Literary Destination Familiarity and Inbound Tourism: Evidence from Mainland China

Guodong Ju, Jiankun Liu, Guangye He, Xinyi Zhang*, and Fei Yan*

Abstract: Destination familiarity is an important non-economic determinant of tourists' destination choice that has not been adequately studied. This study posits a literary dimension to the concept of destination familiarity—that is, the extent to which tourists have gained familiarity with a given destination through literature—and seeks to investigate the impact of this form of familiarity on inbound tourism to Mainland China. Employing the English fiction dataset of the Google Books corpus, the New York Times annotated corpus, and the Time magazine corpus, we construct two types of destination familiarity based on literary texts: affection-based destination familiarity and knowledge-based destination familiarity. The results from dynamic panel estimation (1994–2004) demonstrate that the higher the degree of affection-based destination familiarity with a province in the previous year, the larger the number of inbound tourists the following year. Examining the influence of literature and its consumption on tourism activities sheds light on the dynamics of sustainable tourism development in emerging markets.

Key words: inbound tourism; tourist; destination familiarity; Mainland China; Google Books

1 Introduction

Over the past two decades, globalization has made tourism a popular global leisure activity, and international tourism has become one of the largest and fastest growing industries in the world.

- Guodong Ju is with the Department of Social Policy, London School of Economics and Politial Science, London, WC2A 2AE, UK. E-mail: G.Ju@lse.ac.uk.
- Jiankun Liu is with the Department of Sociology, Harbin Engineering University, Harbin 150001, China. E-mail: jkliu@hrbeu.edu.cn.
- Guangye He is with the Department of Sociology, Nanjing University, Nanjing 210023, China. E-mail: hgy.gloria@nju.edu.cn.
- Xinyi Zhang is with the Business School, Beijing Language and Culture University, Beijing 100083, China. E-mail: zhangxinyi@blcu.edu.cn.
- Fei Yan is with the Department of Sociology, Tsinghua University, Beijing 100084, China. E-mail: feiyan@tsinghua.edu.cn.
- ∗ To whom correspondence should be addressed.
- † Guodong Ju and Jiankun Liu contribute equally to this paper.
  Manuscript received: 2021-06-21; revised: 2021-07-08; accepted: 2021-07-12

Rather than considering political, socioeconomic, and institutional factors, recent research is increasingly examining the cultural factors that predict tourism development. One strand of research links tourists' destination choices with Hofstede (1980)[1]'s "cultural distance", which measures how cultural dimensions in one society differ from those of another[2−9]. Tourism is expected to be higher between host countries/regions and home countries/regions with greater cultural proximity, as the smaller the culture distance, the lower the barriers for tourists to adapt to the destination region.

Another strand of research examines literary tourism, a type of cultural tourism to literary destinations associated with famous writers in their real lives or/and associated with the settings in their novels[10, 11]. Literary tourism hinges on the intimate relationship between books and travel, and the scholarship in this area argues that tourists could be inspired by specific literary texts to engage with the same landscape as the author or the literary characters did[12]. Compared with Hofstede's cultural distance, literary tourism not only captures more detailed cultural variation within a country—which is

particularly important in a large country with significant regional differences such as China[13−16]—it also takes into account the affection-based experiences generally ignored by "cultural distance" scholars. Cultural distance can be more objectively measured, whereas affection is experienced subjectively, but both factors may motivate individuals to travel, even when their understandings and expectations of the destination culture may be inaccurate. In many cases, affection may motivate individuals to embark on the visit even when their understandings or interpretation of the culture of that specific place is not entirely right, in particular, prior to tourists' first visit. For example, a vivid fictional love story may arouse the readers' potent emotions, form the unique literary experience of individuals, and further influence their understanding of the destination culture[5], their desire to connect with literary figures[17] and their attempts to live out the adventures they have read about. Meanwhile, it is the fusion of knowledge from travel guides such as Lonely Planet can provide contemporary instant information on, for example, the route, the transportation method, and fees to help readers find a high-profile Michelin restaurant in the given destination.

Existing studies of literary tourism, however, focus largely on Western countries and often holds divided opinions regarding this form of tourism's value[18], and, more specifically, whether it could promote international tourism to China or at least some places therein (except Refs. [19−21]). Current studies are also largely qualitative case studies that examine the relationship between specific literary works/writers and tourism to specific places based on a specific book and a specific place.

The concept of destination familiarity provides a way to investigate whether and to what extent literary tourism impacts inbound tourism to China. This term explains the key factor that could segment and target certain groups and develop marketing strategy, enabling us to understand how people develop impressions of potential destinations and, in turn, make destination choices[22]. Direct, personal experience is not the only way of gaining familiarity with a locale[23]; indeed, many more people are likely to gain some level of destination familiarity by reading books, news, or magazines, yet, this phenomenon and its impact on actual tourism remain understudied. In this research, we posit a literary

dimension of destination familiarity—that is, experience or information about a tourist destination mediated through literature—and seek to investigate whether and how this form of destination familiarity influences international tourism inflow to specific provinces, states, or municipalities. We chose Mainland China§ as our empirical case because its cultural and socioeconomic development strikingly varies across provinces, and it is both the largest emerging economy and the world's leading destination for inbound tourism, widely estimated to become the biggest inbound market for international tourists by 2030 (The Guardian, November 6, 2018[24]).

While literary tourism is often taken as a form of niche tourism based on sites and events in the novels, some scholars argued that, all tourism is a cultural experience[25, 26], this being said, not only the culture implied from the novels, but also the contemporary information of the destination place, such as, daily news, travel writing, and travel guide should also be taken into account. As such, we identify two different types of literary destination familiarity—affection-based literary familiarity, which results from reading novels and other forms of fiction about a destination; and knowledge-based literary familiarity, which results from reading non-fiction sources of information or knowledge about a destination, such as magazines and travel guides.

To measure affection-based literary destination familiarity quantitatively, we extracted the appearance frequency of Chinese province names from the English-fiction dataset of the Google Books corpus, a text corpora containing a large volume of digitalized English fiction books; likewise, we extracted the same information from the New York Times annotated corpus and the Time magazine corpus to measure knowledge-based literary destination familiarity. In this analysis, we focus exclusively on English-language corpus, as English is the most widely used lingua franca, which plays a critical role in international communication and global policies[27]. Though people might expect that the major source of in-bound tourism to Mainland China is from other Asian countries/regions that is not using English (e.g., Hong Kong, Macau)[28], in fact, English has seen an unparalleled explosion of international

---

§ In this paper, we focus on 31 provinces in Mainland China but exclude Hong Kong, Macau, or Taiwan due to both the availability of accordant statistics and impacts of other potential confounding factors.

activity and achieved a worldwide reach beyond that of any other language[27]. According to the tourism statistics from Ministry of Culture and Tourism of the People's Republic of China, the top 10 source countries of in-bound tourism to Mainland China in the early 2000s were Republic of Korea, Japan, Russia, USA, Malaysia, Singapore, Philippines, Mongolia, Thailand, Australia, with slight changes of rank order each year. Among these countries, half of them are English-speaking countries. For the others, English is the official language. As such, our use of the English-language corpus is appropriate. It should also be noted that such corpus restriction has been widely adopted in many recent studies (e.g., Ref. [29]), providing additional support for this focus.

