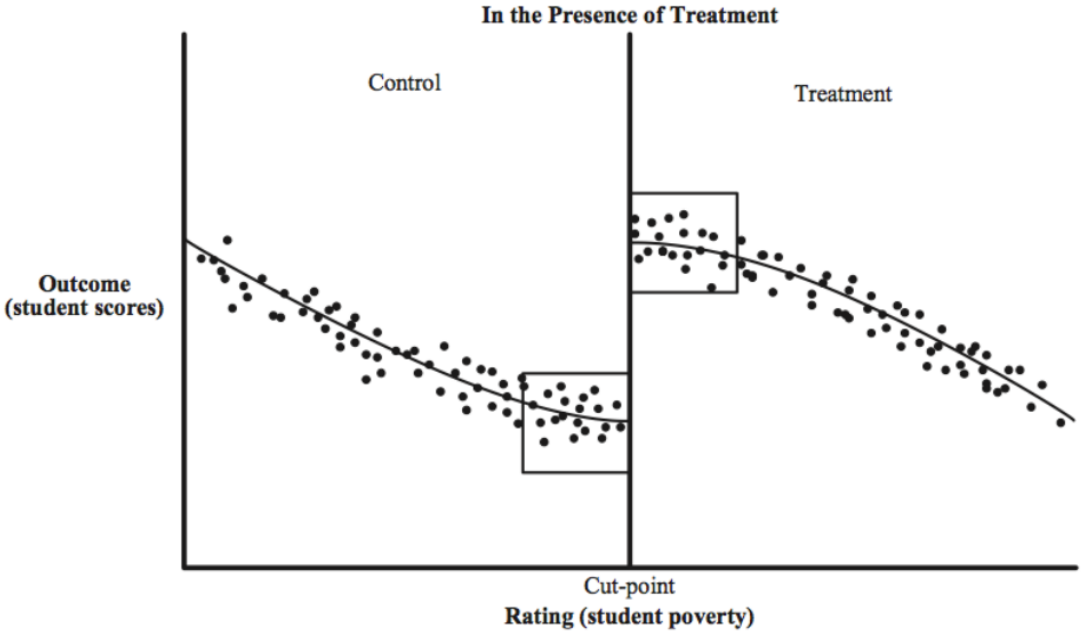

Figure 2. Jumping of the Outcome Variable at the Cut Point

Regression Discontinuity Design can be divided into two types: Sharp Regression Discontinuity (Sharp RDD) and Fuzzy Regression Discontinuity (Fuzzy RDD).

2.1.1 Sharp Regression Discontinuity

Assuming that the intervention variable (*Di*) indicates whether the individual has accepted a certain decision intervention, that is, satisfying (1):

$$D_i = \begin{cases} 0 & if\ X_i \geq X^* \\ 1 & if\ X_i < X^* \end{cases} \quad (1)$$

In formula (1), when the value of the assignment variable *Xi* is greater than or equal to the threshold value X^*, it indicates that the individual is subject to this decision intervention, which is the treatment group. When the value of the assignment variable *Xi* is less than the threshold value X^*, it means that the individual has not received the decision intervention, which is the control group. Individuals on the left and right sides of the cut point shall meet formula (2):

$$\Pr(D_i = 1 | X_{\mathrm{i}} < X^*) = 0, \Pr(D_i = 1 | X_{\mathrm{i}} > X^*) = 1 \quad (2)$$

The model is:

$$Y_i = \alpha + \beta_{01}(X_i - X^*) + \rho D_i + \beta_{11} D_i (X_i - X^*) + \delta Z_{\mathrm{i}} + \eta_i \quad (3)$$

Among them, η_i refers to the conditions that control other variables.

