Appendix 2. Test results of instrumental variable

1. Test for the correlation criterion

The CMP command in STATA 13.0 was used to handle the IV-OPROBIT model. To test the validity of the instrumental variable, this paper uses the first-stage regression of 2SLS to test the correlation between the instrumental variable DIS and the core independent variable LMIGRATION. The statistical value of the F-test is 44.487, which satisfies the empirical rule that the statistical value of the first-stage F-test should exceed 10 (Wooldridge 2016). This result indicates that the instrumental variable DIS is well correlated with the core independent variable LMIGRATION and there is no weak instrumental variable problem.

2. Test for the independence criterion

To further verify whether the instrumental variable is uncorrelated with the stochastic disturbance term, after the estimation results of IV-Oprobit are output, the DIS is used to carry out regression on the residual predicted value of model (5). The result in Fig. A3-1 shows that the regression fitting values are all close to horizontal lines, indicating that DIS is almost uncorrelated with the stochastic disturbance term. Therefore, the authors believe that the selected instrumental variable is reasonable.
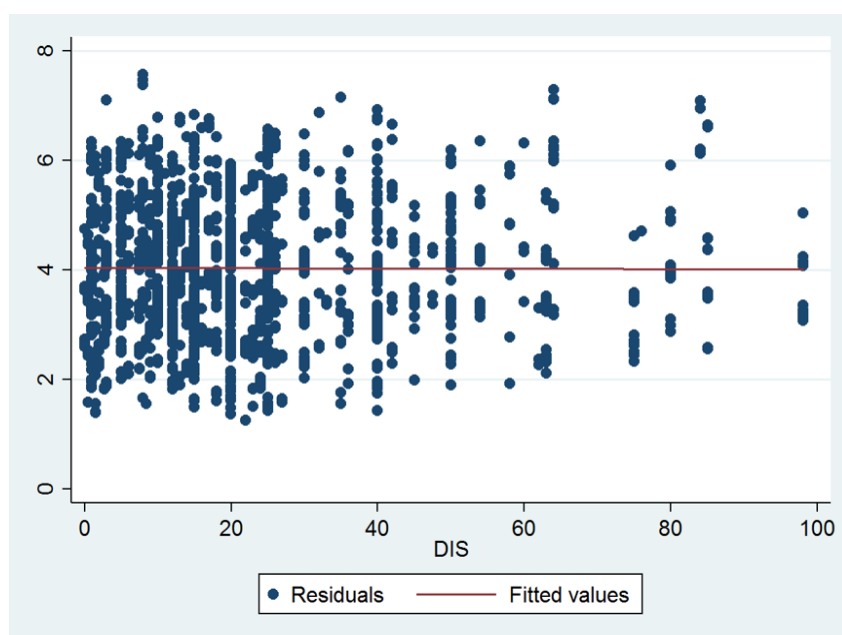

Fig. A3-1 Test results of the correlation between instrumental variable and error terms

3. Test for the exclusion criterion

Theoretically, "the distance from the village to the county center" (DIS) belongs to the geographical distance variable, which will not change with changes of time, space, economy, and society over time. Thus, the authors assumed that there is no evidence that the geographical distance and dependent variable "the frequency of farmers' participation in the construction and repair of irrigation facilities" (ICA) are directly related.

According to existing literature, the distance between village and county center does not constitute a reasonable factor affecting farmers' participation in irrigation collective action. For example, while it has been shown that collective action on irrigation is better organized in areas close to market centers (Wang et al. 2021), other scholars showed that irrigation collective action can also be successful in mountainous areas far from the county center (Chai and Zeng. 2018).