To further investigate how litery destination familiarity might impact inbound tourism flows for 31 Chinese provinces, we employed dynamic panel data models[30, 31]. Specifically, we matched the destination familiarity index with provincial socioeconomic characteristics to create a provincial panel, which covers 31 provinces over eleven-year period (1994−2004). This allows us to establish the causal linkage to examine the effect of literary destination familiarity on inbound tourism. Our study demonstrates that literary destination familiarity as measured by word-search dynamics in Google Books English fiction corpus significantly affects the inflows of international tourists to Mainland China—namely, the higher the degree of literary destination familiarity with a province in the previous year, the larger the number of inbound tourists the following year.

2   Literary Destination Familiarity

Culture is " the collective programming of the mind which distinguishes the members of one human group from another"[1]. In addition to informing people's beliefs, culture is increasingly being studied as something that can influence a country's economic growth. When it comes to international tourism, cultural distance has been widely investigated, with scholars examining whether and to what extent the value system and cultural affinity between host and home countries impacts the potential flow of tourism from the latter to the former[1, 32−34]. Findings remain divided, however, with some suggesting that tourists prefer to travel to countries with cultural values and practices closely aligned to those of their home country (e.g.,

Refs. [7, 35, 36]), and others suggesting that tourists prefer to travel to countries culturally dissimilar culture from their home countries (e.g., Refs. [37−41]). As cultural distance between destination and home countries is not easily or accurately measured or perceived by tourists, these contradictory findings highlight the likelihood that actual cultural distance is not the only cultural factor that influences potential tourism flow.

In many cases, what matters more are individuals' subjective judgement about a potential destination, a type of perceived cultural value rather than a "real" or objective cultural value associated with specific sites[42−44].

As assumed by Samovar and Porter[45], culture refers not only to the cumulative deposit of knowledge and experience, but also the values and attitudes acquired by a group of people over the course of generations of transfer; these values and attitudes can, in turn, be further cultivated and then exert even more influence over behavior. One such means of cultivation is through the reading of literary texts. By creating an imaginative world, literary texts can shape readers' ideas about places and the local culture associated with those specific sites[46]. Popular literature, in particular, might inspire readers to seek "a physical body to enable their senses to connect with objects read"[47, 48]—in short, to engage in literary tourism.

2.1   Literary tourism

Literary tourism associates specific sites with specific literary works (e.g., Refs. [49−51]). These literary works are read as cultural assets that reflect the cultural values or attitudes of specific sites. But because literary texts are created by authors and then further interpreted by readers, the cultural values reflected by such texts are not simply the objective cultural values of the place, but those values as filtered through imagination. As a result, some differentiate literary tourism is from regular tourism, arguing that the former "deals with travelers rather than tourists" and " with the romantic rather than the collective gaze"[10]. Although travelers and tourists tend to share the same social and economic space, they are separated by imagination, which offers new information and cultural experiences. This is because literary tourism in the anthropological sense " involves tourists and visitors identifying with, discovering, and creating

signifiers of cultural values with those people who have become part of the cultural mythologies of places"[52].

2.2   Literary dimension of destination familiarity

In this research, we propose a literary dimension of destination familiarity and hypothesize a close coupling between literary exposure to a destination and the observed inbound tourism flows to that locale in the context of transnational tourism. Literary destination familiarity refers to having prior knowledge or having received information regarding a given destination via various literary texts. This dimension of familiarity has been largely overlooked in tourism studies[53], yet there are reasons to suspect that when choosing an international tourist destination, people will be more likely to select regions/cities they have at least read about in some form, thus gaining affection and/or knowledge about the place, as opposed to those destinations about which they know or care little. The reasons for this are two-fold: (1) transnational tourism involves risks; and (2) some initial cultural images and prior knowledge of a destination can stimulate interest for a prospective visit[54, 55].

Though language familiarity[56] and food familiarity[57] might also promote behavioral preferences, higher literary experiences to a potential destination would create even higher tolerance towards the potential destination. This is because literary experiences could motivate readers' participation, influence their prior perceptions and emotional recognition of a destination, and generally make them more open to visiting that destination (e.g., Refs. [50, 51, 58].

Given the multidimensional nature of familiarity, scholars have tried to establish various typologies. For example, Baloglu[59] (2001) classified three types of familiarity: informational, experiential, and self-rated. Specifically, informational familiarity refers to the extent of received information; experiential familiarity means the number of previous destination experiences; and self-rated familiarity reflects one's observed and self-judged level of familiarity with a particular place. Prentice[53] (2004) further added another two types —educational and proximate familiarity —to measure the extent of one's educational involvement with a destination and one's exposure to the popular media. However, the literary dimension of familiarity has been overlooked to a large extent. Furthermore,

while previous analyses have focused primarily at the micro individual level, our analysis is conducted at the macro provincial level, as we contend that literature resonates at a cultural and societal level.

For this study, we decompose literary destination familiarity into two components —knowledge-based familiarity and affection-based familiarity. The former emphasizes the destination familiarity and knowledge resulting from contemporary non-fiction forms of writing in news and magazines, whereas the latter emphasizes the degree of destination familiarity that arises when fictional forms of literature produce affection for certain destinations.

3   Data, Variables, and Models

3.1   Data

Media, as a primary cultural instrument, carries symbolic messages of cultural knowledge, and gives shape and structure to a society, influencing people's sense of reality. This is particularly the case since the early twentieth century, when mass media transformed the way people interact and communicate with each other, and facilitated learning about other places through "informal mediated learning" (e.g., reading novels or poems). We use three types of media—namely, books, news, and magazines—to construct a measure of literary destination familiarity. For the book measure, we extract the appearance of Chinese provinces from the Google Books English fiction corpus; for the news and magazines, we extract relevant information from the New York Times corpus and the Time magazine corpus, and take them as a proxy of destination familiarity produced through news and magazines, respectively.

At a regional level, the frequency of appearance in books of a particular place can represent the perceived visibility and attractiveness of that given place to the world audience[29, 60−63]. In particular, works of fiction constitutes an important part of social world, but with more imaginary elements based on general sociocultural realities[64, 65]. These works help shape the image of regions, communities, and individuals[63, 66, 67]. In order to examine whether and how literary destination familiarity impacts the inflow of foreign tourism in Mainland China, we extract the appearance of Chinese provinces in Google English fiction books, the New York Times, and Time magazine using the same

methods as Chen and Yan[61]. We argue that individuals have produced and consumed fiction not only for entertainment, but also to spread knowledge and develop values and beliefs that could, collectively, affect people's travel destination choices.

Our focus on the time period from 1994 to 2004 captures two key cultural phenomena: China commitment to openness in 1990, when the Chinese government announced plans to develop the Pudong area in Shanghai into a dynamic new area; and the development of international tourism in Mainland China entered a stage of maturity stage after 1992[68]. The rapid development during this same period of new media has challenged the sociocultural significance of books. This time period also delimits two practical concerns: some key variables are missing in provinces (especially the inland areas) before 1994; and some books published after 2004 remain undigitized[69], thus limiting the comprehensiveness of literary dimension of destination familiarity extracted from fiction books after this date. A panel of 31 provinces spanning 11 years renders sufficient waves for dynamic panel estimation to achieve efficient estimation.

