Annals of Tourism Research 74 (2019) 1–16

Contents lists available at ScienceDirect

Annals of Tourism Research

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

Regional policy and tourism: A quasi-natural experiment

Taotao Deng[a,*], Yukun Hu[a], Mulan Ma[b,*]

[a] School of Urban and Regional Science, Shanghai University of Finance and Economics, No. 777 Guoding Road, Shanghai 200433, China
[b] School of Tourism and Event Management, Shanghai University of International Business and Economics, No. 1900 Wenxiang Road, Shanghai 201620, China

ARTICLE INFO

Associate Editor: Albert Assaf

Keywords:
Western Development Strategy
Policy effects
Regression discontinuity design
Tourism industry

ABSTRACT

China's Western Development Strategy (WDS) has generated stark differences in policy environments for tourism development. Considering the WDS as a quasi-natural experiment, the paper applied an increasingly important regression discontinuity design method to tourism studies and explored the causal effect of the WDS on tourism by comparing treated and untreated cities proximal to the geographic boundary of the WDS. We prove that the WDS has caused a significantly positive effect on tourism development and is estimated to be up to 6% and vary from 5.9% to 6.7%. The mechanism analysis indicates that the WDS can affect tourism development through infrastructure construction and tax incentives.

Introduction

Since the beginning of the 21st century, China has implemented an ambitious Western Development Strategy (WDS) to accelerate the development and reduce regional inequality in its western region. The WDS, which has the longest implementation duration and covers the widest scope of all of China's development strategies, is widely considered the most influential regional development policy in China since the reform and opening up (Fan, Kanbur, & Zhang, 2011). Despite the western region's poor economic performance, this vast, remote, and economically lagging area of China has great potential for economic development due to its abundant natural and cultural resources. In order to promote economic development of the western region, the central government has approved a range of infrastructure projects, strengthened the protections for the ecological environment, and implemented preferential fiscal and tax incentive policies. In many western provinces, the spatial distribution of poverty-stricken areas highly overlaps with tourism resource-rich areas (Ma, 2001). As a labour-intensive industry closely interrelated to the other industries, tourism is recognised as the 'pillar industry' or 'leading industry' in many local governments in the western region.

Since the implementation of the WDS, the western region has experienced rapid economic growth; however, no general consensus has been reached regarding the effect of the WDS on the western region. For example, Song (2013) indicates that although the western region has achieved economic progress, regional inequality with other regions has shown a further trend of deterioration. During the implementation of the WDS, the tourism industry in China's western region made great progress. From 2002 to 2010, the total tourism revenue increased from 18.36 billion USD to 103.7 billion USD, and the number of tourists visiting China's Western region increased from 304.6 million to 1,001.8 million. Conducting an evaluation of the implementation of the WDS is critical to determine its effects and appropriate policy adjustments; thus, the key question of interest becomes how to measure the causal effect of the WDS on tourism growth.

* Corresponding authors at: School of Urban and Regional Science, Shanghai University of Finance and Economics, No. 777 Guoding Road, Shanghai 200433, China.
E-mail addresses: deng.taotao@mail.shufe.edu.cn (T. Deng), mamulan@suibe.edu.cn (M. Ma).

https://doi.org/10.1016/j.annals.2018.10.001
Received 4 May 2018; Received in revised form 19 September 2018; Accepted 2 October 2018

Implementing the WDS has the potential to create opportunities in the tourism industry in the west region. Particularly, many important inter-regional transport projects were completed and strongly enhanced communications between China's inland and coastal areas. The causal impact of the WDS on the tourism industry has not been investigated. Although a few studies (Ma, 2001; Jackson, 2006) have mentioned the potential importance of the WDS to the tourism industry in the western region, insufficient analyses have been presented. In particular, Jackson (2006) calls for richer data and more rigorous methods to analyse how the WDS affects the tourism industry. Moreover, how the WDS may affect the tourism industry (i.e. the mechanism analysis) has not been explored.

With the increasing level of opening up and the rapidly emerging middle class, China's tourism industry has flourished over recent years; thus, credibly evaluating how the WDS affects the tourism industry is challenging because of potential endogenous problems. From an academic perspective, evaluating China's booming tourism industry is insufficient to prove the significance of the degree of the effect of the WDS on tourism development in the western region if the method is simply comparing tourism revenue or tourism flow before and after the WDS in a trend analysis approach. Whether this positive influence is caused by implementing the WDS or because of benefits from the increase in mass tourism must be further explored.

The central question in the policy evaluation is how to identify a causal impact from the confounding effect influencing the outcome of interest (Battistin & Rettore, 2008). To deal with the endogeneity issue, the paper attempts to apply a regression discontinuity design (RDD), which is an advanced method in policy evaluation in a quasi-natural experimental design. An advantage of the RDD method is its effectiveness regarding managing the impact of the endogenous problem and improving the estimation accuracy of the policy effect. Compared with the traditional regression methods, an RDD is a good method to estimate a causal relationship in the 'jump' effect caused by the policy. As a regional development strategy, the WDS has generated stark differences in policy environments for developing tourism economics. The implementation scope of the WDS is determined by the exogenous geographical boundary, which provides relatively credible natural experimental evidence for this study. To explore the policy effects of the WDS on the tourism industry, the paper compares the tourism industry in cities adjacent to the programme boundary (i.e. the cut-off point) between the 'west' (who benefits from WDS) and other regions (who does not benefit from the WDS).

The paper contributes to the tourism literature in three aspects. First, the paper proposed a conceptual analysis framework for policy evaluation in tourism studies. Second, based on data from China's prefecture-level cities, the paper used empirical-based evidence to investigate the causal effects of the WDS on the tourism industry in China's western region. Third, since the RDD method has become increasingly important in policy evaluation, the paper has applied a novel RDD to tourism research, which enriches application methods in tourism economics.

The remainder of the paper is structured as follows. Section 2 presents a literature review. Section 3 specifies the research background, conceptual framework, and econometric models. Section 4 describes the empirical results. Section 5 draws conclusions.

Literature review

Impact of public policy on tourism development

Early studies on how public policy affects tourism development have mainly focused on theoretical analysis. Pearce (1998) and Vernon (2005) emphasise the role of government in promoting tourism development regarding institutional mode selection, policy formulation, and promotion of public and private sector cooperation, respectively. Alipour and Kilic (2005) indicate that the failure of policy institutions has caused the difference in tourism development between two similar regions. Dredge (2006) shows that policy networks can shape tourism development by harnessing public–private partnerships. In general, these studies have confirmed that public policy affects tourism development.

With the application of econometric methods in tourism studies, empirical studies began exploring how policy affects tourism. For example, Gülcan, Kuştepeli, and Akgüngör (2009) use an econometric method to examine the effect of public investment on value added in the tourism industry. The results show that tourism development is significantly promoted by public policies. Lejárraga and Walkenhorst (2013) explore the determinants of the linkages between tourism and the national economy from developing countries and find that public policy can enhance these linkages through improving business environments and trade regulations.

Public policies often play a key role in developing a country's tourism industry (Lejárraga & Walkenhorst, 2013). As the largest developing country, China's tourism development is strongly motivated by politics (Richter, 1983). The early studies have mainly focused on the initial stage of tourism in the context of the reform and opening up in China. Chow (1988) reveals that the initial development of tourism is strongly related to China's institutional transition. Sofield and Li (1998) indicate that China's tourism development is related to its cultural policies. Zhang, Chong, and Ap (1999) summarise the role of policy in developing inbound tourism in China as five aspects: infrastructure, regulations, financial incentives, coordination among government departments, tourism education, and training. He (2013) explores the influence mechanism of tax policies by replacing a business tax with a value-added tax on the tourism service industry in China. It concludes that the tax policy will increase the tax burden of tourism service enterprises.

The literature has improved our understanding of how public policy affects tourism development; however, current literature has some limitations. First, most studies confine their perspectives to tourism policy and ignore the comprehensive policies that determine the development environment of tourism, such as regional policies. The comprehensive policies may have more complex and profound effects on tourism development. On the one hand, the formulation and implementation of tourism policies requires comprehensive consideration of environmental factors, such as the economy, society and culture (Meethan, 1998; Wang & Ap, 2013), namely, the comprehensive policy can affect the policy environment of tourism. On the other hand, the comprehensive policy can

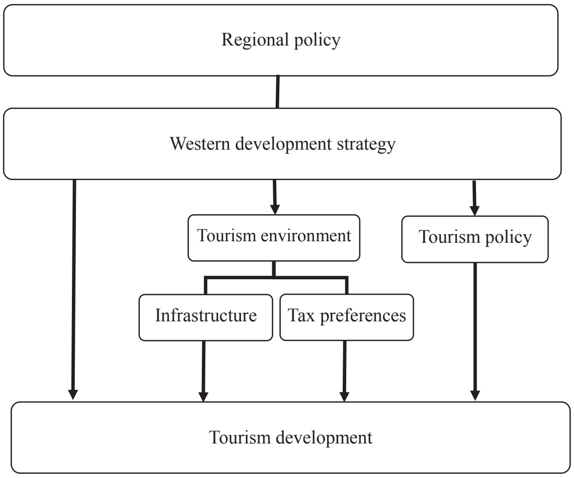

Fig. 1. Theory framework for the causal impact of regional policy on tourism.

determine the industrial environment of tourism, such as infrastructure, taxes, market. Second, the literature has mainly used simple statistical analyses to estimate the policy effect on tourism development, and there is a lack of analysis based on rich data and rigorous methods, especially the causal inference method.