2.1.2 Fuzzy Regression Discontinuity

In the fuzzy RDD, the intervention variable (*Di*) is not a definite function about the threshold, and the probability of an individual receiving processing at the cut point has a jump to satisfy equation (4):

$$0 < |\Pr(D_i = 1 | X_i > X^*) - \Pr(D_i = 1 | X_i < X^*)| < 1 \quad (4)$$

In fuzzy RDD, cut-point can be used as the instrumental variables of intervention variables, that is, T_i is instrumental variables of *di*, and two-stage least squares standard error can be used for statistical inference.

The model is:

$$Y_i = \alpha + \beta_{01}(X_i - X^*) + \rho T_i + \beta_{11} T_i (X_i - X^*) + \delta Z_i + \eta_i \quad (5)$$

Among them, the first stage of regression is:

$$D_i = \alpha_D + \rho_D T_i + \delta_D (X_i - X^*) + \lambda_D T_i (X_i - X^*) + \xi_i \quad (6)$$

Among them, ρ estimates Intent -to-Treat Effect (ITT).

2.2 Five Criteria for Regression Discontinuity Design

Research using cut-point regression design must meet the five standards specified by WWC (Clearinghouse, 2014), as follows:

Criterion 1: Integrity of Assigned Variables (or Mandatory Variables)

Specifically, it includes:

(1) Is the threshold already known? Can a force variable operate on a threshold? Internal efficiency may be affected if individuals are able to manipulate mandatory variables (e.g., baseline test scores) and know thresholds related to intervention status. If there is a violation of internal validity, it must be able to systematically classify the unobservable variables related to the outcomes. If the operation is feasible, the covariate balance check can also provide more information.

(2) Describe the process of intervention allocation in detail

(3) The density test method is used to present statistical and graphical evidence describing the density smoothness of the forced variable near the threshold. By checking whether the distribution of forced variables is smooth, the clustering distribution of observed values above and below the threshold value means the feasibility of cut-point regression operation. The "density test" tests the original assumption that the density discontinuity at the threshold is zero. If the test results reject the original hypothesis, it indicates that the cut-point regression operation is not feasible. Calonico (2016) extended the method.

Criterion 2: Sample Attrition

Calculate the total attrition of ITT samples (e.g., samples in the intervention \ control group) and the attrition difference at cut point.

Criterion 3: Continuity of Outcome Variables and Mandatory Variables

Auxiliary regression was used to check the covariate balance near the threshold. In the Regression Discontinuity Design model, Z represents a fixed covariant:

$$Y_i = \alpha + \beta_{01}(X_i - X^*) + \rho D_i + \beta_{11} D_i (X_i - X^*) + \delta Z_i + \eta_i$$

One method is to check the covariant balance by estimating the auxiliary regression discontinuity design at each covariant.

$$Z_i = \alpha + \beta_{01}(X_i - X^*) + \rho D_i + \beta_{11} D_i (X_i - X^*) + \varepsilon_i$$

A zero result means the covariates are balanced. However, if we have "many" covariates, we may occasionally reject some assumptions H_0: $\rho = 0$ ("multiple comparison" problem).

Another method is to do regression about Y and Z and get the predicted value of Y (that is, a single regression weighted covariant index), and use this index in the covariant balance check:

$$\hat{Y}_i = \alpha + \beta_{01}(X_i - X^*) + \rho D_i + \beta_{11} D_i (X_i - X^*) + \varepsilon_i$$

Criterion 4: Function Form and Bandwidth

(1) Display graphical evidence (i.e., scatter diagram and fitting curve describing the relationship between outcome variables and forced variables).

The importance of graphical evidence is mainly manifested in the following two aspects. One is to visualize the data form in the first stage. According to the interval width of variables near the threshold and the average value of the corresponding outcome variable (Y_i) and distribution variable (*Xi*), and draw appropriate lines based on the graph, so that the results are transparent and effective, and provide guidance for selecting functional equations. Second, choose the appropriate interval width. If the interval width near the threshold based on variables X_i is too narrow, it will show that the data volatility is stronger (the data distribution is not smooth), and if the interval width is too large, it is easy to cover up the data volatility (the data distribution is too smooth).