3.2   Variables

3.2.1   Dependent variable

The dependent variable in this analysis is the annual number of foreign tourists in each province from 1993 to 2005, obtained from the Yearbook of China Tourism Statistics, published jointly by China National Tourism Administration and the National Bureau of Statistics. Though differentiating the source of inbound tourism is important to identify tourism from which countries are most likely to be influenced by fiction books, the data of this sort are confidential. We thus only report the overall international tourism rather than country-specific one. For ease of the estimation, we took the natural logarithm of the number of foreign tourists to make its distribution approximately normal. Moreover, we also used the annual foreign exchange revenue from tourism of the province from 1993 to 2005 as an alternative measure of tourism development, which was obtained from the China Statistical Yearbook, published by the National Bureau of Statistics. For this measure, we convert the data into relevant values with 1994 as the base year, and then took the natural logarithm of original variables to adjust for potential heteroscedasticity.

3.2.2   Literary destination familiarity

As discussed in the previous review of literary tourism scholarship, literary destination familiarity involves two components: affection-based destination familiarity and knowledge-based destination familiarity. Affection-based familiarity can result from reading novels or other forms of fiction. Because fiction, as a genre, represents a process of reconstruction of self-discovery and cultural recovery, it often captures and conveys the cultural symbol of the destination[17, 66, 67, 70−72] and is characterized by long-lasting cultural authenticity. Contemporary knowledge, by contrast, can be gained from reading news or magazines that deliver instantaneous information about a destination country, region, or society using contemporary language.

Normally, destination familiarity at the societal level is almost impossible to measure using traditional survey methods. We took advantage of the English fiction books dataset of the Google Books corpus to construct affection-based destination familiarity, and used the New York Times corpus and the Time magazine corpus to construct knowledge-based destination familiarity.

3.2.2.1   Affection-based destination familiarity

We constructed the index of affection-based destination familiarity (ADF) with 31 provinces in Mainland China from the English fiction sub-corpora in the Google Books Ngram corpus. With more imaginary elements, fictional stories can strongly reflect and cultivate people's mental and moral development, and take part in shaping the cultural identity of regions, communities, and individuals[63, 66, 67]. Rather than providing knowledge, fictional narratives would increase one's empathy and compassion, which is more affection-based. In this vein, we take advantage of English fiction book corpus to construct affection-based similarity.

By searching the text corpora for any province name, we can learn how often that province is mentioned in a given period. Similar analytical strategy has widely been adopted in recent studies (e.g., Refs. [29, 60−62]).

We assumed that in any given year, the more frequently a province's name appears in fiction books, the more prominent its visibility and level of destination familiarity. We were careful to use multiple spellings of province names to capture every reference. Since 1949, the English-speaking world has relied on the Chinese Hanyu Pinyin system to spell out the names of locations. For example, before the World War II, "Beijing" was

called "Peking" . To calculate the word frequency of Beijing, we counted both. Moreover, terms are capitalized differently, so our search was inclusive of those differences. To obtain the most accurate data, we synthesized all search results. The detailed search terms are presented in the appendix (Table A1).

As a standard practice (e.g., Refs. [60, 73]), we normalized the total appearance of a key word using the frequency of "the" in the same year to correct for the potential bias coming from the influx of data. The normalized annual frequency of the word usage of Chinese provinces fiction-based destination familiarity (FDF) can be computed by

$$FDF_{i\_t} = \frac{C_{i\_t}}{C_{the\_t}} \qquad (1)$$

where C_{i_t} represents the frequency of province name appeared in English fictions. C_{the_t} is the number of appearances of "the".

3.2.2.2 Knowledge-based destination familiarity

To measure the knowledge-based destination familiarity, we constructed the New York Times (NYT)-based destination familiarity index (NDF) and the Times-based destination familiarity index (TDF) using the New York Times corpus, a daily newspaper founded in 1851 with worldwide influence and readership, and Time magazine, a weekly magazine published in New York City, which has the world's largest circulation. From the New York Times, we searched the annual number of articles containing the names of Chinese provinces and divided it by the total number of articles produced per year. The data can be accessed from "Chronicle" developed by the New York Times (the webpage can be accessed at http://chronicle.nytlabs.com/). From Time magazine, we conducted the similar searching and found 100 million words of text in 275000 articles from 1923 to 2006 (it can be accessed at http://corpus.byu.edu/time/). Based on this, we created an index with the total mentions of each province for each year.

3.2.3 Control variables

To reveal the effect of literary destination familiarity, the variables of economic development level, regions' infrastructure, international tourism, labor costs, and institutional environment were properly controlled in the estimation of inbound tourism. Economic Development Levels (GDP), captured by the provincial GDP, are shown to be correlated with tourism development. Again,

we adopted the similar operationalization as tourism development by converting the data into relevant values using 1994 as the base year and taking the natural logarithm of them. Infrastructure (INF), calculated by provincial road and highway density, represents the accessibility of transportation in China that allows travelers and tourism companies to design travel packages. A reasonable extrapolation is that better infrastructure construction stimulates the growth of foreign tourists as they can travel more conveniently, which refers that the infrastructure level can be an endogenous factor. Existing findings show that regional GDP and tourism have positive association[74], which suggests that regional GDP can simultaneously affect/interact with local tourism alongside literature familiarity. Thus, we control regional GDP to eliminate the potential confounding bias. The inflow of Foreign Direct Investment (FDI), measured by "actually utilized" provincial per capita FDI in a given province, is also controlled as trade openness. Finally, institutional attributes, proxied by the number of Non-Governmental Organizations (NGOs) per 100000 residents in a given province, is also controlled because it represents the strength of the civil society and the robustness of institutional building. We took the natural logarithm for all the values of original variables. The descriptive statistics of the aforesaid variables are summarized in Table 1.

3.3 Dynamic panel models

The analysis is conducted based on provincial panel data comprising 31 provinces over the period 1994−2004, which is conducive to take explicit account of province-specific heterogeneity. Given that tourism mobility is largely path-dependent, we employ dynamic panel model by including the lagged-value of number of foreign tourists in the model. The model is written as

$$
\begin{aligned}
Tour_{it} = {} & \gamma \times Tour_{i(t-1)} + \alpha \times FDF_{i(t-1)} + \\
& \gamma_1 \times NDF_{i_{t-1}} + \gamma_2 \times TDF_{i_{t-1}} + \\
& \beta_1 \times GDP_{i(t-1)} + \beta_2 \times INF_{i(t-1)} + \\
& \beta_3 \times NGO_{i(t-1)} + \beta_4 \times FDI_{i(t-1)} + \\
& \theta_t + c_i + \mu_{it} \qquad (2)
\end{aligned}
$$

In Eq. (2), $Tour_{it}$ is the dependent variable, representing the number of foreign tourists or foreign exchange earnings from tourists in year t , while $Tour_{i(t-1)}$ is its corresponding one-year lag value in the same province i in the previous year t−1. $FDF_{i(t-1)}$,

Table 1 Descriptive statistics of variables.