Causal impact of regional policy on tourism development

Regional policy is likely to generate a causal impact on tourism development through the following ways, as shown in Fig. 1.

1) Regional policy directly influences tourism development. First, the main purpose of a regional policy is to alleviate regional imbalance in China (Hong, Liu, & Huang, 2014). In practice, especially for developing countries, tourism plays an important role in reducing regional inequality (Li, Chen, Li, & Goh, 2016). Therefore, there is a reason to believe that the regional policy embraces some items directly related to tourism to achieve regional objects. In the context of WDS, tax preferences for tourism enterprises, priority to tourism projects and protection of cultural resources are all exactly regional policy terms that directly determine tourism development. Second, the targets of regional policy are the undeveloped regions, which typically possess stronger potentials and intentions of tourism development. On the one hand, the weak industries stimulate the motives for developing tourism; on the other hand, the mild external intervention is conducive to foster tourism potentials. Taking western China as an example, the distribution of tourism resources is highly correlated with the distribution of poverty (Ma, 2001). Therefore, tourism is likely to become a prior instrument for regional policy to support the undeveloped regions. Moreover, with Pro-poor Tourism (PPT) proposed, tourism poverty alleviation has become the focus of tourism researches (Zeng & Ryan, 2012). The wide practice of tourism poverty alleviation further enhances the integration of tourism and regional policies.
2) Regional policy influences the development environments of tourism. As a comprehensive public policy, regional policy has multiple impacts on local economy in China (Hong et al., 2014). Tourism, as an important part of regional economy, will inevitably be influenced directly or indirectly by regional policy in various ways. That is to say, besides the direct impacts, regional policy may have indirect impacts on tourism through the external environmental factors. In the context of WDS, the instruments such as infrastructure investment (e.g., transport, information) and tax preference are not only the main lines of the regional policy, but the key factors determining the development environment of tourism. Therefore, regional policies are likely to improve the development environment of tourism mainly through infrastructure improvement and tax incentives in this case. Previous studies have partly demonstrated these indirect mechanisms from different perspectives. First, massive infrastructure improvement create a favorable environment for tourism (Imikan & Ekpo, 2012). Typically, Khadaroo and Seetanah (2007) demonstrate that transport infrastructure promote the growth of tourist arrivals; Law, Leung, and Buhalis (2009) find that information technologies are widely adopted by tourism sectors to reduce costs and improve service quality. Besides the hard infrastructures, there is a necessity to improve the infrastructures of soft environment for tourism development (Thapa, 2012), such as human resource (Baum & Szivas, 2008; Liu & Wall, 2005), business environments (Lejárraga & Walkenhorst, 2013). Second, tax preferences involved in WDS contribute to improve the business environment of tourism, by stimulating innovation (Bloom, Griffith, & Reenen, 2002), optimizing industrial location (Baldwin & Okubo, 2010) and promoting enterprise entry (Rin, Giacomo, & Sembenelli, 2009).
3) Regional policy influences the formulation and implementation of tourism policy. Tourism policy is a complex system composed of multidimensional factors such as economic, social, and cultural factors (Meethan, 1998). Naturally, the formulation and implementation of tourism policy will be affected jointly by a set of comprehensive factors (Krutwaysho & Bramwell, 2010), including economic and social environment, institution and other factors (Wang & Ap, 2013). As we summarized above, regional policy is a type of comprehensive policies involving various aspects of regional development, then a regional policy will inevitably

Fig. 2. Coverage of China's WDS. *Note*: Grey area indicates the coverage of WDS.

affect the formulation and implementation of tourism policies to a certain extent.

Policy background, methodology, and data

Policy background

As shown in Fig. 2, the western region[1] in China refers to 12 provincial-level administrative regions that cover six provinces (Gansu, Qinghai, Shanxi, Sichuan, Guizhou, and Yunnan), five autonomous regions (Inner Mongolia, Ningxia, Tibet, Xinjiang and Guangxi) and one municipality (Chongqing). The western region is rich in natural and cultural resources and holds a critical strategic position. Due to natural and historical reasons, the economic situation of the western region has consistently lagged behind the other regions of China. To promote economic development in the western region, in addition to profiting from the mineral resources while preserving the natural environment, local governments have considered the tourism industry as a pillar industry.

Since the early 2000s, the State Council has released policy instruments and investment plans to increase economic development in China's vast, remote, and economically lagging western region. Those relevant policies can be roughly divided into three categories: public infrastructure investment, fiscal transfers, and preferential tax policies.

Public infrastructure investment. The transport industry is the basic sustentation in the process of economic development and has great significance to the tourism industry. In the WDS framework, the central government has made significant investments in public construction projects, including the construction and renovation of transport, energy, and hydraulic infrastructure projects as well as projects for ecological protection, education, and public health. In particular, some national key transport projects, such as the Qinghai–Tibet railway project and some airport and highway construction projects, were completed in the western region.

Fiscal transfers. The economically less developed western region has received increasingly more central financial transfer payments. In the WDS framework, the central government has enhanced fiscal transfers (e.g. general transfers, special transfers, and fiscal interest discounts) to the local governments in the western region. Those fiscal transfers have improved urban environment and service facilities for the western region.

Preferential tax policies. Tax incentives are one of the most important preferential policies in the WDS. Tax reduction and exemption have a positive effect on attracting external investments. In the WDS framework, the enterprise income tax reduction is the main tax incentive measure. During the implementation of the WDS, many policy documents on tax reduction and exemption were produced by the central government.

[1] In addition to western provinces, the coverage of WDS includes three ethnic minority areas (Xiangxi Tujia-Miao Autonomous Prefecture, Enshi Tujia-Miao Autonomous Prefecture, and Yanbian Korean Autonomous Prefecture).

Methodology: Introduction of the regression discontinuity design

Background and basic structure

Regression discontinuity design (RDD) was first exploited by Thistlethwaite and Campbell (1960) to identify treatment effects in the context of non-experiment. Since the late 1990s, an increasing number of researches have used RDD for policy evaluation (Angrist & Lavy, 1999; DiNardo & Lee, 2004; Chay, McEwan, & Urquiola, 2005; McEwan & Shapiro, 2008). Especially after Hahn, Todd, and Klaauw (2001) formally demonstrated the identification conditions and methods, RDD began to be widely used in economics, politics and sociology, and has become one of the most popular identification strategies in empirical analysis.

The literature often frames the causal inferences in the context of potential outcomes (Holland, 1986; Rubin, 1974). In the basic setting of causal inference, what we are interested in are the causal effects of a binary treatment. Let Y_{i0} and Y_{i1} represent the potential outcomes for individual i: Y_{i0} is the outcome when i is not treated, and Y_{i1} is the outcome when i is treated. The causal effect for i is exactly the difference between Y_{i1} and Y_{i0}. Because Y_{i0} and Y_{i1} can't be observed simultaneously, we need to estimate the average treatment effects (ATE) as Eq. (1):

$$\tau_{ATE} = E[Y_{i1} - Y_{i0}] \tag{1}$$

Let D_i denotes the treatment variable, with $D_i = 0$ if city i is not treated by regional policy, and $D_i = 1$ otherwise; Y_i denotes the observed outcomes, expressed by (Y_{i0}, Y_{i1}) as Eq. (2):

$$Y_i = (1 - D_i)\cdot Y_{i0} + D_i \cdot Y_{i1} \tag{2}$$

The basic idea of RDD is that D_i is determined by whether an observed running variable X_i exceed a given cutoff point c. The definition of X_i is often determined by specific policy rules. In this paper, X_i refers to the geographic distance between city i and the WDS boundary, and the cutoff point refers to the WDS boundary. Assuming at the cutoff point, $E(D_i|X_i)$ has a sharp discontinuity, while the distribution of any other factors affecting Y_i are continuous, then any discontinuity of $E(Y_i|X_i)$ can be interpreted as treatment effect of D_i. In this case, if there is a linear relationship between Y_i and X_i, the treatment effect τ we focus on can be estimated by a liner regression as Eq. (3):

$$Y_i = \alpha + \tau D_i + \beta X_i + \varepsilon_i \tag{3}$$

In the RDD, there are two potential relationships between (Y_{i1}, Y_{i0}) and X_i: $E[Y_{i1}|X_i]$ and $E[Y_{i0}|X_i]$. But by Eq. (2), we can only observe $E[Y_{i1}|X_i, D_i = 1]$ and $E[Y_{i0}|X_i, D_i = 0]$, which leads to the failure to directly obtain ATE: $E[Y_{i1}\text{-}Y_{i0}|X_i]$. Fortunately, if $E[(Y_{i0}, Y_{i1})|X_i]$ are continuous with respect to X_i, RDD can obtain a local ATE (LATE) at the cutoff point c as Eq. (4):

$$\lim_{\varepsilon \to 0^+} E[Y_i\ |X_i = c + \varepsilon] - \lim_{\varepsilon \to 0^-} E[Y_i\ |X_i = c + \varepsilon] = E[Y_{i1} - Y_{i0}\ |X_i = c] \tag{4}$$