(2) Provide evidence of proper function form (e.g., higher order polynomials, local linear regression), and use Akaike Information Theory Criterion (AIC).

Higher order polynomials above/below threshold:

$$\tilde{X}_i = (X_i - X^*)$$

$$Y_i = \alpha + \beta_{01} \tilde{X}_i + \beta_{11} \tilde{X}_i^2 + \ldots \beta_{0p} \tilde{X}_i^p + \rho D_i$$
$$+ \beta_{11} D_i \tilde{X}_i + \beta_{12} D_i \tilde{X}_i^2 + \ldots \beta_{1p} D_i \tilde{X}_i^p + \eta_i$$

Note: It is necessary to concentrate the mandatory variables on the threshold value for directly estimating the parameters. In practical experience, *p* rarely exceeds 2 or 3; graphic evidence provides some guidance for information standards (Schochet & Chiang, 2010).

Local linear regression:

The nonparametric method with limited data, weighted by the "triangle kernel", shows that this method emphasizes the observation of the observed values near the threshold and reduces the observation of the far-end observed values (Fan & Gijbels, 1995). Lee and Lemiuex (2010) think that the standard linear regression closely related to bandwidth is more transparent and easier to explain.

(3) Check the robustness of different bandwidths and select the appropriate bandwidth.

Choosing bandwidth for nonparametric estimation involves two problems. One is that too wide bandwidth will lead to bias; secondly, if the bandwidth is too narrow, some accuracy will be lost. A method called "cross-validation" is often used in RDD literature. This method can predict the result of each observed sample with a given nonparametric estimator and a given bandwidth. Using this method, the appropriate bandwidth can be determined from multiple candidate bandwidths (Lee & Lemieux, 2010).

The formula is as follows:

$$\hat{h}_{opt} = C_K \cdot \left(\frac{\hat{\sigma}_-^{2}(c)+\hat{\sigma}_+^{2}(c)}{\hat{f}(c)(\hat{m}_+^{(2)}(c)-(\hat{m}_-^{(2)}(c))^2+\hat{r}_-+\hat{r}_+}\right)^{1/5} \cdot N^{-1/5}$$

(4) The relationship between the outcome variable and the forced variable on both sides of the threshold is not necessarily the same.

Criterion 5: take fuzzy RDD as tool variable

(1) Measure the effectiveness of constraints

(2) Calculate the impact of the first-stage regression on distribution

(3) Check the validity of weak tool variables (Hausman, Stock, & Yogo, 2005)

3. Application of RDD in Traffic Policy Evaluation

This paper reviews the related articles on traffic policy published in top journals in recent years, which are mainly manifested in four aspects: toll roads, road safety, driving, vehicle restriction policies, and the impact of traffic congestion mitigation measures.

3.1 Application in Toll Road Field

Shihe Fu and Yizhen Gu (2017) studied the impact of highway toll pricing on air pollution for the first time. During the National Day holiday in 2012, they implemented the policy of exempting highway tolls nationwide. Using the daily pollution and weather data of 98 cities in China in 2011 and 2012, using RDD and difference method, and taking the National Day of 2011 as a control, the study found that the toll-free has increased air pollution by 20%, and the visibility decreased by 1 km, thus estimating the impact of road pricing charges on air pollution.

In Los Angeles, the labeling policy of clean energy vehicles has been implemented. Bento et al. (2014) used RDD to estimate the interaction between labeling policy and congestion. Although the policy of allowing the use of hybrid vehicles on high occupancy motor vehicles is a "free" method to stimulate the use of hybrid vehicles, the author's empirical evidence shows that California's labeling policy has caused a large number of welfare losses, and most of the losses caused by the interaction between policy and congestion are due to the adoption of new types. Adding a hybrid vehicle to the road at 7 o'clock every morning will increase the social cost by 4,500 dollars/year, reflecting the trade-off between welfare loss and environmental gain.