	Variable	Abbr.	Description	Mean	S.D.
Dependent Variable	Inbound Tourism	$Tour\_N_t$	Number of inbound tourists (1994–2004)	93.000	195.305
		$Tour\_R_t$	Annual foreign exchange revenue from tourism (1993–2005)	21470.37	40789.40
Explanatory Variable	Affection-based Destination Familiarity	$FDF_{(t-1)}$	Appearance in English fiction (1994–2003)	0.720	1.714
	Knowledge-based Destination Familiarity	$NDF_{(t-1)}$	Appearance in *New York Times* (1994–2003)	0.704	1.739
		$TDF_{(t-1)}$	Appearance in *TIME* magazine (1994–2003)	2.496	9.659
Control Variables	Lagged Inbound Tourism	$Tour\_N_{(t-1)}$	One-year lag of number of inbound tourists (1994–2003)	86.880	182.240
	Economic Development	GDP_{t-1}	GDP (1994–2003)	1489.449	1253.734
	Infrastructure	INF_{t-1}	Road/Highway density (1994–2003)	30.5305	19.809
	NGO Development	NGO_t	Number of NGO per capita (1994–2004)	1.312	0.504
	Openness	$FDI_{(t-1)}$	Lagged FDI (1994–2003)	69.546	111.900

Note: For readability, FDF variables are multiplied by 100000; NDF variables are multiplied by 1000.

$NDF_{i(t-1)}$, and $TDF_{i(t-1)}$ are the core explanatory variables. $FDF_{i(t-1)}$ is the index of affection-based destination familiarity with 31 provinces extracted from Google English Fiction Corpus. $NDF_{i(t-1)}$ and $TDF_{i(t-1)}$ are the indices of knowledge-based destination familiarity constructed using New York Times corpus and Time magazine corpus, respectively. In addition, c_i is the unobserved regional effect, and μ_{it} is the idiosyncratic shock, which is assumed to be serially uncorrelated over time. θ_t denotes different period intercepts by controlling year dummies; in doing so, we assumed that there is no correlation of the idiosyncratic disturbances across provinces[75].

The big challenge of dynamic panel model is the possible correlation between the lagged dependent variable $Tour_{it}$ and the time invariant error terms c_i. To deal with this issue, we adopt GMM estimation, specifically, the difference GMM (DGMM) and the system-GMM (SGMM) estimators, which were developed by Arellano and Bond (1991)[30] and Holtz-Eakin et al. (1990)[31]. DGMM fits the first-difference specifications of the model to subtract out c_i and take appropriate lagged dependent and independent variables as "internal" instrumental variables to solve the potential correlation between lagged error terms μ_{it} and $Tour_{it}$ in the estimation. However, as Arellano and Bover (1995)[76], and Griliches and Hausman (1986)[77], DGMM estimators might suffer from a weak instruments problem when the autoregressive parameter approaches unity[76], or when the variance of the unobserved unit-specific effects is large[77], SGMM was

further adopted. SGMM not only uses first differences of dependent variable to construct the orthogonality condition of the error term, it also introduces suitable lags of the lagged independent variables in levels, which can be taken as potential instruments. Compared with DGMM, SGMM augments DGMM by estimating simultaneously in differences and levels, when overidentification test has passed, the aforementioned bias can be largely attenuated in SGMM[76, 78]. The additional condition of SGMM do not come without a cost, as such operationalization may lead to proliferation of instruments, bringing about overfitting of endogenous variables due to a finite sample bias[79]. To choose a more suitable estimation, we also conducted a series of standard tests for GMM estimations: (1) the unit-root test to test stationary time series; (2) an AR(2) test to check the serial correlation of μ_{it}; and (3) the Hansen/Sargan over-identification test to examine the instruments validity.

4 Result

We present the GMM estimation in Table 2 , where Models 1−5 are the results of DGMM, and Model 6 is the result of SGMM. We first employed the DGMM method by using the lagged number of foreign tourists as the instruments for the first-differenced number of foreign tourists.

First, we fit pooled Ordinary Least Squares (OLS) regression and FE models and found that the coefficients of lagged foreign tourist number are 1.004 and 0.456,

Table 2   Dynamic panel regressions of predicting the effect of literary destination familiarity on inbound tourism (1994–2004).

Variable	DGMM					SGMM
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$Tour\_N_{(t-1)}$	0.487***	0.530***	0.518***	0.457***	0.444***	0.642***
	(0.061)	(0.058)	(0.061)	(0.067)	(0.069)	(0.058)
$FDF_{(t-1)}$	0.022***	0.023***	0.034***	0.031***	0.029***	0.048***
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.010)
$NDF_{(t-1)}$	0.001	0.001	−0.002	−0.001	−0.002	0.070***
	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)	(0.018)
$TDF_{(t-1)}$	−0.003***	−0.003***	−0.003***	−0.003***	−0.003***	−0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
$GDP_{(t-1)}$		−0.018	0.164	0.122	0.140	0.255***
		(0.086)	(0.148)	(0.162)	(0.159)	(0.050)
$FDI_{(t-1)}$			0.043***	0.049***	0.054***	0.070***
			(0.009)	(0.010)	(0.014)	(0.021)
$INF_{(t-1)}$				−0.232***	−0.233***	−0.205***
				(0.065)	(0.069)	(0.062)
$NGO_{(t-1)}$					−0.010	−0.029
					(0.074)	(0.081)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	1.791***	1.766***	0.457	1.573	1.504	−0.061
	(0.214)	(0.656)	(1.114)	(1.259)	(1.200)	(0.270)
AR(2) *p*-value	0.376	0.361	0.680	0.952	0.870	0.407
Hansen test *p*-value	0.213	0.265	0.238	0.398	0.428	1.000
N	246	246	241	241	241	272

Note: (1) The numbers in the parentheses are adjusted robust standard errors; (2) $p < 0.05$*, $p < 0.01$**, $p < 0.001$*** (two-tailed tests); (3) AR(2) is the Arellano-Bond test (H0: there is no autocorrelation), and the Hansen test is the overidentification test (H0: there are no invalid overidentifying restrictions).

respectively. As our estimates from DGMM and SGMM lie between these two values, this implies that both DGMM and SGMM estimates are adaptable. From Table 2, we can see that the coefficient of the lagged number of tourists ($Tour_{i(t-1)}$) is 0.487, falling exactly between 1.004 and 0.456 as expected. Moreover, the DGMM model yielded a statistically significant coefficient of 0.022 for the estimated effect of FDF on number of foreign tourists, and a statistically significant but small effect of TDF ($\gamma\_1=-0.003$), and an insignificant effect of NDF on foreign tourists. That being said, affection-based destination familiarity is important in predicting the in-bound tourism in the subsequent year, however, the effect of knowledge-based destination familiarity is trivial. This is likely because, compared with fiction, news reports current events in a more objective manner and on a wider scale. For periodicals, such as Time magazine, when involving Chinese cities, it often addresses political, ethical, and

social questions or involves debates on a range of issues, which may do little to promote in-bound tourism and/or could even reduce it. As we further controlled for GDP per capita, FDI, infrastructure, and institutional environment, we see that NGO development has shown no significant effect on the number of foreign tourists. But provincial GDP per capita and FDI in the previous year are positively associated with tourism development in that province the current year; by contrast, the infrastructure in the previous year is negatively associate with tourism development in the current year. This result is unexpected, our hunch is rather than road/highway density, which is important for domestic tourism, the development of other industry, for example, airline industry may be associated with international tourism more. Moreover, the positive effect of FDF is consistently shown ($p < 0.001$), holding constant of all socioeconomic characteristics of provinces in the previous year.