In general, the condition of continuity enables RDD to use the observed outcomes on one side of cutoff point (e.g. $E[Y_i|X_i, D_i = 0]$) as a valid estimation of the unobservable potential outcomes on the other side of cutoff point (e.g. $E[Y_{i0}|X_i, D_i = 1]$) (Lee & Lemieux, 2010).

In practice, RDD can be divided into two categories by the assignment rules of treatment (Hahn et al., 2001; Trochim, 1984): Sharp RDD and Fuzzy RDD. In Sharp RDD, the assignment to treatment is determined completely by X_i: $D_i = 1$ if $X_i \geq c$ and $D_i = 0$ if $X_i < c$. In Fuzzy RDD, the assignment to treatment is not only determined by X_i, but also by other unobservable factors: $D_i = D(T_i, \varepsilon_i)$, where $T_i = 1$ if $X_i \geq c$ and otherwise $T_i = 0$, and ε_i is the unobserved factors.

Basic identification conditions

In order to obtain unbiased causal effects by RDD, some basic identification conditions have to be satisfied (Hahn et al., 2001). First, there exists a obvious gap in the probability of treatment receipt at cutoff point, as shown in Eq. (5):

$$\lim_{x \to c^+} E[D_i\ |X_i = x] \neq \lim_{x \to c^-} E[D_i\ |X_i = x] \tag{5}$$

In the Sharp RDD, the probability of treatment receipt change from zero to one at the cutoff point; while in the Fuzzy RDD, the probability only need a smaller jump. In general, a valid RDD requires an obvious discontinuity in the distribution of treatment at the cutoff point, otherwise it will fail due to the weakness of treatment.

Second, except for D_i, all other factors affecting Y_i have to be continuous with respect to X_i, otherwise the treatment effects will be confounded by other factors. In the framework of potential outcomes, this condition requires that $E[(Y_{i0}, Y_{i1})|X_i]$ are continuous at the cutoff point, as expressed by Eq. (6):

$$\lim_{x \to c^+} E[Y_{ij}\ |X_i = x] = \lim_{x \to c-} E[Y_{ij}\ |X_i = x],\ j = 0, 1 \tag{6}$$

In practice, the RDD often require that $E[Y_{ij}|X_i]$ are continuous at all values of X_i (Imbens & Lemieux, 2008).

Third, the key condition of RDD is that the individuals can't precisely control X_i near the cutoff point, which ensures RDD to be close to a local randomized experiment at the cutoff point. It is the condition that make RDD more transparent and convincing (Lee & Lemieux, 2010). In the framework of potential outcomes, this condition denotes that the assignment to treatment is independent on the potential outcomes near the cutoff point, as expressed by Eq. (7):

$$(Y_{i1}, Y_{i0}) \perp D_i\ |X_i \in (c - \delta, c + \delta) \tag{7}$$

where $\delta > 0$ can be made arbitrarily small.

Estimation methods

The estimation methods of RDD mainly include nonparametric regression at the boundary, local linear regression (LLR), local polynomial regression (LPR). Because the nonparametric regression does not work well at the boundary, LLR is regarded as a better selection (Hahn et al., 2001; Imbens & Lemieux, 2008; Lee & Lemieux, 2010). In Sharp RDD, the nonparametric estimator can be expressed as Eq. (8), where $K(x)$ refers to the kernel function, h represents the bandwidth.

$$\tau_{LATE} = E[Y_{i1} - Y_{i0} \mid X_i = c] = \frac{\sum_{i:X_i \geq c} Y_i \cdot K(\frac{X_i - c}{h})}{\sum_{i:X_i \geq c} K(\frac{X_i - c}{h})} - \frac{\sum_{i:X_i < c} Y_i \cdot K(\frac{X_i - c}{h})}{\sum_{i:X_i < c} K(\frac{X_i - c}{h})} \tag{8}$$

When $K(x)$ represents a rectangle kernel function, the value of Eq. (8) is actually equal to the difference of the mean of Y_i between the individuals near both sides of the cutoff point (i.e., Eq. (4)). However, by the assumption of local randomization, the nonparametric regression is valid only within a limit bandwidth near the cutoff point. Moreover, when X_i has impacts on (Y_{i1}, Y_{i0}), the differences between the treatment group and the control group tend to cause a biased nonparametric estimator. In this case, LLR can reduce the estimation bias through linear regression locally on both sides of the cutoff point respectively (Hahn et al., 2001) as Eqs. (9) and (10).

$$\min_{\alpha_l, \beta_l} \sum_{i=1}^{N} [Y_i - \alpha_l - \beta_l \cdot (X_i - c)]^2 \cdot K(\frac{X_i - c}{h}) \cdot 1(X_i < c) \tag{9}$$

$$\min_{\alpha_r, \beta_r} \sum_{i=1}^{N} [Y_i - \alpha_r - \beta_r \cdot (X_i - c)]^2 \cdot K(\frac{X_i - c}{h}) \cdot 1(X_i \geq c) \tag{10}$$

If $K(x)$ represents a rectangle kernel function, by estimating Eqs. (9) and (10), we can see that the LLR estimator is the difference between the two fitted values at cutoff point c as Eq. (11):

$$\tau_{LATE} = [\hat{\alpha}_r + \hat{\beta}_r \cdot (c - c)] - [\hat{\alpha}_l + \hat{\beta}_l \cdot (c - c)] = \hat{\alpha}_r - \hat{\alpha}_l \tag{11}$$

As long as the bandwidth is narrow enough, the LLR is a optimal linear approximation to the conditional expectation function. However, when the bandwidth needs to be widen for richer samples, LPR can be used to capture the high-order nonlinear relationship between Y_i and X_i. Similar with LLR, LPR run regressions respectively on both sides of cutoff point, as Eqs. (12) and (13):

$$\min_{b_l} \sum_{i=1}^{N} (Y_i - b_l^T x)^2 \cdot K(\frac{x_i}{h}) \cdot 1(x_i < 0) \tag{12}$$

$$\min_{b_r} \sum_{i=1}^{N} (Y_i - b_r^T x)^2 \cdot K(\frac{x_i}{h}) \cdot 1(x_i \geq 0) \tag{13}$$

where $x_i = X_i\text{-}c$, $x = (1, x_i, x_i^2, x_i^3, \ldots, x_i^p)$, $b_j = (b_{j1}, b_{j2}, \ldots, b_{jp})$, $j = r, l$. The LPR estimator can be expressed as Eq. (14):

$$\tau_{LATE} = \hat{b}_{r0} - \hat{b}_{l0} \tag{14}$$

In Fuzzy RDD, the above methods are still applicable, but Y_i need to be regressed respectively on X_i and T_i, from which the ratio of two parameters is the estimator of Fuzzy RDD. In addition, if the two regressions both adopt rectangular kernel function and same bandwidth, the Fuzzy RDD can be estimated by 2SLS method with T_i as the instrument variable of D_i (Angrist & Pischke, 2008; Cook, 2008).

Validity and robustness test

Since bandwidth selection is a trade-off between estimation bias and variance, the validity and robustness of RDD rely on the bandwidth. In early researches, the cross validation (CV) method (Imbens & Lemieux, 2008; Ludwig & Miller, 2007) are commonly used to estimate the optimal bandwidth. Here we apply the improved CV methods (Calonico, Cattaneo, & Titiunik, 2014; Imbens & Kalyanaraman, 2012) to obtain more credible estimators.

After estimation, some tests for validity and robustness are necessary. First, the pseudo outcome is used to test the assumption of continuity; second, the probability density function of X_i can be used to test the assumption of local randomization (McCrary, 2008); third, the pseudo cutoff point is used to test whether the treatment effects are confounded with other unobserved factors; fourth, the robustness of estimation and the sensitivity of bandwidth selection can be judged by using different bandwidths to estimate RDD.