3.2 Road Safety Field

Lindo et al. (2016) took the legal drinking age of 18 as the threshold and applied multichannel data to comprehensively analyze the impact of the minimum legal drinking age on adolescent health in New South Wales. In this study, the restricted use data of Australian family income and labor dynamics collected from 2001 to 2011 were used to evaluate the impact of the minimum legal drinking age on drinking behavior. Using the data of driver's license authorization days and traffic accidents involving personal injury or vehicle towing provided by NSW Highway and Maritime Center from 2000 to 2010, this paper evaluates the influence of the minimum legal drinking age on the traffic accident rate of each driver; using the hospitalization data of the National Hospital Incidence database (including the hospitalization period from 2001 to 2010), this paper evaluates the impact of traffic accidents caused by alcohol or alcoholism, personal attacks and other injuries on admission or hospitalization. It is found that when young people in NSW reach the legal drinking age, drinking behavior will greatly increase, but there is no evidence that this will have an impact on any type of motor vehicle accidents, although alcohol abuse has a great impact on drinking and hospitalization.

Paola et al. (2012) used the method of RDD to evaluate the causal relationship between the penalty system of illegal driving and road safety. Based on the data of road traffic accidents, traffic accident deaths, and driving crimes in Italy from 2001 to 2005, this study estimated the impact of the penalty system for driving violations introduced in July 2003. The results show that the introduction of the penalty system for illegal driving reduces about 9% of traffic accidents, 18% of traffic accident injuries, and 30% of traffic accident deaths by controlling seasonal factors, weather conditions, traffic intensity, police patrol times, speeding cameras, gasoline prices and unemployment rate. In addition, the introduction of the new penalty system for illegal driving has led to great changes in the sanctions scheme, and the new system has a relatively weak effect on the crimes subject to nonmonetary sanctions under the old system, such as the revocation of driver's licenses.

3.3 Evaluation of Vehicle Restriction Policy

During the 2008 Olympic Games, Beijing implemented a driving restriction measure based on motor vehicle license numbers to reduce air pollution and traffic congestion. After the Olympic Games, this restriction has been revised several times, and there are mainly two policy changes: first, the short time of limiting the vehicle number leads to the weakening of the revised policy; second, the higher penalty for violators and the restriction of vehicles lead to the strengthening of the reformed policy. Lu (2016), based on two policy changes, used RDD in Tobit model to evaluate the impact of these two policy changes on air pollution and traffic congestion. It is found that weak traffic policies will lead to more environmental pollution, while strong traffic policies will improve air quality. In addition, it is found that the number limit policy increases people's use of public transportation and alleviates traffic congestion to a certain extent.

The effectiveness of driving restriction policies in improving air quality is relatively lacking. Ye (2017), based on Lanzhou's upgraded driving restriction policy, collected real-time data of six pollutants including PM2.5, PM10, SO_2, NO_2, O_3 and CO by using RDD method and hourly data of observation stations, and found that driving restriction was upgraded from one day per week to odd days per week, which did not improve the overall air quality of Lanzhou. On the contrary, the air pollution level in Lanzhou deteriorated immediately after the implementation of the stricter driving restriction policy. In addition, drivers respond to policy restrictions by changing the itinerary, acquiring/renting additional vehicles and rescheduling vehicles on unrestricted streets.

3.4 Traffic Congestion Mitigation and Its Impact on Other Areas

Will the input and use of public transportation reduce traffic congestion? Yang et al. (2018) used RDD method to study how six subway lines opened in Beijing from 2009 to 2015 affected the congestion in Beijing. The study found that after the opening of subway lines, the traffic congestion dropped sharply, and the opening of each subway reduced the daily driving delay time by 15% in Beijing in a short period of time, while the driving speed of cars along the subway increased. From the perspective of causal inference, the study shows that the opening of the subway plays a role in reducing traffic diversion in early peak hours.

