**Appendix A. Variables definition and measurement**

Exhibit A. Variables definition and measurement

|  |  |
| --- | --- |
| Variable | Definition and measurement |
| Core variables |  |
| Energy poverty | See manuscript |
| Tourism Income | See manuscript |
| Individual and household characteristics |  |
| Age | Age of head of household |
| Marriage | Marital status of the head of household, 1 for married, 0 otherwise |
| Education | Years of formal education of the head of household |
| Health | The health status of the head of household, ranging from 1 to 5, indicating a progressive deterioration in health status |
| Household assets | The size of the household's assets, expressed in logarithms |
| Household size | Number of persons living together |
| Urban | Living in rural or urban areas, 0 for rural, 1 for urban |
| Urban characteristics |  |
| Gross domestic product | Gross domestic product (GDP) of the city, expressed in logarithms |
| Population | Total population of the city, expressed in logarithms |
| Industrial structure | Share of secondary GDP in total GDP |

**Appendix B. Descriptive statistics**

Exhibit B. Descriptive statistics

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Mean | Std. Dev. | Min | Max |
| Core variables |  |  |  |  |
| Energy poverty | 0.179 | 0.384 | 0 | 1 |
| Tourism Income | 5.674 | 1.272 | 2.515 | 8.736 |
| Individual and household characteristics |  |  |  |  |
| Age | 47.551 | 10.758 | 22 | 65 |
| Marriage | 0.888 | 0.316 | 0 | 1 |
| Education | 8.059 | 4.416 | 0 | 22 |
| Health | 3.08 | 1.209 | 1 | 5 |
| Household assets | 57.373 | 101.392 | -8.4 | 652.213 |
| Household size | 3.907 | 1.806 | 1 | 21 |
| Urban | 0.486 | 0.5 | 0 | 1 |
| Urban characteristics |  |  |  |  |
| Gross domestic product | 7.567 | 1.195 | 4.81 | 10.549 |
| Population | 6.138 | 0.737 | 3.914 | 8.047 |
| Industrial structure | 0.424 | 0.109 | 0.095 | 0.794 |

**Appendix C. Endogeneity problem**

The problem of model endogeneity comes from two main sources: reverse causality and omitted variables. Since tourism development is a city-level variable, it is difficult to argue that energy poverty at the household level affects tourism development at the city level, so the likelihood of reverse causality in the model of this paper is low. However, there may be a problem of endogeneity of unobservable omitted variables. In this paper, we use instrumental variables approach for 2SLS estimation. Hence, we use two instrumental variables to try to mitigate the endogeneity problem. The first instrumental variable (IV1) is the number of national scenic spots, key cultural relic protection units (Chinese scenic spots) and intangible cultural heritage within 100 km of the household's city. This instrumental variable is reasonable because, on the one hand, Chinese scenic spots are selected by the state based on local natural vegetation, geomorphological features and ancient architecture, and intangible cultural heritage is inherited by history and culture, both of which mostly depend on natural and historical factors. This implies that IV1 is not directly related to the household's current energy poverty status, satisfying the exclusivity of the instrumental variable. On the other hand, Chinese scenic spots and intangible cultural heritages can increase the degree of local popularity, which is an important factor in attracting tourists, and they have a direct relationship with tourism development. For example, research by Prada-Trigo (2016) shows that intangible cultural heritage is attracting more and more tourists. Intangible heritage contains both culinary and cultural elements that attract tourists. This satisfies the correlation of instrumental variables. However, during the years of our sample, the annual change in the number of famous sites and intangible cultural heritages in China is almost negligible. Because the last two directory changes were in 2013 and 2019. If unchanged, that can be recognized as a variable that does not change over time. This will affect the estimation of the fixed effects model. Therefore, this subsection is first estimated by the 2SLS method with mixed cross sections.

Columns (1) and (2) of Exhibit C present the results of the 2SLS estimation for the first and second stages, respectively. The results in Column (1) show that there is a significant correlation relationship between IV1 and tourism development, which proves that the instrumental variables fulfil the correlation condition. Meanwhile, the Cragg-Donald Wald F statistic in Column (1) is 717.787, which is greater than 16.38, ruling out the possibility of weak instrumental variables (Stock and Yogo, 2005). The estimated coefficient of Tourism Income in the second stage in Column (2) indicates that tourism development still significantly reduces household energy poverty when IV1 is used to attenuate endogeneity problems.