From the sample selected in this paper, as shown in Fig. A3-2, if DIS is divided into five sections of 0-20 km, 20-30 km, 40-60 km, 60-80 km, and above 80 km, there is no significant difference in the degree of farmers participating in irrigation collective action in each section. This implies that there is no case where a certain degree of participation is concentrated in a certain distance section. Therefore, there is no significant evidence for a direct correlation between DIS and ICA.
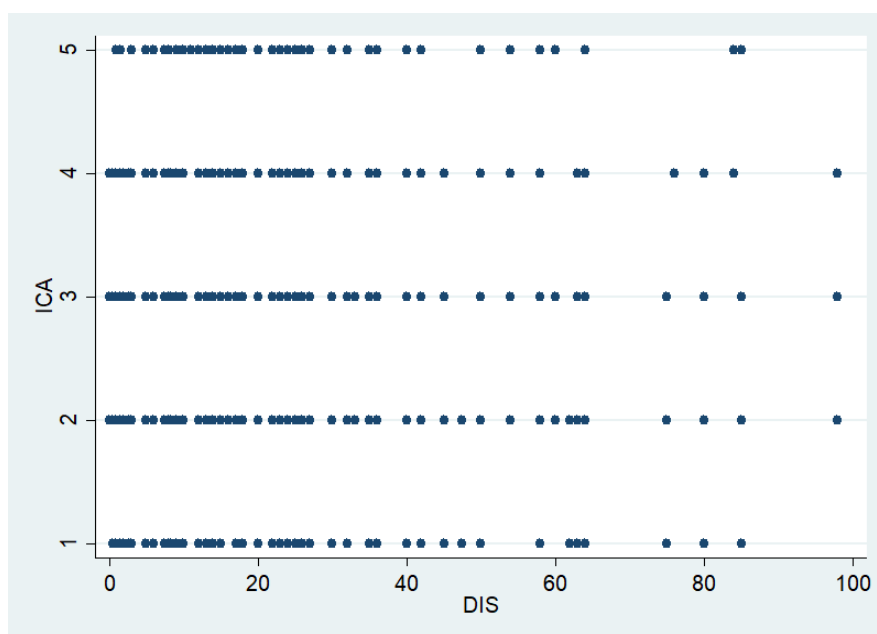
Fig. A3-2 Correlation test between instrumental variable and independent variable

Finally, From the perspective of empirical testing, the mediating effect model (see in Appendix 2) can also be used to verify whether the DIS affects ICA solely through its impact on LMIGRATION (based on the utilized samples). With ICA as dependent variable, DIS as independent variable, and LMIGRATION as mediating variable, a mediating effect model is constructed as shown in Fig. A3-3. This figure also reports the estimated results. The estimated results of the correlation between DIS, LMIGRATION, and ICA show that the direct effect of DIS on ICA is not significant, indicating that no evidence has been found that DIS directly affects ICA. The impact of DIS on LMIGRATION and the impact of LMIGRATION on ICA are statistically significant, indicating that DIS will impact ICA through LMIGRATION. Such a result is called a complete mediation effect in the estimation of the mediation effect model, namely, DIS only affects the outcome variable ICA through LMIGRATION.
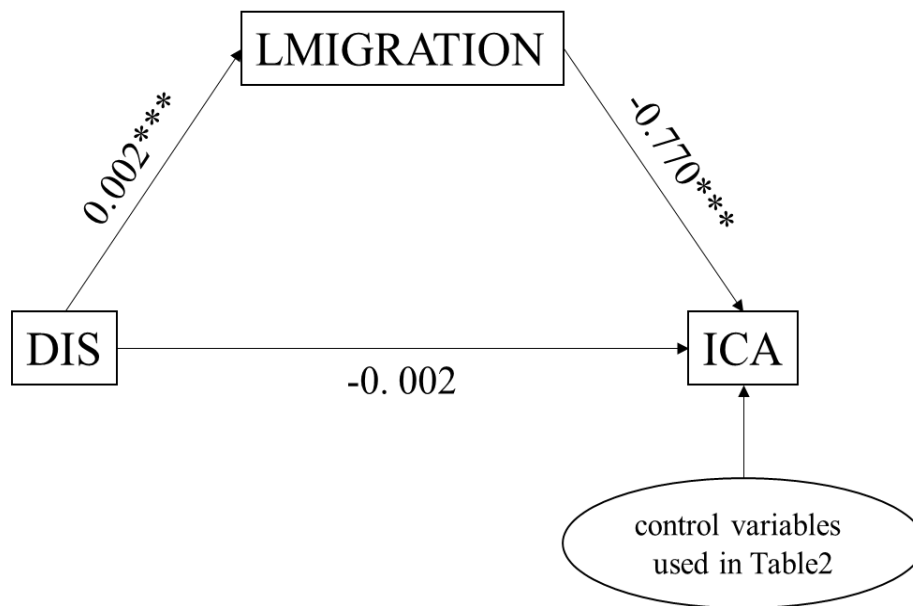
* p < 0.1, ** p < 0.05, *** p < 0.01

Fig. A3-3 The distance from the village to the county center (DIS) only affects the outcome variable irrigation collective action (ICA) through LMIGRATION.

4. Test summary

Based on the above discussion, the authors believe that the instrumental variable selected in this paper meets all requirements of a reasonable instrumental variable.