Strength and progress

Randomized experiment is a gold standard for causal inferences (Angrist & Pischke, 2008; Imbens & Wooldridge, 2009). Because RDD is close to a randomized experiment, the causal inferences by RDD are more credible than those based on other strategies, such as difference-in-differences and instrumental variables (Lee & Lemieux, 2010). In particular, first, RDD requires mild assumptions compared to other non-experimental approaches (Hahn et al., 2001). Second, RDD doesn't require that the assignment to treatment is

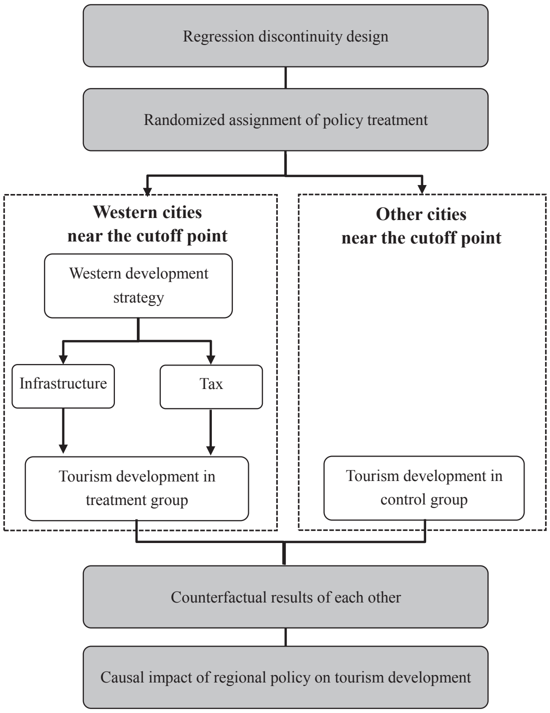

Fig. 3. Conceptual model of causal inference by RDD.

completely randomized, as long as individuals can't precisely control the running variable (Lee, 2008). Third, RDD can be naturally tested using simple graphs, which ensures its credibility and transparency (Lee & Lemieux, 2010). In the light of emerging problems, RDD is being constantly improved, such as RDD with multiple running variables (Caliendo, Tatsiramos, & Uhlendorff, 2013), RDD with measurement error in running variable (Pei & Shen, 2016), quantile treatment effects in RDD (Frandsen, Frölich, & Melly, 2012), RDD with sample selection (Dong, 2017) and regression kink design (Card, Lee, Pei, & Weber, 2015).

Conceptual model of causal inference by RDD

This part presents a conceptual model of causal inference by RDD, which describe the specific application of the spatial RDD in current context.

As shown in Fig. 3, the solid arrows indicate the action relation; the solid lines indicate the progressive relationship; the white boxes represent the factors or variables of this empirical model; and the dark boxes represent the components of the identification process. The two dotted boxes represent, respectively, two types of policy environments separated by a cut-off point, namely, the geographic boundary between the western and central provinces in China. The dotted box on the left represents the western cities affected by the WDS near the cut-off point. The dotted box on the right represents other cities not affected by the WDS near the cut-off point.

The paper aims to investigate the causal impact of regional policy on tourism development. Due to other related extant policies and confounding factors, the causal inference based on conventional econometric methods tends to be biased. Therefore, based on the quasi-natural experiment evidence provided by the WDS in China, the paper uses a spatial RDD method to eliminate confoundings and identify the real causal impact of regional policy on tourism development.

Based on the natural experiment that the assignment of policy treatment is approximately randomised near the two sides of the cut-off point, the spatial RDD chooses the samples near the cut-off point to ensure there is no systematic difference between the two groups of samples before the WDS is proposed. Then, the right group (control group) can provide counterfactual evidence for the left group (treatment group), which is the core condition of causal inference. In addition, according to the policy background of the WDS and the literature, regional policies may affect tourism development mainly through the improvement of infrastructure and tax incentives, which this paper examines in the mechanism analysis.

Econometric model

To investigate the causal effect of regional policies on tourism development, the paper first regress $\Delta Revenue$ on *West* by using OLS, as shown in Eq. (15). We then use this result as benchmark for comparison, which is used to make the initial identification regarding the causal effect and improve the main empirical model.

$$\Delta Revenue_{\mathrm{i}} = \alpha + \beta\, West_{\mathrm{i}} + \lambda Z_{\mathrm{i}} + \mu_{\mathrm{i}} + \varepsilon_{\mathrm{i}} \tag{15}$$

ΔRevenue refers to the outcome variable, measured by the increase in the percentage of total tourism revenue to GDP from 2002 to 2010. *West* refers to the treatment variable. When a city benefits from WDS, *West* equals 1 and 0 otherwise. Z_i represents a set of covariates for avoiding the potential confounding bias. β is the estimator that we are interested in, severed as an initial judgement on the genuine causal effect. μ represents the provincial fixed effect and is for controlling the unobserved environmental factors affecting tourism development in the city. ε refers to the random disturbance term.

Notably, the OLS regression cannot control the influence of unobservable or unobserved factors, such as natural resource endowment, local customs, moral values and religious culture, which are key factors affecting tourism development. In this situation, β will over or underestimate the causal effect of WDS on tourism development in the city. Notably, although OLS may cause estimation bias, the results serve as a benchmark or reference and help improve the validity of econometric modelling.

To avoid estimation bias and identify the true causal effect of the WDS on tourism development, the paper uses the RDD method (Lee and Lemieux, 2010) to identify the regional policy effect on tourism development, with the WDS as a natural experiment. Taking WDS as a natural experiment, the paper uses a spatial RDD to investigate the causal effects of the WDS on tourism development by referring to Almond, Chen, Greenstone, and Li (2009). Specifically, the empirical strategy in the paper considers geographic boundaries between the central and western provinces as the cut-off point. The sample cities on the left side of the cut-off point belong to the RDD treatment group (and the three cities aforementioned), and the sample cities on the right side of the cut-off point belong to the RDD control. In terms of the identification condition, because we use the cities as the sample and constant location as the running variable, the observed characteristics of the samples near the cut-off point can be located on the map. This trait guarantees the validity of the spatial RDD and is one of the advantages of the model (Keele & Titiunik, 2015). In practice, the applicability and validity of an RDD can be tested by a simple graphical method.

Notably, the cut-off point in our context is denoted by the geographic boundary between the middle and western provinces in China. As shown in Fig. 2, grey areas represent the sample cities of the treatment group, and the white areas represent the control group; specifically, the boundary between the grey and white areas is the cut-off point of our model. Our empirical model of a spatial RDD is denoted as Eq. (16). *distance* refers to running variable and is denoted by a shortest linear distance from a sample city to the cutoff point. *F*(*distance*) represents the polynomial terms of the running variable and is used to control the unobservable confounding factors on both sides of the cutoff point. τ is the treatment effect we are interested in.

Regarding the model pattern, because the treatment of the WDS is entirely dependent on the location of cities, we adopt Sharp RDD in the paper. Z_i represents a set of control variables, including some predetermined variables of urban characteristics that likely affect tourism development. According to the identification condition, the effectiveness of RDD method is not dependent on CIA, namely, controlling other confoundings is unnecessary. Notably, introducing X is conductive to improving the estimation efficiency of RDD.

$$\Delta Revenue_{\mathrm{i}} = \theta + \tau West_{\mathrm{i}} + \delta F(distance) + \gamma Z_{\mathrm{i}} + \mu_{\mathrm{i}} + \varepsilon_{\mathrm{i}} \tag{16}$$

Variable and data

This paper chose 2002–2010[2] as the research period. Considering data availability, this paper chose 285 cities in China as the whole samples. Cities located within the scope of the policy belong to the treatment group, and other cities belong to the control group. In particular, according to the principle of an RDD, the practical estimation process will extract city samples near the cut-off point from the whole samples as the effective samples according to the optimal bandwidth. Table 1 summarises all the variables used in this paper.

The main data sources are *China Statistical Yearbook for Regional Economy* and *China City Statistical Yearbook*. The GDP deflator index was used to eliminate the influence of price (based on the price in 2000). Per capita GDP, per capita fiscal revenue, road network density, and population size take logarithms. The statistical characteristics of variables are shown in Table 2.

Empirical results

OLS estimates

Table 3 summarises the benchmark results of OLS estimates under different bandwidth and control variables. Columns (1)–(5) control all the urban characteristics and provincial fixed effects, but the sample range (bandwidth) is different. Columns (6) and (7) release the control vectors in the full sample.

As presented in column (1) under the full sample, regional policy makes the percentage of total tourism revenue to GDP in western cities 8.9% higher than that in other cities. This finding indicates that the regional policy has significantly promoted tourism development, after controlling the observable influence factors. With the gradual reduction of the sample range, the results from columns (2)–(5) still show that the regional policy has positive effects on tourism development, and the estimated results are relatively robust. After releasing the provincial fixed effect, column (6) indicates that although the endogeneity leads to the decrease of

[2] Although the WDS plan was proposed in 2000, the implementation scope was finally formalized in September 2001 by the State Council.