Considering the fact that the lagged levels may be weak instruments of first-differenced variables, we proceeded to the SGMM estimator to utilize lagged first-differences as instruments in levels equation. To compare DGMM and SGMM, we further conducted a series of standard tests for GMM estimation. Both DGMM and SGMM have passed an AR(2) test, meaning that serial correlation poses a less serious threat to our estimations. The results from the Hansen test shows that the p-value of DGMM estimates is larger than 0.1, meaning that the internal instruments are valid. However, in SGMM, the value is 1, this implies that there are too many invalid instruments, which may result in model validity issues. The DGMM estimates are thus preferred, but the result of SGMM is still presented as a comparison. In addition, we also use foreign exchange earnings from tourism as the alternative measure of tourism development, the results remain consistent (see Table 3).

Note that fiction books, compared with news and magazines, often reflect the long-term accumulation of knowledge, experiences, and attitudes. Their slow circulation makes them less likely to have an immediate effect on tourism. However, our results show that after controlling for the destination familiarity constructed from contemporary media (i.e., news and magazines), destination familiarity constructed from fiction—even with low circulation and low exposure based on our operationalization compared with contemporary media—still exerts a significantly positive impact on in-bound tourism, offering strong evidence to support our theory.

We then conducted static models as robust checks for our main findings. Taking the Hausman-test for both dependent variables, the fixed-effect model is preferred compared to the random-effect model in our case. Models with both independent variables show positive and significant effect on outcomes, which suggests that

Table 3   Dynamic panel regressions of predicting the effect of literary destination familiarity on foreign exchange earnings from tourism (1994–2004).

Variable	DGMM					SGMM
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
$Tour\_N_{(t-1)}$	0.203***	0.215***	0.200***	0.112	0.061	0.704***
	(0.029)	(0.030)	(0.055)	(0.069)	(0.083)	(0.056)
$FDF_{(t-1)}$	0.035***	0.033***	0.035***	0.038***	**0.032*****	**0.065*****
	(0.006)	(0.006)	(0.007)	(0.006)	**(0.005)**	**(0.010)**
$NDF_{(t-1)}$	0.008	0.025*	0.009	−0.016	0.009	0.077**
	(0.011)	(0.015)	(0.011)	(0.026)	(0.016)	(0.037)
$TDF_{(t-1)}$	−0.005***	−0.006***	−0.004***	−0.004***	−0.004***	−0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
$GDP_{(t-1)}$		0.069	0.129	−0.112	0.126	0.216**
		(0.088)	(0.144)	(0.221)	(0.192)	(0.095)
$FDI_{(t-1)}$			0.037***	0.054***	0.041***	0.017
			(0.013)	(0.015)	(0.014)	(0.018)
$INF_{(t-1)}$				−0.424***	−0.352***	0.062
				(0.115)	(0.101)	(0.140)
$NGO_{(t-1)}$					−0.107	−0.203**
					(0.073)	(0.097)
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	6.828***	6.054***	6.422***	9.767***	7.911***	0.779*
	(0.273)	(0.661)	(1.366)	(2.300)	(1.526)	(0.418)
AR(2) p-value	0.873	0.991	0.633	0.547	0.400	0.503
Hansen test p-value	0.241	0.317	0.165	0.418	0.519	1.000
N	275	275	268	268	268	299

Note: (1) The numbers in the parentheses are adjusted robust standard errors; (2) $p < 0.05*$, $p < 0.01**$, $p < 0.001***$ (two-tailed tests); (3) AR(2) is the Arellano-Bond test (H0: there is no autocorrelation), and the Hansen test is the overidentification test (H0: there are no invalid overidentifying restrictions).

literature familiarity has positive impact on choices of tourists (see the appendix (Table A2)). That further indicates the validity of our major findings.

5 Discussion and Conclusion

In the past, cultural distance was a primary factor affecting cross-border trade behaviors. However, cultural distance largely ignores the cognitive dimension of culture. This is problematic because, in many cases, it is not the "real" or objective cultural differences between home and destination countries, but how visitors perceive the cultural value of the destination that may influence tourists' choice of potential locations. We propose a literary dimension of destination familiarity as a concept correlated with, but essentially different from cultural distance, which is a predictor of development of inbound tourism.

Considering the multidimensionality of destination, we decompose literary destination familiarity into two types, namely, knowledge-based destination familiarity and affection-based destination familiarity. To construct indices of these two types of destination familiarity, we extract the appearance frequency of Chinese province names from 1994 to 2004 using the English fiction dataset in the Google Books corpus, the New York Times annotated corpus and the Time magazine corpus. We discovered that affection-based destination familiarity plays an essential role in attracting foreign tourists. The higher the level of affection-based destination familiarity, the larger the inflow of foreign tourists, and the higher the foreign exchange earnings from tourism.

Note that our results show a distinct impact from affection-based destination familiarity (by appearing in literature) and knowledge-based destination familiarity (especially for the Time based familiarity index) on local tourism. While the former factor keeps a positive and significant effect on local tourism, the later one has an opposite impact. As a premature hypothesis, we propose that this distinction may due to the difference of sentiments/attitudes in literature and the Time magazine when they mention Chinese regions. It is possible that reports in Time may contain more negative attitudes and stereotypes for China, which hinders tourists from travelling to China. But this paper does not have enough necessary evidence to support this extrapolation. We suggest that a valuable option for future researchers is to dig deeper to subtle mechanisms by using the text and sentiment analysis.

Still, our findings have some limitations that deserve further examination. First, our construction of destination familiarity is primarily based on the premise that the appearance of the names of Chinese provinces in three corpora (i.e., Google English fiction corpus, New York Times corpus, and Time magazine corpus) which makes them real and familiar to readers and that the effect of destination familiarity is the same across all home countries. In this regard, we equalized the province exposure to different home countries/regions across the world to a large degree. Second, when it comes to destination familiarity with the potential destination, two measures are of central importance: book-specific measure and sentiment measure. The former, through linking the keywords to specific books, can identify the nationalities of the authors and publishers that are crucial to reveal the relevance of these cultural products. The latter can reflect how the keywords are described in the books and how likely the keywords are to resonate significantly. However, owing to the copyright problem, the Google Books corpus only provides the aggregate searching information each year rather than specific searching information from each book each year; thus, we cannot link the word searching to specific books and extract the relevant sentiment from the content. It remains to be studied why and how fiction-based destination familiarity would affect in-bound tourism—is it because of the narrative style, the positive sentiment associated with the locale, or some other factor? And do these effects relate only to fiction books, or could they also be found for the movies based on fictions? These are important questions that should be investigated in future research. Until then, this study provides a first step to build the linkage between literary destination familiarity and in-bound tourism. To enrich the theory relevant to literary tourism, more careful research is needed. Third, we ignored the varying effect of destination familiarity at the different stages of tourism development. In particular, the effect of culture is likely to dominate in the early stage of tourism development, when foreign tourists have little prior experiences and knowledge about Mainland China. As tourists gain more actual experience in an area over time, the effect of destination