With the investment in transportation infrastructure construction, it brings convenience for people to travel, but also brings noise troubles. Ahlfeldt et al. (2016) analyzed the positive accessibility effect and negative noise effect brought by the construction of the first ground rail transit in Berlin, Germany, in 1902, and found that the distance of 1 km from the rail transit station will increase the housing and land price by 21%, and the increase of 10 dB noise will reduce the house price by 5%. If these effects are not estimated as conditions, they are underestimated by 40%-80%. The analysis of real estate transaction data shows that in the 20th century, people's preference for travel accessibility remained stable, but their sensitivity to noise increased significantly. This is also the reason urban construction tends to build underground transportation instead of above-ground rail transportation.

4. Comparison of Advantages and Disadvantages of RDD and Other Methods in the Field of Traffic Policy Evaluation

In the field of transportation policy evaluation, since policy implementation and development often have time nodes, and RDD is usually allocated according to threshold rules, there are more opportunities to use RDD method. RDD is a quasi-experimental method, which has incomparable advantages in policy inference evaluation. Its most obvious feature is that RDD can balance unobserved factors like randomized experiments, but OLS-based and matching methods must rely on the strong conditional independent mean hypothesis, that is, there are no other factors that need to be measured.

Compared with randomized experiments, RDD is widely used for the following reasons:

First, RDD is to use the existing data for analysis, instead of spending a lot of manpower and material resources on project investigation, intervention, treatment, evaluation, etc. just like randomized experiments, which greatly saves the research cost and time from the perspective of cost.

Second, RDD can reflect the effect of policy intervention on marginal units near cut-points, and the estimation of RDD has obvious advantages in explaining the average effect of samples, which is conducive to studying the impact of breakpoints on subjects and has obvious enlightenment on whether the intervention intensity should be expanded or reduced.

Third, the RDD can avoid the constraints and troubles in ethics, politics, and morality. It is necessary to set up a control group and an experimental group to carry out randomized experiments, among which there are ethical constraints, whether the subjects are fully informed, whether the subjects in the control group know that they are assigned to the control group, and so on. These problems have always been difficult to carry out experiments. However, RDD can effectively avoid this constraint, or when it is difficult to carry out randomized experiments, RDD is undoubtedly the best choice.

Last, RDD is usually closer to reality than the idealized randomized experiments and can be used to evaluate the effect of implementing relevant policies and decrees. For the impact evaluation of large-scale transportation public policy projects, the government and decision-makers may be more concerned about the implementation effect in the actual social environment, rather than the effect under ideal conditions.

In the following, we use RDD to compare with other commonly used causal inference methods on observation data and illustrate their advantages and disadvantages, as shown in Table 1.

Table 1. Comparison of RDD and Other Causal Inference Methods

Examples	Methods adopted in research	Research content	Comparison of advantages and disadvantages of research methods
Nidup (2016)	Difference-in-difference method (DID)	The study uses TO measure the impact of improving rural road connectivity in Bhutan on income. Based on the data of Bhutan's Living Standard survey in 2007 and 2012, the author constructs the income index by principal component analysis and finds that road connectivity helps to	Advantages: (1) Based on the national large-scale data survey, the sample size is sufficient and representative; (2) Allow the existence of unobservable factors, and allow unobservable factors to influence the decision-making of whether to

		increase the income of low-quantile groups, but has no significant effect on the median group, and at the same time reduces the income effect of high-quantile groups.	accept intervention, thus relaxing the conditions of policy evaluation, making the application of policy evaluation closer to the economic reality, so it is widely used. (3) The effect of intervention can be evaluated in the real environment. Disadvantages: (1) The treatment group and the control group need to follow the parallel trend assumption (2) The DID method is based on the panel data model, which requires not only cross-sectional unit data, but also individual time series data (3) ATT can be identified, but ATE cannot be identified
Habyarimana and Jack (2011)	Randomized experiment	A randomized experiment was designed to improve road safety driving in Kenya. In the experimental group, drivers who are dissatisfied with driving skills can express their riding experience by writing stickers and placing them in obvious positions in the car. In the control group, there is no way for passengers to express their dissatisfaction with drivers who are dissatisfied with driving skills. The experimental results show that, compared with the control group, the insurance	Advantages: (1) Randomization can ensure the balance between observed and unobserved covariates. (2) Form an effective counterfactual, which is convenient to estimate the average causal effect of ITT and compliance. Disadvantages: (1) It is rarely encountered in real economic activities, and the implementation cost is high and the time is long;