Table 1
Variable definitions.

Outcome variable	Symbol	Definition/Unit
Tourism development	*Revenue*	Differences in percentage of total tourism revenue to GDP (%)
Tourists growth	*Tourist*	Average annual growth rate of total tourist (%)
Infrastructure investment	*infrastructure*	Differences in per capita invest in municipal infrastructure (%)
Tax burden	*Tax*	Differences in percentage of total tax to GDP (%)
Treatment variable		
Regional policy dummy	*West*	1, cities in policy scope; 0, otherwise
Running variable		
Geographical distance	*Distance*	The shortest linear distance from the city to the cutoff point (km)
Predetermined variable		
Per capita GDP	*Pgdp*	GDP/total population (Yuan)
Per capita fiscal revenue	*Govern*	Fiscal revenue/total population (Yuan)
Industrial structure	*Industry*	Tertiary Industry add value / GDP (%)
Road density	*Road*	Road mileage/area (km/km^2)
Infrastructure	*Infrastructure*	Infrastructure investment/GDP (%)
Tourism resources	*Resource*	the number of 5A and 4A attractions
Per capita park area	*Park*	Total area of urban parks/total population (hectare/10000 people)
Population size	*Psize*	Total population
Government size	*Gsize*	Government expenditure/GDP (%)
Investment in fixed assets	*Invest*	Total investment in fixed assets/GDP (%)
Education	*Edu*	Number of students in middle school/number of middle school teacher
Medicine facility	*Medicine*	Number of hospital bed/total population
Provincial fixed effect	*fixed*	Provincial dummy

Note: All predetermined variables control for urban characteristics in 2000.

Table 2
Variables' summary statistics.

Variable	Observations	Mean	Standard deviation	Min	Max
Revenue2002-2010	285	0.030	0.060	−0.340	0.260
Revenue2002-2013	285	0.070	0.080	−0.350	0.490
Tourist	285	0.190	0.090	−0.200	0.590
Infrastructure	285	0.430	1.080	−3.090	3.300
Tax	285	0.020	0.020	−0.060	0.080
West	285	0.300	0.460	0	1
Distance	285	335.2	296.1	13	2100
Pgdp	285	8338	7029	1160	42498
Govern	285	507.7	1179	46.55	18013
Industry	285	0.350	0.080	0.090	0.720
Road	285	0.370	0.240	0.030	2.890
Infrastructure	285	0.010	0.010	0.001	0.090
Resource	285	2.740	4.250	0	41
Park	285	0.890	1.570	0.005	21.08
Psize	285	403.4	282.0	15.97	3091
Gsize	285	0.090	0.040	0.030	0.350
Invest	285	0.290	0.100	0.110	0.860
Edu	285	18.52	3.160	9.690	32.13
Medicine	285	0.280	0.130	0.060	0.860

Table 3
Effect of regional policy on tourism – OLS estimates.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample range	All	750 km	500 km	250 km	100 km	All	All
West	0.089*** (0.016)	0.090*** (0.019)	0.108*** (0.022)	0.110*** (0.018)	0.123*** (0.026)	0.016* (0.009)	0.014* (0.008)
R-square	0.443	0.432	0.464	0.662	0.765	0.191	0.013
Controls	YES	YES	YES	YES	YES	YES	NO
Provincial FE	YES	YES	YES	YES	YES	NO	NO
Observations	285	259	214	145	70	285	285

Notes: Standard errors are in parentheses. Significance levels: ***1%, **5%, *10%.

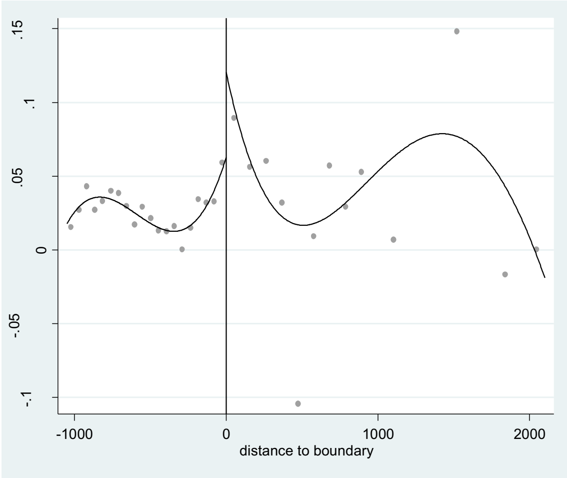

Fig. 4. Distribution of the outcome variable near the cut-off point. *Notes*: The positive value denotes the location of western city, and the negative value denotes locations of other cities. Each point represents the mean of outcome variable in each bin.

the estimated coefficient, the influence of the regional policy on tourism economy remains significant. After continuing to release the urban feature variables, column (7) indicates that the positive regional policy effect still exists. The decrease of this effect may be due to the negative correlation between urban characteristics and policy treatment, namely, the development of a western city is relatively backward, which is in line with the practical situation. In summary, after controlling the observable urban characteristics and the provincial fixed effect, the regional policy has a significant positive effect on tourism development. We observe that the results of benchmark regression were consistent with the expectation.

RDD estimates

Applicability test

Two conditions should be satisfied before using an RDD: (1) the individual cannot manipulate the running variable precisely, to ensure the randomness of policy treatment assignment; and (2) the distribution of an outcome variable at the cut-off point is discontinuous, and other variables (i.e. control variables) are continuous at the cut-off point to ensure that the estimates of treatment effect are not influenced by the confounding factors (Lee & Lemieux, 2010).

For condition (1), because the sample and running variable are, respectively, the city and geographical distance, the values of these elements around the geographic cut-off point are clear and constant. Therefore, the research design fulfils the conditions of an RDD (Keele & Titiunik, 2015). For condition (2), following the approaches by Lee & Lemieux (2010) and Meng (2013), we use graphical methods to verify the continuity of the outcome variable at the cut-off point. The sample was divided into 20 bins, and the distribution of the mean of the outcome variable in each bin was drawn using scatter plots and polynomial fitting lines. As shown in Fig. 4, the outcome variable is apparently discontinuous at the cut-off points and satisfies the condition of RDD.

Optimal bandwidth estimates

Bandwidth selection is the key to ensuring the robustness of an RDD. In a small bandwidth, the individual differences are small, but the estimation variance is large; otherwise, the opposite is true. Therefore, the bandwidth selection is a trade-off between the estimation robustness and efficiency. The commonly used optimal bandwidth estimation method is the cross validation method (CV) proposed by Imbens & Lemieux (2008). After that, Imbens & Kalyanaraman (2012), Calonico et al. (2014) further improved the CV and proposed two optimal bandwidth estimation methods (IK and CCT). The basic idea of these methods is to minimise the mean square error (MSE); thus, the bandwidths obtained from these methods are called MSE-optimal bandwidths. Calonico, Cattaneo, and Farrell (2018) proposed an optimal bandwidth estimation method through minimising the coverage error and called the bandwidths coverage error optimal bandwidths.

To achieve a robust estimation, we use ten algorithms of the two types of methods to estimate the optimal bandwidths and RDD. As shown in Table 4, the MSE-optimal bandwidths are from 125 to 159, and the coverage error optimal bandwidths are from 94 to 120; both types of estimates are robust, but the latter is smaller than the former. Therefore, we chose different optimal bandwidths to ensure the robustness of the RDD estimation.

RDD estimates

Although the RDD estimation does not rely on the observable influence factors, controlling important factors can make the treatment and control groups more comparable and improve the efficiency of the RDD estimates (Angrist & Pischke, 2008). Regarding optimal bandwidth, because MSE-optimal bandwidths and coverage error optimal bandwidths are respectively similar, we chose one from each bandwidth group, respectively, to run the RDD. In addition, to further test the robustness of the RDD estimation, we

Table 4
Estimation of optimal bandwidths with various methods.

Method		Optimal bandwidth (Unit: km)	
		Left side of cutoff point	Right side of cutoff point
MSE optimal bandwidth	(1)	130.1	130.1
	(2)	124.6	159.0
	(3)	129.6	129.6
	(4)	129.6	129.6
	(5)	129.6	130.1
Coverage error optimal bandwidth	(6)	98.08	98.08
	(7)	93.89	119.8
	(8)	97.71	97.71
	(9)	97.71	97.71
	(10)	97.71	98.08

Notes: Methods (1)–(5) are based on the minimisation of the MSE proposed by Imbens & Lemieux (2008), Imbens & Kalyanaraman (2012) and Calonico et al. (2014). Methods (6)–(10) are based on the minimisation of coverage error proposed by Calonico et al. (2018).

manually chose bandwidths for the RDD estimation near the estimated optimal bandwidths.