familiarity is likely to decline. Fourth, due to the limitation of our data, we were not able to differentiate the specific characteristic differences of first-time tourists vs. experienced tourists. Previous studies show that experienced tourists who repeatedly visited a given destination had a relatively higher loyalty to that place than first-timers[80, 81]. As such, how destination familiarity influences destination choices among different groups still needs further research. Fifth, although we recognize that many foreign tourists coming to Mainland China are not English speakers, we can not separate their nationalities or add literature that written in other languages due to the data limitation. That may raise potential biases for our findings. However, we argue that as English is the most prevalent language in the world, even people from other countries (e.g., Japan & South Korea) can be influenced by English literature. Still, we recommend readers to keep mind in that potential bias when interpreting our findings. We highlight that further research can contribute to this issue by getting finer measurements for the proportion of tourists from different countries and adding sources from other languages such as Japanese, South Korean, French, etc.

Despite these drawbacks, our research is an innovative attempt to reveal how cultural mechanisms might influence tourism activities in a given destination. Theoretically, the proposition of literary destination familiarity complements and further enriches the research on literary tourism. Methodologically, we take advantage of a corpus method to construct an index of literary destination familiarity, a phenomenon that would be difficult to quantify using a traditional survey. As a recent study asserts, this novel method "has content validity and can be used as a proxy measure for previously difficult-to-research phenomena and questions"[82]. We believe that with the optimization of high-throughput digitized resources, we can utilize word-search dynamics to study sociocultural trends and individual practices with rich spatial and temporal scales.

Yet choices of travel site in a country/region that is geographically and culturally distant from the home country involves a rational and intuitive decision-making process, which might vary in different contexts and at different stages of tourism development. This result, at least to some extent, suggests the possibility of promoting fiction-related literary tourism to Mainland China. As Mainland China becomes more and more globalized, there is increasing exposure of its cities in various provinces. If fiction can have an enduring impact on international tourism, one could develop new fictional cultural products that use previously successful narratives and similar characters and images in order to promote Mainland China. With the help of social media marketing, the diffusion of such fiction-typed cultural products could spur the development of tourism in Mainland China. Hopefully, this study will inspire more investigation of such questions in the future.

Appendix

Table A1 Search terms for Chinese provincial-level regions.

Chinese name	Searching terms
北京	Beijing, Peking
天津	Tianjin, Tientsin
河北	Hebei, Hopei
山西	Shanxi
内蒙古	Inner Mongolia, Neimenggu
辽宁	Liaoning
吉林	Jilin, Kirin
黑龙江	Heilongjiang
上海	Shanghai
江苏	Jiangsu, Kiangsu
浙江	Zhejiang, Chekiang
安徽	Anhui
福建	Fujian, Fukien, Hokkien
江西	Jiangxi, Kiangsi
山东	Shandong, Shangtung
河南	Henan, Honan
湖北	Hubei, Hupei
湖南	Hunan
广东	Guangdong, Kwangtung, Canton
广西	Guangxi, Kwangsi
海南	Hainan
重庆	Chongqing, Chungking
四川	Sichuan, Szechwan
贵州	Guizhou, Kweichow
云南	Yunan
西藏	Tibet, Xizang
陕西	Shaanxi
甘肃	Gansu, Kansu
青海	Qinghai
宁夏	Ningxia, Ningsia
新疆	Xinjiang, Sinkiang

Table A2   Result by static panel model (fixed-effect for both year and province).

Variable	Number of Foreign Tourists	Foreign Exchange Earnings
$FDF_{(t-1)}$	**0.054*****	**0.093*****
	(0.015)	**(0.028)**
$NDF_{(t-1)}$	**−0.000**	**0.007**
	(0.026)	**(0.032)**
$TDF_{(t-1)}$	**−0.003****	**−0.005****
	(0.001)	**(0.002)**
$GDP_{(t-1)}$	−0.102	0.085
	(0.266)	(0.345)
$FDI_{(t-1)}$	0.032	0.077**
	(0.028)	(0.030)
$INF_{(t-1)}$	0.034	−0.027
	(0.151)	(0.216)
$NGO_{(t-1)}$	−0.146	−0.164*
	(0.106)	(0.095)
Year Fixed	Yes	Yes
Province Fixed	Yes	Yes
Constant	3.754*	7.668***
	(1.898)	(2.412)
Hausman Test chi2	34.66	28.98
p-value	0.0001	0.0007
N	308	308

Acknowledgment

References

[1]   G. Hofstede, *Culture's Consequences: International Differences in Work-Related Values*. Beverly Hills, CA, USA: Sage, 1980.

[2]   J. C. Crotts, The effect of cultural distance on overseas travel behaviors, *J. Travel Res.*, vol. 43, no. 1, pp. 83–88, 2004.

[3]   J. C. Crotts and A. Pizam, The effect of national culture on consumers' evaluation of travel services, *J. Tour. Cult. Commun.*, vol. 4, no. 1, pp. 17–28, 2003.

[4]   E. Kastenholz, 'Cultural proximity' as a determinant of destination image, *J. Vacat. Market.*, vol. 16, no. 4, pp. 313–322, 2010.

[5]   H. B. Liu, X. R. Li, D. A. Cárdenas, and Y. Yang, Perceived cultural distance and international destination choice: The role of destination familiarity, *J. Destinat. Market. Manag.*, vol. 9, pp. 300–309, 2018.

[6]   A. Y. Liu, D. X. F. Fan, and R. T. Qiu, Does culture affect tourism demand? A global perspective *J. Hospital. Tour. Res.*, vol. 45, no. 1, pp. 192–214, 2021.

[7]   M. M. Qiang, M. H. Shen, and H. M. Xie, Cultural diffusion and international inbound tourism: Evidence from China, *Tour. Econom.*, vol. 25, no. 6, pp. 884–903, 2019.

[8]   Y. Reisinger and J. C. Crotts, Applying Hofstede's national culture measures in tourism research: Illuminating issues of divergence and convergence, *J. Travel Res.*, vol. 49, no. 2, pp. 153–164, 2010.

[9]   Y. Yang and K. K. F. Wong, The influence of cultural distance on China inbound tourism flows: A panel data gravity model approach, *Asian Geogr.*, vol. 29, no. 1, pp. 21–37, 2012.

[10]   D. T. Herbert, Heritage as literary place, in *Heritage, Tourism and Society*, D. T. Herbert, ed. London, UK: Mansell Publishing, 1995, pp. 32-48.

[11]   L. Stiebel, Hitting the hot spots: Literary tourism as a research field with particular reference to KwaZulu-Natal, South Africa, *Crit. Arts*, vol. 18, no. 2, pp. 31–44, 2004.

[12]   L. Buell, D. Lodge, and N. Wood, *Place Modern Criticism and Theory*. London, UK: Pearson Longman, 1989.