To address the fact that IV1 does not vary over time, we use the logarithm of the number of A-rated tourist attractions in the city where the household is located as an instrumental variable (IV2). Class A tourist attractions are also rated by the government, and an average of 13.7% have been added each year in the last decade. IV2 satisfies the exclusivity and correlation of the instrumental variables, which is similar to IV1. Column (3) shows the first-stage estimation results of 2SLS based on two-way fixed effects. The results show that IV2 is significantly and positively associated with tourism development, and the Cragg-Donald Wald F statistic is 145.522, which is greater than 16.38. The results in Column (4) indicate that tourism development significantly reduces household energy poverty after mitigating the endogeneity problem of the model.

Exhibit C. Results of endogeneity issues

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variables | (1) | (2) | (3) | (4) |
| IV-2SLS | | IV-2SLS(panel) | |
| First | Second | First | Second |
| Tourism Income | EP | Tourism Income | EP |
| IV1 | 0.002\*\*\* |  |  |  |
|  | (0.000) |  |  |  |
| IV2 |  |  | 0.104\*\*\* |  |
|  |  |  | (0.009) |  |
| Tourism Income |  | -0.195\*\*\* |  | -0.364\*\*\* |
|  |  | (0.025) |  | (0.097) |
| Control variables | YES | YES | YES | YES |
| Time fixed effect | YES | YES | YES | YES |
| Individual fixed effect | NO | NO | YES | YES |
| Cragg-Donald Wald F statistic | 714.787 |  | 145.522 |  |
| Observations | 30,822 | 30,822 | 27,566 | 27,566 |

Note: Since IV1 is time-invariant, columns (1) and (2) do not control for individual fixed effects.

**Appendix D. Robustness checks**

This subsection examines the reliability of the baseline estimates by varying the energy poverty measure and the estimation methodology of the model.

First, we replace EP in Equation (1) with the income-expenditure approach (EP1) and the multidimensional poverty indicator (EP2). The estimates in Columns (1) and (2) of Exhibit D are both significantly negative, indicating that tourism development continues to significantly reduce household energy poverty after using EP1 and EP2 to measure energy poverty.

Second, we include city-level fixed effects in Equation (1) to reduce the impact of city economic and social factors that do not vary over time in the model estimation. Column (3) of Exhibit D presents the estimation results of this approach. It can be seen that the results remain robust.

Thirdly, we use logit model estimates that better apply the explanatory variables as dichotomous dummy variables. Column (4) is the estimation of a logit model that does not control for individual fixed effects but additionally controls for city fixed effects. Column (5) presents the estimation results of the logit model with two-way fixed effects. The results in both columns indicate that tourism development significantly reduces household energy poverty.

Fourthly, we use the original tourism income as an explanatory variable and the results remain unchanged (column 6).

Finally, while the proportion of elderly people and children in households may influence energy poverty, these variables are generally stable over the short term and are effectively controlled by household fixed effects. Additionally, our model includes household size as a control variable, which captures variations in household composition. Robustness checks incorporating the proportion of elderly people and children confirm that our main results are not significantly affected by their exclusion (column 7).

Exhibit D. Robustness checks results

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| EP1 | EP2 | City FE | logit | FE-xtlogit |  |  |
| Tourism Income | -0.018\*\* | -0.026\*\*\* | -0.028\*\*\* | -0.109\* | -0.257\*\*\* |  | -0.028\*\*\* |
|  | (0.009) | (0.005) | (0.009) | (0.065) | (0.079) |  | (0.010) |
| Or\_Tourism Income |  |  |  |  |  | -0.057\*\*\* |  |
|  |  |  |  |  |  | (0.008) |  |
| Control variables | YES | YES | YES | YES | YES | YES | YES |
| Time fixed effect | YES | YES | YES | YES | YES | YES | YES |
| Individual fixed effect | YES | YES | YES | NO | YES | YES | YES |
| City fixed effect | NO | NO | YES | YES | NO | NO |  |
| Observations | 30,822 | 30,822 | 30,822 | 30,822 | 9,636 | 30,822 | 30,822 |
| R-squared | 0.030 | 0.016 | 0.031 | - | - | 0.018 | 0.035 |

Note: \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001; Standard errors at the clustered household level are reported in parentheses; Results not reported for other control variables.