Regarding the selection of kernel functions, Imbens and Lemieux (2008) assert there is no significant difference between rectangular kernel function and other kernel functions. Therefore, this paper has adopted the rectangular kernel function. For the polynomial term, because the optimal bandwidth is too small relative to the whole sample range in our study, local linear regression (LLR) is more effective (Angrist & Pischke, 2008). In addition, three types of robust standard error estimators are used in this paper: conventional estimator, bias-corrected estimator, robust estimator.

Table 5 summarises the main empirical results from the RDD estimation. For different bandwidths and robust standard error, the RDD estimates are very robust. The effect of regional policy on the development level of tourism is 5.9–6.7% and statistically significant. The results show that in the first decade of the 21st century, the increase in the percentage of total tourism revenue to GDP in western China is 6% higher than in the other cities due to the influence of regional policies. This is a very significant policy effect, because the average increase in percentage of tourism revenue for all cities in China is 3.1% during the study period. We observe that cities near the both sides of the cut-off point show distinct trends of tourism development due to differing regional policies.

The RDD estimates are basically consistent with the OLS estimates; however, the coefficient of regional policy in RDD is smaller due to the positive bias of the OLS estimator. For example, the omission of regional culture, customs, topography, and other unobservable factors may lead to the positive bias of OLS estimation. Fortunately, the RDD can obtain the consistent estimators through limiting the samples near the cut-off point and randomizing the policy treatment approximately.

To further verify the robustness of RDD estimation, we randomly chose 85 km and 200 km beyond the estimated bandwidth range to run RDD estimates. The results are shown in Table 6. When choosing a smaller bandwidth (85 km), the sample size is reduced, but the cities are more strongly affected by the policy. When the larger bandwidth is selected (200 km), the number of samples increases, but the policy effect shows the distance attenuation. In short, RDD estimations are robust with different bandwidths and robust standard error. The robustness test confirmed again that the regional policy has a significant positive effect on tourism economy.

Validity test

According to the principles of effectiveness proposed by Imbens and Lemieux (2008), this paper uses three commonly used test methods. 1) The fundamental premise of RDD estimation is that the individual cannot absolutely manipulate cut-off points, which have been verified in Section 4.2. 2) The estimated treatment effect may be biased due to confounding factors; thus, the continuity of some important urban features at the cut-off point must be verified. We selected four important urban features variables, namely,

Table 5
Effect of regional policy on tourism – RDD estimates.

Optimal bandwidth	MSE optimal bandwidth			Coverage error optimal bandwidth		
Method	(1) Conventional	(2) Bias-corrected	(3) Robust	(4) Conventional	(5) Bias-corrected	(6) Robust
West	0.062*** (0.024)	0.067*** (0.024)	0.067** (0.028)	0.059* (0.032)	0.063** (0.032)	0.063* (0.033)
Controls	YES	YES	YES	YES	YES	YES
Provincial FE	YES	YES	YES	YES	YES	YES
Polynomial terms	linear	linear	linear	linear	linear	linear
Bandwidth	130	130	130	98	98	98
All observation	285	285	285	285	285	285
Effective observation	87	87	87	63	63	63

Table 6
Effect of regional policy on tourism – other bandwidths.

Bandwidth	85 km			200 km		
Method	(7) Conventional	(8) Bias-corrected	(9) Robust	(10) Conventional	(11) Bias-corrected	(12) Robust
West	0.079** (0.031)	0.084*** (0.031)	0.084*** (0.033)	0.033* (0.019)	0.044** (0.019)	0.044* (0.024)
Controls	YES	YES	YES	YES	YES	YES
Provincial FE	YES	YES	YES	YES	YES	YES
Polynomial terms	linear	linear	linear	linear	linear	linear
Bandwidth	85	85	85	200	200	200
All observation	285	285	285	285	285	285
Effective observation	55	55	55	124	124	124

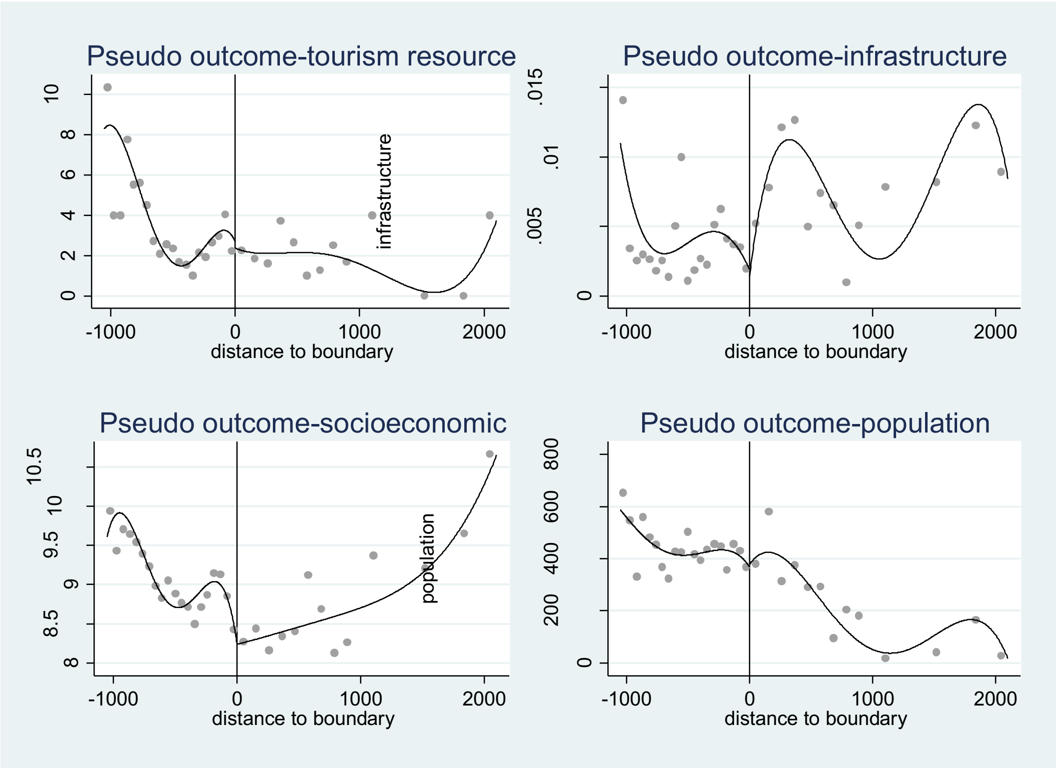

Fig. 5. Distribution of pseudo outcome variables near the cut-off point.

tourism resources, infrastructure, economic development, and population size, to run the pseudo outcome test. Firstly, in Fig. 5, four pseudo outcome variables do not obviously 'jump' at the cut-off point, which preliminarily verified the reliability of the main regression results. Further, in Table 7, all four pseudo outcome estimations are nonsignificant. These results imply that urban characteristics possibly affecting the policy effect have no systematic differences near the cut-off point and thus they will not cause confounding bias.

Table 7
Pseudo outcome test for the RDD estimation.

Outcome variables	*resources*	*infrastructure*	*economy*	*population*
(1) Conventional	1.955 (2.001)	0.00153 (0.00267)	−0.153 (0.0950)	0.0102 (0.218)
(2) Bias-corrected	1.833 (2.001)	0.000774 (0.00267)	−0.0846 (0.0950)	−0.0445 (0.218)
(3) Robust	1.833 (2.187)	0.000774 (0.00311)	−0.0846 (0.119)	−0.0445 (0.241)
Bandwidth	147.2	95.4	87.6	116.7
All observation	285	285	285	285
Effective observation	91	61	56	79

Notes: Each regression is run with the given pseudo outcome variable. Each regression takes the optimal bandwidth respectively.

Table 8
Pseudo treatment test for the RDD estimation.

Cutoff point	−75 km	75 km	−100 km	100 km	−175 km	175 km	−250 km	250 km
(1) Conventional	0.005 (0.013)	−0.002 (0.023)	−0.020 (0.012)	0.032 (0.027)	−0.017 (0.013)	0.031 (0.033)	0.009 (0.018)	0.025 (0.027)
(2) Bias-corrected	0.010 (0.013)	−0.008 (0.023)	−0.018 (0.012)	0.029 (0.027)	−0.022 (0.012)	0.015 (0.033)	0.002 (0.018)	0.029 (0.026)
(3) Robust	0.010 (0.015)	−0.008 (0.025)	−0.018 (0.014)	0.029 (0.029)	−0.022 (0.014)	0.015 (0.033)	0.002 (0.021)	0.029 (0.029)
Effective observations	129	71	40	40	94	48	58	45
Bandwidth	188.1	151.0	113.9	113.8	121.9	136.9	104.9	165.7

Notes: Each regression is run with the given pseudo cut-off point. The positive values of cut-off point denote the location of the pseudo cut-off point on the left side of the true cut-off point, and the negative values denote the cut-off points' location on the right side.