[13]   B. L. Kirkman, K. B. Lowe, and C. B. Gibson, A quarter century of culture's consequences: A review of empirical research incorporating Hofstede's cultural values framework, *J. Int. Busin. Stud.*, vol. 36, pp. 285–320, 2006.

[14]   F. M. S. Li, Culture as a major determinant in tourism development of China, *Curr. Issues Tour.*, vol. 11, no. 6, pp. 492–513, 2008.

[15]   K. Sivakumar and C. Nakata, The stampede toward Hofstede's framework: Avoiding the sample design pit in cross-cultural Research, *J. Int. Bus. Stud.*, vol. 32, no. 3, pp. 555–574, 2001.

[16]   H. G. Xu, P. Y. Ding, and J. Packer, Tourism research in China: Understanding the unique cultural contexts and complexities, *Curr. Issues Tour.*, vol. 11, no. 6, pp. 473–491, 2008.

[17]   S. J. Squire, The cultural values of literary tourism, *Ann. Tour. Res.*, vol. 21, no. 1, pp. 103–120, 1994.

[18]   L. Ghetau and L. V. Esanu, Literary tourism as a promoter of cultural heritage, *WIT Trans. Ecol. Environ.*, vol. 150, pp. 345–353, 2011.

[19]   C. Ryan, Y. N. Zhang, H. M. Gu, and S. Ling, Tourism, a classic novel, and television: The case of Cao Xueqin's dream of the red mansions and grand view gardens, Beijing, *J. Travel Res.*, vol. 48, no. 1, pp. 14–28, 2009.

[20]   X. J. Yu and H. G. Xu, Ancient poetry in contemporary Chinese tourism, *Tour. Manag.*, vol. 54, pp. 393–403, 2016.

[21]   X. J. Yu and H. G. Xu, Moral gaze at literary places: Experiencing 'being the first to worry and the last to enjoy' at Yueyang Tower in China, *Tour. Manag.*, vol. 65, pp. 292–302, 2018.

[22]   C. C. Chen and Y. H. Lin, Segmenting mainland Chinese tourists to Taiwan by destination familiarity: A factor-cluster approach, *Int. J. Tour. Res.*, vol. 14, no. 4, pp. 339–352, 2012.

[23]   T. K. Srull, The role of prior knowledge in the acquisition, retention, and use of new information, in *Advances in*

Consumer Research, R. Bagozzi and A. Tybout, eds. Ann Arbor, MI, USA: Association for Consumer Research, 1983, pp. 572–576.

[24] T. Guardian, China will be the world's most visited country by 2030, https://www.theguardian.com/travel/2018/nov/06/china-will-be-the-worlds-most-popular-holiday-destination-by-2030, 2018

[25] D. MacCannell, *Empty Meeting Grounds: The Tourist Papers*. London, UK: Routledge, 1992.

[26] J. Urry, *The Tourist Gaze: Leisure and Travel in Contemporary Societies*. London, UK: Sage Publications Ltd, 1990.

[27] D. Crystal, *English as a Global Language*. 2nd ed. New York, NY, USA: Cambridge University Press, 2003.

[28] C. M. Hall and S. J. Page, *Tourism in South and Southeast Asia: Issues and Cases*. London, UK: Butterworth-Heinemann, 2000.

[29] Y. S. Chen, F. Yan, and Y. Zhang, Local name, global fame: The international visibility of Chinese cities in modern times, *Urban Stud.*, vol. 54, no. 11, pp. 2652–2668, 2017.

[30] M. Arellano and S. Bond, Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Rev. Econom. Stud.*, vol. 58, no. 2, pp. 277–297, 1991.

[31] D. Holtz-Eakin, W. Newey, and H. S. Rosen, Estimating vector autoregressions with panel data, *Econometrica*, vol. 56, no. 6, pp. 1371–1395, 1990.

[32] M. J. Ahn and B. McKercher, The effect of cultural distance on tourism: A study of international visitors to Hong Kong, *Asia Pacif. J. Tour. Res.*, vol. 20, no. 1, pp. 94–113, 2015.

[33] G. Hofstede, The cultural relativity of organizational practices and theories, *J. Int. Busin. Stud.*, vol. 14, no. 2, pp. 75–89, 1983.

[34] Y. Reisinger, *International Tourism: Cultures and Behavior*. New York, NY, USA: Routledge, 2012.

[35] J. Fourie and M. Santana-Gallego, Ethnic reunion and cultural affinity, *Tour. Manag.*, vol. 36, pp. 411–420, 2013.

[36] M. M. Li, Cross-cultural tourist research: A meta-analysis, *J. Hospital. Tour. Res.*, vol. 38, no. 1, pp. 40–77, 2014.

[37] J. L. Crompton, Motivations for pleasure vacation, *Ann. Tour. Res.*, vol. 6, no. 4, pp. 408–424, 1979.

[38] G. R. Larsen and J. W. Guiver, Understanding tourists' perceptions of distance: A key to reducing the environmental impacts of tourism mobility, *J. Sustain. Tour.*, vol. 21, no. 7, pp. 968–981, 2013.

[39] B. McKercher and H. D. Cros, Testing a cultural tourism typology, *Int. J. Tour. Res.*, vol. 5, no. 1, pp. 45–58, 2003.

[40] S. O'Leary and J. Deegan, People, pa ce, place: Qualitative and quantitative images of Ireland as a tourism destination in France, *J. Vacat. Market.*, vol. 9, no. 3, pp. 213–226, 2003.

[41] L. L. Wu, J. Y. Zhang, and A. Fujiwara, A tourist's multi-destination choice model with future dependency, *Asia Pacif. J. Tour. Res.*, vol. 17, no. 2, pp. 121–132, 2012.

[42] H. Lefebvre and D. Nicholson-Smith, *The Production of Space*. Hoboken, NJ, USA: Wiley-Blackwell, 1992.

[43] A. N. Seaman, Exploring the connections between literary places, literary texts, and tourist performance, PhD

Dissertation, Clemson Univ., Clemson, SC, USA, 2016.

[44] Y. F. Tuan, *Topophilia: A Study of Environmental Perception, Attitudes, and Values*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1974.

[45] L. A. Samovar and R. E. Porter, *Intercultural Communication: A reader*. Belmont, CA, USA: Wadsworth Publishing Co, 1994.

[46] R. Williams, *Keywords: A Vocabulary of Culture and Society*. New York, NY, USA: Oxford University Press, 1976.

[47] G. Busby and J. Klug, Movie-induced tourism: The challenge of measurement and other issues, *J. Vacat. Market.*, vol. 7, no. 4, pp. 316–332, 2001.

[48] J. P. Robertson and L. A. Radford, The private uses of quiet grandeur: A meditation on literary pilgrimage, *Chang. Engl.*, vol. 16, no. 2, pp. 203–209, 2009.

[49] D. T. Herbert, Literary places, tourism and the heritage experience, *Ann. Tour. Res.*, vol. 28, no. 2, pp. 312–333, 2001.

[50] N. Watson, *The Literary Tourist*. New York, NY, USA: Palgrave Macmillan, 2006.

[51] P. A. Westover, Inventing the London of literary tourist: Walking the romantic city in Leigh Hunt's 'Wishing Cap' essays, *Eur. Romant. Rev.*, vol. 23, no. 1, pp. 1–19, 2012.