		claims of the experimental group have decreased by half to two-thirds, and the claims involving injury or death have decreased by 60%. Compared with other means of reducing road mortality in the public health field, the effect is remarkable.	(2) Assessing in a controlled environment, but may not provide real-world validity information. (3) It may face with moral and ethical problems due to screening criteria and choice of consent.
Paola, Scoppa and Falcone (2013)	RDD	Based on the daily information data of road traffic accidents, traffic accident deaths, and driving crimes provided by Italian police, this paper compares the traffic accident deaths before and after the introduction of penalty system by using RDD. After controlling the seasonal factors, weather conditions, traffic intensity, police patrol time, time trends, and other factors, it is found that the penalty system has a negative impact on all samples. The introduction of the penalty system reduced about 9% of road traffic accidents, while the impact on traffic injuries and deaths decreased by 18% and 30%, respectively. It is found that it remains robust in time span.	Advantages: (1) it is relatively easy to realize (2) It is convenient to obtain representative large samples (3) It has good external effectiveness (4) Random variability produces local randomization with the threshold; balance test. (5) It can establish an effective counterfactuals. Disadvantages: (1) the technical operation is slightly more difficult. (2) The conclusion is confined to the scope of the problem studied, lacking external validity.

5. Application and Prospect of Regression Discontinuity Design in Developing Countries' Transport Policy Evaluation Field

Regression Discontinuity Design has not been fully applied in the field of transportation policy evaluation in China as well as developing countries. Literature search found that the international research in the field of transportation policy based on RDD mainly appeared after 2012. In addition, the retrieval of the literature found that most of the literature using causal inference research methods for traffic policy impact assessment appeared after 2014, and few papers used RDD. Why has not RDD been widely used? On the one hand, it might be that the RDD method is slightly more difficult to operate and master. On the other hand, the RDD method is mostly applied to the field of economics, and the researchers in the field of transportation policy do not have a good grasp of this method. Most of the scholars in the field of transportation policy evaluation are experts on transportation or transportation management. Few people can apply this method to empirical research, therefore it has not been widely used in the field of transportation.

In developing countries, such as China, the causal inference method of difference-in-difference method has been applied to assess the "business tax" on the impact of the tax burden of transport listed companies (Li & Li, 2016), and the effect of intelligent transportation system to improve traffic congestion in Beijing (Wu & Wang, 2015). RDD is applied to evaluate the effect of improving vehicle emission standards on smog management (Sun, 2017). The impact and effect evaluation of RDD in the field of transportation policy has great application potential (Hutchinson, 2007). Combined with the international research in this field, the areas that can be applied are as follows: (1) Carry out the impact assessment of toll roads: domestic expressways have been following the toll model, but foreign expressways are regarded as public good, and no toll is taken. The impact of toll roads on the occurrence rate of safety accidents, environmental pollution, and travel choice is worth investigating; (2) Evaluation on the effect of road traffic policy implementation: whether the traffic restrictions in winter have played the expected effect of easing traffic pressure and controlling smog, whether the safe road advocacy of "letting people drive" has got the actual effect, and whether strengthening the strictness of driver's license examination and training is beneficial to traffic safety and other topics are worthy of in-depth study; (3) The impact of various modes of transportation on the individual traffic behavior: There has been broad research space, such as whether the addition of subway lines to ease the traffic pressure, the impact of the network about the traffic travel behavior.