3) There may be 'jump' at other pseudo cut-off points, that is, the regional policy effect may be a pseudo treatment effect. This paper randomly selected eight false cut-off points around the true cut-off point for pseudo treatment test. As shown in Table 8, all pseudo treatment estimations are nonsignificant, indicating that the 'jump' of outcome variable at the true cut-off point (i.e., the WDS boundary) is induced by the regional policy.

Robustness test

The main regression results show that the regional policy has significantly positive effects on tourism development, and the estimators are robust. To further explore this causal effect, this section uses, respectively, the increase in percentage of total tourism revenue to GDP from 2002 to 2013 (a wider range of time) and the average annual growth of tourists from 2002 to 2010 as the outcome variables to run the RDD estimation again. First, as presented in Fig. 6, both outcome variables are apparently discontinuous at the cut-off point. Second, as shown in Table 9, the regression results for two tourism development indicators are consistent with the main regression results; thus, the positive impact of regional policy on tourism development is confirmed again.

Firstly, the long-term effect of the WDS on the increase in percentage of total tourism revenue to GDP is 7.9–8.6%, indicating that the regional policy has a sustained effect on the tourism economy. Moreover, the larger estimation coefficient indicates that the regional policy effect does not show time attenuation during the study period. Secondly, the effect of the regional policy on the average annual growth of tourists is 6.6–6.9%, indicating that the regional policy enhances the attraction for tourists, for example, the regional policy has improved urban infrastructure. In summary, the regional policy may promote long-term tourism development, and this also proves the robustness of the main results of this paper.

Mechanism analysis of the regional policy effect

The regression results show that the WDS has positive effects on multiple indicators of tourism development, indicating that regional policies may affect tourism development through various channels. According to the planning of the WDS, the programme mainly promotes regional development through infrastructure improvement, tax incentives, and transfer payments. Infrastructure improvement and tax incentives mainly work for urban and enterprise development, and the main beneficiaries of transfer payments are local residents.

Therefore, Section 4.5 analyses the influence mechanism of regional policies on tourism development from two aspects: infrastructure and tax. We conducted an RDD estimation considering the change of per capita municipal infrastructure investment and

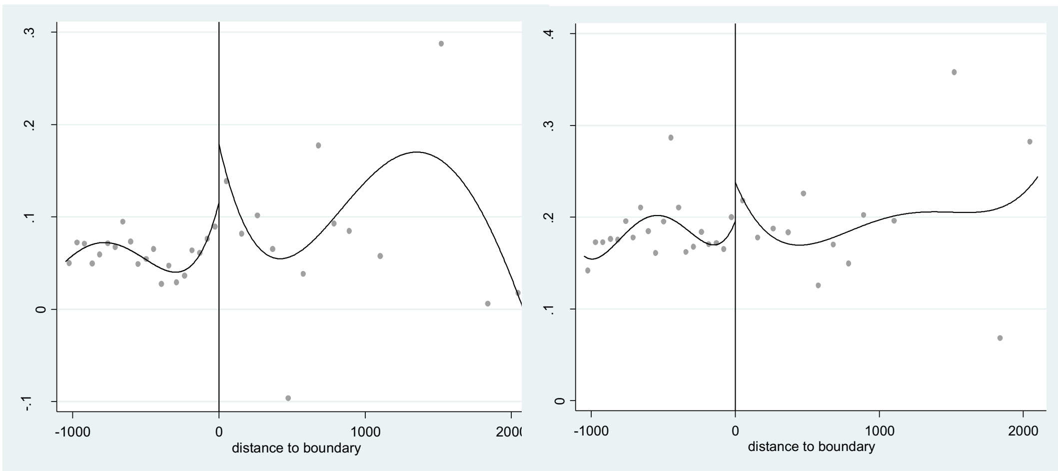

Fig. 6. Distribution of other outcome variables near the cut-off point.

Table 9
Extensive effect of regional policy on tourism.

Outcome variable	*Tourist*			*Revenue2013*		
Method	Conventional	Bias corrected	Robust	Conventional	Bias corrected	Robust
West	0.0685* (0.0351)	0.0663* (0.0351)	0.0663* (0.0403)	0.0793** (0.0341)	0.0863** (0.0341)	0.0863** (0.0395)
Controls	YES	YES	YES	YES	YES	YES
Provincial FE	YES	YES	YES	YES	YES	YES
Polynomial terms	linear	linear	linear	linear	linear	linear
Bandwidth	190	190	190	125	125	125
All observation	285	285	285	285	285	285
Effective observation	121	121	121	85	85	85

Table 10
Mechanism analysis of regional policy effect.

Outcome variable	*Infrastructure*			*Tax*		
Method	Conventional	Bias corrected	Robust	Conventional	Bias corrected	Robust
West	0.712* (0.402)	0.803** (0.402)	0.803* (0.454)	−0.014** (0.006)	−0.015** (0.006)	−0.015** (0.007)
Controls	YES	YES	YES	YES	YES	YES
Provincial FE	YES	YES	YES	YES	YES	YES
Polynomial terms	linear	linear	linear	linear	linear	linear
Bandwidth	188	188	188	154	154	154
All observation	285	285	285	285	285	285
Effective observation	110	110	110	99	99	99

urban tax burden from 2002 to 2010 as the outcome variables, respectively. We observe that the results are robust in Table 10. First, the WDS has a significant positive effect on municipal infrastructure investment, indicating that regional policies may enhance the attractiveness and competitiveness of cities‘ tourism industry by improving the urban infrastructure. Second, the WDS has a significant negative effect on urban tax burden, indicating that regional policies may improve the competitiveness of tourism enterprises by reducing the tax burden.

Conclusions and discussions

Taking China's WDS as a natural experiment, the paper adopted a spatial RDD method to estimate the causal effect of the regional policy on tourism development. Adequate tests have been conducted to confirm the robustness of the empirical results. The major conclusions are summarised as follows. First, the WDS has caused a significant positive effect on tourism development. Compared with other cities, western cities affected by the WDS have higher increase in the percentage of tourism revenue, estimated to be up to 6% and varying from 5.9% to 6.7%. Relative to the average increase for all the sample cities in China, this estimated policy effect is extremely significant. Second, as a comprehensive public policy, the WDS has a sustained a promoting effect on tourism development in the long run. Third, regarding the indicators of tourism development, the WDS enhances the level of tourism specialisation, as indicated by the percentage of tourism revenue, and increases the number of tourists. Finally, the WDS promotes tourism development through infrastructure improvement and tax incentives. In summary, improvements to a city's infrastructure can enhance the attractiveness of a city's tourism, and tax burden reductions can improve the competitiveness of tourism enterprises.

This paper has provided an evaluation framework for identifying the causal effects of regional policy on tourism in developing countries. We now present additional discussions on our evaluation framework for further studies. As for our methodology, this paper attempts to analyse how policy affects tourism development from the perspective of regional policy. Compared with specialised tourism policy, regional policy may have a more complex and continuous impact on tourism development. As summarized, many of the quantitative results in the literature have been based on simple statistical analysis or case evidence, and these types of results are often disturbed by confounding factors. The main purpose of our empirical strategy is to eliminate this confounding bias. The estimated coefficients based on rigorous identification strategy and statistical data in the paper are a type of net effect.

An RDD is the quasi-experiment method most similar to a randomised experiment. The paper is the first attempt to use an RDD method in tourism studies, in which the adequate validity tests have confirmed the applicability of the RDD method in tourism research. Therefore, the paper provides a new, valid empirical analysis framework for future tourism research. Notably, this paper takes the geographic boundary as the spatial cut-off point for the RDD estimation, which is only one of the specific applications of an RDD in tourism research. The flexibility and applicability of an RDD make it an important method to identify causal effects in tourism empirical research. For example, taking the tourists' legal retirement age as a cut-off point, we can use RDD method to estimate how retirement policy affects tourists' travel choices.

In practice, as has been emphasised in the literature, tourism-related policy is a complex system affected by other policies and economic and social factors. Although identifying the impact of a single policy on tourism development is possible, the joint impact of various policies on tourism development in the long term is often not easily identified. The WDS is a large-scale regional policy implemented in China. China's empirical research conclusion has good external validity for other developing countries. Developing countries experiencing industrialisation and urbanisation may incur unbalanced regional development and resources misallocation across space. China's experiences suggest that tourism may provide an effective industrial support for regional policy solutions to those problems.

Acknowledgements

This work was supported by MOE (Ministry of Education of the People’s Republic of China) Project of Humanities and Social Sciences (Project No. 17YJC790021) and the Innovation Fund of Shanghai University of Finance and Economics for Postgraduate (NO. 2018110645).