[52] H. C. Anderson and M. Robinson, *Literature and Tourism: Reading and Writing Tourism Texts*. London, UK: Continuum, 2002.

[53] R. Prentice, Tourist familiarity and imagery, *Ann. Tour. Res.*, vol. 31, no. 4, pp. 923–945, 2004.

[54] J. L. Crompton, Structure of vacation destination choice sets, *Ann. Tour. Res.*, vol. 19, no. 3, pp. 420–434, 1992.

[55] A. D. A. Tasci and W. C. Gartner, Destination image and its functional relationships, *J. Travel Res.*, vol. 45, no. 4, pp. 413–425, 2007.

[56] E. Cohen and R. L. Cooper, Language and tourism, *Ann. Tour. Res.*, vol. 13, no. 4, pp. 533–563, 1986.

[57] S. Seo, O. Y. Kim, S. Oh, and N. R. A. Yun, Influence of informational and experiential familiarity on image of local foods, *Int. J. Hospital. Manag.*, vol. 34, pp. 295–308, 2013.

[58] H. I. Lowe, *Mark Twain's Homes and Literary Tourism*. Columbia, MO, USA: University of Missouri Press, 2012.

[59] S. Baloglu, Image variations of Turkey by familiarity index: Informational and experiential dimensions, *Tour. Manag.*, vol. 22, no. 2, pp. 127–133, 2001.

[60] R. A. Bentley, A. Acerbi, P. Ormerod, and V. Lampos., Correction: Books average previous decade of economic misery, *PLoS One*, vol. 9, no. 1, p. e83147, 2014.

[61] Y. S. Chen and F. Yan, International visibility as determinants of foreign direct investment: An empirical study of Chinese provinces, *Soc. Sci. Res.*, vol. 76, pp. 23–39, 2018.

[62] Y. S. Chen and F. Yan, Economic performance and public concerns about social class in twentieth-century books, *Soc. Sci. Res.*, vol. 59, pp. 37–51, 2016.

[63] J. B. Michel, Y. K. Shen, A. P. Aiden, A. Veres, M. K. Gray, The Google Books Team, J. P. Pickett, D. Hoiberg, D. Clancy, P. Norvig, et al., Quantitative analysis of culture using millions of digitized books, *Science*, vol. 331, no. 6014, pp. 176–182, 2011.

[64] E. J. Marsh, M. L. Meade, and H. L. Roediger III,

Learning facts from fiction, *J. Memory Lang.* , vol. 49, no. 4, pp. 519–536, 2003.

[65] J. R. Shortridge, The concept of the place-defining novel in American popular culture, *Profess. Geogr.* , vol. 43, no. 3, pp. 280–291, 1991.

[66] P. M. Bal and M. Veltkamp, How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation *PLoS One* , vol. 8, no. 1, p. e55341, 2013.

[67] D. C. Kidd and E. Castano, Reading literary fiction improves theory of mind, *Science* , vol. 342, no. 6156, pp. 377–380, 2013.

[68] K. L. Macpherson, The head of the dragon: The Pudong new area and Shanghai's urban development, *Plann. Perspect.*, vol. 9, no. 1, pp. 61–85, 1994.

[69] M. K. Chalmers and P. N. Edwards, Producing 'one vast index': Google Book Search as an algorithmic system, *Big Data Soc.*, vol. 4, no. 2, pp. 1–6, 2017.

[70] B. Earl, Literary tourism: Constructions of value, celebrity and distinction, *Int. J. Cult. Stud.* , vol. 11, no. 4, pp. 401–417, 2008.

[71] A. Hoppen, L. Brown, and A. Fyall, Literary tourism: Opportunities and challenges for the marketing and branding of destinations? *J. Destinat. Market. Manag*, vol. 3, no. 1, pp. 37–47, 2014.

[72] C. Mansfield, *Researching Literary Tourism*. Plymouth, UK: Shadows Books & Media, 2015.

[73] A. Acerbi, V. Lampos, P. Garnett, and R. A. Bentley, The expression of emotions in 20th century books, *PLoS One*, vol. 8, no. 3, p. e59030, 2013.

[74] M. Castro-Nuño, J. A. Molina-Toucedo, and M. P. Pablo-Romero, Tourism and GDP: A meta-analysis of panel data studies, *J. Travel Res.*, vol. 52, no. 6, pp. 745–758, 2013.

[75] D. Roodman, How to do Xtabond2: An introduction to difference and system GMM in Stata, *Stata J.* , vol. 9, pp. 86–136, 2009.

[76] M. Arellano and O. Bover, Another look at the instrumental variable estimation of error-components models, *J. Econometr.*, vol. 68, no. 1, pp. 29–51, 1995.

[77] Z. Griliches and J. A. Hausman, Errors in variables in panel data, *J. Econometr.* , vol. 31, no. 1, pp. 93–118, 1986.

[78] R. Blundell and S. Bond, Initial conditions and moment restrictions in dynamic panel data models, *J. Econometr.*, vol. 87, no. 1, pp. 115–143, 1998.

[79] L. P. Hansen, Large sample properties of generalized method of moments estimators, *Econometrica* , vol. 50, no. 4, pp. 1029–1054, 1982.

[80] G. Lee and I. P. Tussyadiah, Exploring familiarity and destination choice in international tourism, *Asia Pacif. J. Tour. Res.*, vol. 17, no. 2, pp. 133–145, 2012.

[81] X. X. Sun, C. G. Q. Chi, and H. G. Xu, Developing destination loyalty: The case of Hainan Island, *Ann. Tour. Res.*, vol. 43, pp. 547–577, 2013.

[82] S. Richey and J. B. Taylor, Google Books Ngrams and political science: Two validity tests for a novel data source, *Polit. Sci. Polit.*, vol. 53, no. 1, pp. 72–77, 2020.

Guodong Ju is a PhD candidate in the Department of Social Policy, London School of Economics and Political Science. His research interests focus on quantitative methods, causal inference, social network, and minority rights.

Jiankun Liu is an associate professor at the Department of Sociology, Harbin Engineering University. He earned the PhD degree in sociology from Tsinghua University. His main research interest lies in organizational sociology and big data in social science. He has published in *The Journal of Chinese Sociology* and *Chinese Journal of Sociology*.

Guangye He is an associate professor at the Department of Sociology, Nanjing University. Her research focuses on family sociology, social stratification, and quantitative methodology in sociology. She has published in *Social Science Research*, *Poetics*, *Chinese Sociological Review*, *China Review*, and *Journal of Contemporary China*.

Xinyi Zhang is an associate professor at Business School, Beijing Language and Culture University. She received the PhD degree in economics from the Chinese Academy of Fiscal Sciences. Her research interests focus on public finance, tourism research, and quantitative methods. She has published broadly in internationally and nationally recognized academic journals.

Fei Yan is an associate professor in sociology at Tsinghua University. His research focuses on political sociology, historical sociology, and sociology of development. His work has appeared in *Social Science Research*, *Urban Studies*, *The Sociological Review*, *Poetics*, *Social Movement Studies*, and *Oxford Bibliographies in Sociology*.