In this era of "knowing what it is, but also knowing why it is", there is no doubt that the causal inference method gives us a powerful tool to explore the causal relationship between social relations and the evolution of public policy. As one of the easy-to-realize, highly reliable, and advanced methods, RDD will lead the evaluation of transportation policy, and even the evaluation of policy effects at all levels of society into a new stage of development, with infinite potential.

6. Conclusion

In this paper, the origin, basic idea, model, and applications of RDD are summarized, and its application in the field of transportation is analyzed through literature review. The advantages and disadvantages of RDD method and other causal inference methods in the study of transport policy evaluation are compared. The future application prospect of RDD in the field of transportation evaluation in developing countries is prospected, which provides basic support for the application in the field of transport economy in developing countries.

Acknowledgements

The authors are supported by Excellent Graduate Training Program of Shaanxi Normal University (LHRCCX23124), Natural Science Foundation of Shaanxi Province (2023-JC-QN-0784), Xi'an Soft Science Research Program (24RKYJ0039). The authors would also like to thank scholars from Stanford University for their teaching and instruction on Regression Discontinuity Design during corresponding author’s visiting research.

References

Ahlfeldt, G., Nitsch, V., & Wendland, N. (2016). Ease vs. Noise: On the Conflicting Effects of Transportation Infrastructure. *Working Paper*. CESifo Group Munich. https://doi.org/10.2139/ssrn.2851720

Angrist, J. D., & Pischke, J. S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives*, *24*(2), 3-30. https://doi.org/10.1257/jep.24.2.3

Angrist, J., Bettinger, E., Bloom, E., et al. (2002). Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment. *American Economic Review*, *92*(5), 1535-1558. https://doi.org/10.1257/000282802762024629

Aragón, M. I. B., Jiménez, D. J., & Valle, R. S. (2014). Training and performance: The mediating role of organizational learning. *Business Research Quarterly*, *17*(3), 161-173. https://doi.org/10.1016/j.cede.2013.05.003

Bento, A., Kaffine, D., Roth, K., et al. (2014). The effects of regulation in the presence of multiple unpriced externalities: Evidence from the transportation sector. *American Economic Journal: Economic Policy*, *6*(3), 1-29. https://doi.org/10.1257/pol.6.3.1

Berk, R. A., & Rauma, D. (1983). Capitalizing on nonrandom assignment to treatments: A regression-discontinuity evaluation of a crime-control program. *Journal of the American Statistical Association*, *78*(381), 21-27. https://doi.org/10.1080/01621459.1983.10477917

Bor, J., Fox, M. P., Rosen, S., et al. (2017). Treatment eligibility and retention in clinical HIV care: A regression discontinuity study in South Africa. *PLoS medicine*, *14*(11), e1002463. https://doi.org/10.1371/journal.pmed.1002463

Calonico, S., Cattaneo, M. D., Farrell, M. H., et al. (2016). Regression discontinuity designs using covariates. *Working Paper*, University of Michigan.

Clearinghouse, W. W. (2014). *Procedures and standards handbook* (Version 3.0). Washington, DC: US Department of Education.

De Paola, M., Scoppa, V., & Falcone, M. (2013). The deterrent effects of the penalty points system for driving offences: a regression discontinuity approach. *Empirical Economics*, *45*(2), 965-985. https://doi.org/10.1007/s00181-012-0642-9

Fan, J., & Gijbels, I. (1995). Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society*. Series B (Methodological), 371-394. https://doi.org/10.1111/j.2517-6161.1995.tb02034.x

Fu, S., & Gu, Y. (2017). Highway toll and air pollution: Evidence from Chinese cities. *Journal of Environmental Economics and Management*, *83*, 32-49. https://doi.org/10.1016/j.jeem.2016.11.007

Habyarimana, J., & Jack, W. (2011). Heckle and Chide: Results of a randomized road safety intervention in Kenya. *Journal of Public Economics*, *95*(11-12), 1438-1446. https://doi.org/10.1016/j.jpubeco.2011.06.008

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, *69*(1), 201-209. https://doi.org/10.1111/1468-0262.00183