References

Alipour, H., & Kilic, H. (2005). An institutional appraisal of tourism development and planning: The case of the Turkish Republic of North Cyprus. *Tourism Management, 26*(1), 79–94.
Almond, D., Chen, Y., Greenstone, M., & Li, H. (2009). Winter heating or clean air? Unintended impacts of China’s Huai river policy. *American Economic Review, 99*(2), 184–190.
Angrist, J. D., & Lavy, V. (1999). Using maimonides’ rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics, 114*(2), 533–575.
Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist’s companion.* Princeton University Press.
Baldwin, R., & Okubo, T. (2010). Tax reform, delocation, and heterogeneous firms. *Scandinavian Journal of Economics, 111*(4), 741–764.
Battistin, E., & Rettore, E. (2008). Ineligibles and eligible non-participants as a double comparison group in regression-discontinuity designs. *Journal of Econometrics, 142*(2), 715–730.
Baum, T., & Szivas, E. (2008). HRD in tourism: A role for government? *Tourism Management, 29*(4), 783–794.
Bloom, N., Griffith, R., & Reenen, J. V. (2002). Do R&D tax credits work? Evidence from a panel of countries 1979–1997. *Journal of Public Economics, 85*(1), 1–31.
Caliendo, M., Tatsiramos, K., & Uhlendorff, A. (2013). Benefit duration, unemployment duration and job match quality: A regression-discontinuity approach. *Journal of Applied Econometrics, 28*(4), 604–627.
Calonico, S., Cattaneo, M. D., & Farrell, M. H. (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. *Journal of the American Statistical Association,* 1–13.
Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica, 82*(6), 2295–2326.
Card, D., Lee, D. S., Pei, Z., & Weber, A. (2015). Inference on causal effects in a generalized regression kink design. *Econometrica, 83*(6), 2453–2483.
Chay, K., McEwan, P., & Urquiola, M. (2005). The central role of noise in evaluating interventions that use test scores to rank schools. *American Economic Review, 95*(4), 1237–1258.
Chow, W. S. (1988). Open policy and tourism between Guangdong and Hong Kong. *Annals of Tourism Research, 15*(2), 205–218.
Cook, T. D. (2008). “Waiting for life to arrive”: A history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics, 142*(2), 636–654.
DiNardo, J., & Lee, D. S. (2004). Economic impacts of new unionization on private sector employers: 1984–2001. *Quarterly Journal of Economics, 119*(4), 1383–1441.
Dong, Y. (2017). Regression discontinuity designs with sample selection. *Journal of Business & Economic Statistics,* 1–16.
Dredge, D. (2006). Policy networks and the local organisation of tourism. *Tourism Management, 27*(2), 269–280.
Fan, S., Kanbur, R., & Zhang, X. (2011). China’s regional disparities: Experience and policy. *Review of Development Finance, 1*, 47–56.
Frandsen, B. R., Frölich, M., & Melly, B. (2012). Quantile treatment effects in the regression discontinuity design. *Journal of Econometrics, 168*(2), 382–395.
Gülcan, Y., Kuştepeli, Y., & Akgüngör, S. (2009). Public policies and development of the tourism industry in the Aegean region. *European Planning Studies, 17*(10), 1509–1523.
Hahn, J., Todd, P., & Klaauw, W. V. D. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica, 69*(1), 201–209.
He, J. M. (2013). Research on mechanism and impact of replacing business tax by value added tax on China’s tourism service industry. *Tourism Science, 27*(1), 29–40 (in Chinese).
Holland, P. W. (1986). Statistics and causal inference: Rejoinder. *Journal of the American Statistical Association, 81*(396), 968–970.
Hong, J. J., Liu, Z. Q., & Huang, W. (2014). Regional revitalization strategies and the industrial spatial structure changes in China. *Economic Research Journal, 49*(8), 28–40 (in Chinese).
Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies, 79*(3), 933–959.
Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity design: A guide to practice. *Journal of Econometrics, 142*(2), 615–635.
Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature, 47*(1), 1–81.
Imikan, A. M., & Ekpo, K. J. (2012). Infrastructure and tourism development in Nigeria: A case study of rivers state. *International Journal of Economic Development Research and Investment, 3*(2), 53–60.
Jackson, J. (2006). Developing regional tourism in China: The potential for activating business clusters in a socialist market economy. *Tourism Management, 27*, 695–706.
Keele, L. J., & Titiunik, R. (2015). Geographic boundaries as regression discontinuities. *Political Analysis, 23*(1), 127–155.
Khadaroo, J., & Seetanah, B. (2007). Transport infrastructure and tourism development. *Annals of Tourism Research, 34*(4), 1021–1032.
Krutwaysho, O., & Bramwell, B. (2010). Tourism policy implementation and society. *Annals of Tourism Research, 37*(3), 670–691.
Law, R., Leung, R., & Buhalis, D. (2009). Information technology applications in hospitality and tourism: A review of publications from 2005 to 2007. *Journal of Travel & Tourism Marketing, 26*(5–6), 599–623.
Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. house elections. *Journal of Econometrics, 142*(2), 675–697.
Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature, 48*(2), 281–355.
Lejárraga, I., & Walkenhorst, P. (2013). Economic policy, tourism trade and productive diversification. *International Economics, 135*, 1–12.
Li, H., Chen, J. L., Li, G., & Goh, C. (2016). Tourism and regional income inequality: Evidence from china. *Annals of Tourism Research, 58*, 81–99.
Liu, A., & Wall, G. (2005). Human resources development in China. *Annals of Tourism Research, 32*(3), 689–710.
Ludwig, J., & Miller, D. L. (2007). Does head start improve children’s life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics, 122*(1), 159–208.
Ma, Z. Y. (2001). On integrating poverty alleviation into framework of tourism development in the Project of Western China’s Development. *Journal of Natural Resources, 16*(2), 191–195 (in Chinese).
Mccrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics, 142*(2), 698–714.
Mcewan, P. J., & Shapiro, J. S. (2008). The benefits of delayed primary school enrollment: Discontinuity estimates using exact birth dates. *Journal of Human Resources, 43*(1), 1–29.
Meethan, K. (1998). New tourism for old? Policy developments in Cornwall and Devon. *Tourism Management, 19*(6), 583–593.
Meng, L. (2013). Evaluating China’s poverty alleviation program: A regression discontinuity approach. *Journal of Public Economics, 101*(1), 1–11.

Pearce, D. G. (1998). Tourism development in Paris: Public intervention. *Annals of Tourism Research, 25*(2), 457–476.
Pei, Z., & Shen, Y. (2016). The devil is in the tails: Regression discontinuity design with measurement error in the assignment variable. *Working Papers, 106*(1), 75–101.
Richter, L. R. (1983). Political implications of Chinese tourism policy. *Annals of Tourism Research, 10*(3), 395–413.
Rin, M. D., Giacomo, M. D., & Sembenelli, A. (2009). Entrepreneurship, firm entry, and the taxation of corporate income: Evidence from Europe. *Journal of Public Economics, 95*(9–10), 1048–1066.
Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*(5), 688.
Sofield, T. H. B., & Li, S. (1998). Tourism development and cultural policies in China. *Annals of Tourism Research, 25*(2), 362–392.
Song, Y. (2013). Rising Chinese regional income inequality: The role of fiscal decentralization. *China Economic Review, 27*(1), 215–231.
Thapa, B. (2012). Soft-infrastructure in tourism development in developing countries. *Annals of Tourism Research, 39*(3), 1705–1710.
Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology, 51*(6), 309–317.
Trochim, W. (1984). *Research design for program evaluation: The regression-discontinuity design.* Beverly Hills, CA, USA: Sage Publications.
Vernon, J., Essex, S., Pinder, D., & Curry, K. (2005). Collaborative policymaking – Local sustainable projects. *Annals of Tourism Research, 32*(2), 325–345.
Wang, D., & Ap, J. (2013). Factors affecting tourism policy implementation: A conceptual framework and a case study in China. *Tourism Management, 36*(6), 221–233.
Zeng, B., & Ryan, C. (2012). Assisting the poor in china through tourism development: A review of research. *Tourism Management, 33*(2), 239–248.
Zhang, H. Q., Chong, K., & Ap, J. (1999). An analysis of tourism policy development in modern China. *Tourism Management, 20*(4), 471–485.

Taotao Deng is an associate professor in Regional Economics at Shanghai University of Finance and Economics. He received his PhD from University of Aberdeen, UK. His research interests lie in urban and regional studies, tourism economics and transport economics. As the lead author, he has authored 12 scholarly papers indexed by SSCI. His publications have appeared in *Tourism Economics*, *Asia Pacific Journal of Tourism Research, Transportation*, *Transport Reviews*.