**Appendix E. Mechanism**

**Purchasing power**

Household income is key to the purchasing power of household energy. Much of the literature suggests that tourism development can raise the incomes of local residents by expanding employment and increasing the demand for tourism-related artisanal and agricultural products. Columns (1) through (3) of Exhibit E report the mechanisms through which tourism development affects energy poverty reduction by boosting household income. Here, we use a three-step model for mechanism testing (see Baron and Kenny, 1986). The first step tests the effect of tourism development on household income (Column 1). The second step tests the effect of household income on energy poverty (Column 2). The third step incorporates the household income variable into Equation (1) to observe the change in the coefficient of the Tourism Income variable (Column 3). If the coefficient of Tourism Income decreases or is statistically insignificant compared to the coefficient in Column (6) of Table 1, the mechanism is shown to hold. It is easy to see from the results in Columns (1) through (3) that tourism development does reduce household energy poverty by increasing household income. This finding further supports that tourism development has a poverty reduction effect in China.

**Energy accessibility**

Tourism depends on a good ecological environment, which partly depends on energy efficiency and the degree of green innovation in the city. To promote high-quality and sustainable development of the local tourism industry, the local government restricts the development of high-polluting, high-energy-consuming and inefficient enterprises to optimize the ecological environment (Nguyen and Su, 2021). This will increase the overall energy efficiency of the city as well as the degree of green innovation in the city, which will help reduce the energy consumption of enterprises and thus increase the energy accessibility of residents. Due to data limitations, we cannot directly observe that households are constrained at the energy supply level. Therefore, we explore the impact of tourism on household energy accessibility in terms of the energy use efficiency of the energy suppliers. To estimate the impact of tourism development on household energy accessibility, we constructed city panel data from 2011-2019. The estimated model is as follows:

 （2）

In the above equation, Mechanism is the mechanism variable, including two variables of urban energy consumption, energy use efficiency. Control includes the city's population size, industrial structure, financial level, technological level and human capital investment. The meanings of the remaining variables and parameters are the same as those in Equation (1).

The results in column (4) of Exhibit E report the estimated results of tourism development on urban energy consumption. Energy consumption is the total amount of electricity, oil and gas consumption converted to standard coal and is expressed in logarithms in the estimation model. Not surprisingly, the estimated coefficients for tourism development are not significant, indicating that tourism development does not increase energy consumption in the city. However, although the total amount did not change it does not mean that tourism did not deprive rural households of their energy supply. For this reason, columns (5) and (6) further estimate the "savings" from tourism development on urban energy resources. Column (5) presents the results of the estimation of tourism development on the energy use efficiency of the city. The energy utilization efficiency variable is measured by energy consumption/gross local product. The results show that the estimated coefficients are significantly positively correlated at the 5% level, implying that tourism development significantly improves energy use efficiency in cities.

To further demonstrate that tourism development can improve energy efficiency, we expand the conclusion by looking at green technological innovation and public transportation development. Column (6) reports the estimated results of the impact of tourism development on green innovation in the city. The city green innovation variable is measured by the logarithm of the number of green patent applications in the city. The results show that tourism development promotes urban green innovation at the 5% significance level. Column (7) estimates the impact of tourism development on the number of public transportation vehicles, and the results indicate that tourism development can promote public transportation development. Columns (4)-(7) show that tourism development, rather than increasing local energy consumption, improves energy efficiency thanks to its contribution to the advancement of green technologies and the development of public transportation.

Exhibit E. Mechanism analysis results

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) |  | (4) | (5) | (6) | (7) |
| Purchasing power | | |  | Energy accessibility | | | |
| Household income | Energy poverty | |  | Energy consumption | Energy efficiency | Green innovation | Public transportation |
| Tourism Income | 0.067\*\*\* |  | -0.014\* |  | -0.002 | 0.617\*\* | 0.132\*\* | 0.078\*\*\* |
|  | (0.024) |  | (0.008) |  | (0.032) | (0.308) | (0.061) | (0.027) |
| Household income |  | -0.207\*\*\* | -0.206\*\*\* |  |  |  |  |  |
|  |  | (0.005) | (0.005) |  |  |  |  |  |
| N | 30,822 | 30,822 | 30,822 |  | 2,481 | 2,481 | 2,482 | 2,456 |

Note: Columns (4)- (7) present estimates at the prefecture-level; All other variables are controlled.

References

Nguyen CP, Dinh Su T (2021) Tourism, institutional quality, and environmental sustainability. *Sustain Prod Consum* 28:786-801.

Prada-Trigo J, Jesús CPG, López-Guzmán T, Sandra EPL (2016) Tourism and Motivation in Cultural Destinations: towards those Visitors Attracted by Intangible Heritage. *Almatourism (Rimini)* 7(14):17-37.

Stock JH, Yogo M (2005) Testing for Weak Instruments in Linear IV Regression, in *In Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*,(Stock JHE, Andrews DWK eds), pp 80-108. Cambridge University Press.