Hanushek, E. A. (1999). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis*, *21*(2), 143-163. https://doi.org/10.3102/01623737021002143

Hausman, J., Stock, J. H., & Yogo, M. (2005). Asymptotic properties of the Hahn–Hausman test for weak-instruments. *Economics Letters*, *89*(3), 333-342. https://doi.org/10.1016/j.econlet.2005.06.007

Hsieh, C. T., & Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of public Economics*, *90*(8-9), 1477-1503. https://doi.org/10.1016/j.jpubeco.2005.11.002

Hutchinson, T. P. (2007). Concerns about methodology used in real-world experiments on transport and transport safety. *Journal of transportation engineering*, *133*(1), 30-38. https://doi.org/10.1061/(ASCE)0733-947X(2007)133:1(30)

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, *48*(2), 281-355. https://doi.org/10.1257/jel.48.2.281

Li, W., & Li, W. (2016). Research on the Influence of Business Reform on Tax Burden and Finance of Listed Companies in Transportation Industry. *Shanghai Journal of Economics*, *2016*(1), 61-70.

Lindo, J. M., Siminski, P., & Yerokhin, O. (2016). Breaking the link between legal access to alcohol and motor vehicle accidents: evidence from New South Wales. *Health economics*, *25*(7), 908-928. https://doi.org/10.1002/hec.3197

Lu, X. (2016). Effectiveness of government enforcement in driving restrictions: a case in Beijing, China. *Environmental Economics and Policy Studies*, *18*(1), 63-92. https://doi.org/10.1007/s10018-015-0112-7

Moers, F. (2005). Discretion and bias in performance evaluation: the impact of diversity and subjectivity. Accounting, *Organizations and Society*, *30*(1), 67-80. https://doi.org/10.1016/j.aos.2003.11.001

Müller, T., & Shaikh, M. (2018). Your Retirement and My Health Behavior: Evidence on Retirement Externalities from a fuzzy regression discontinuity design. *Journal of health economics*, *57*, 45-59. https://doi.org/10.1016/j.jhealeco.2017.10.005

Nidup, J. (2016). Road Accessibility and Wealth in Rural Bhutan: A Difference in Difference Approach. *Journal of Economic Development*, *41*(3), 55. https://doi.org/10.35866/caujed.2016.41.3.004

Rubin, D. B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of educational Statistics*, *2*(1), 1-26. https://doi.org/10.3102/10769986002001001

Schochet, P. Z., & Chiang, H. S. (2010). *Error Rates in Measuring Teacher and School Performance Based on Student Test Score Gains*. NCEE 2010-4004. National Center for Education Evaluation and Regional Assistance.

Sels, L., De Winne, S., Delmotte, J., et al. (2006). Linking HRM and small business performance: An examination of the impact of HRM intensity on the productivity and financial performance of small businesses. *Small business economics*, *26*(1), 83-101. https://doi.org/10.1007/s11187-004-6488-6

Sun, K. (2017). Research on smog control effect of motor vehicle emission standards: Analysis based on RDD. *Soft Science*, *2017*(11), 93-97.

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, *51*(6), 309. https://doi.org/10.1037/h0044319

Wu, L., & Wang, Y. (2015). Implementation of Intelligent Transportation System (ITS) and Improvement of Urban Traffic Smoothness —— DID Estimation Analysis Based on Panel Data. *Logistics Technology*, *2015*(22), 113-116.

Yang, J., Chen, S., Qin, P., et al. (2018). The effect of subway expansions on vehicle congestion: Evidence from Beijing. *Journal of Environmental Economics and Management*, *88*, 114-133. https://doi.org/10.1016/j.jeem.2017.09.007

Ye, J. (2017). Better safe than sorry? Evidence from Lanzhou's driving restriction policy. *China Economic Review*, *45*, 1-21. https://doi.org/10.1016/j.chieco.2017.05.009

Zhao Xiliang. (2017). *Mostly Useful Science of Econometrics*. Beijing: Peking University Press.