**Notes on data collection and context for project components**

The main goal of the project was to collaborate with local communities to develop interventions, called Children’s Centres. The Centres provide activities (sport, educational, music, crafts) which were tailored to, and met the needs of communities in both Zhejiang and Guizhou provinces. These provinces are geographically, socio-economically and culturally very different. The county level Women’s Federation (a governmental organisation) took responsibility for the project in the villages. The project was carried-out in 20 villages in Zhejiang, and 20 in Guizhou. Because of different conditions and different approaches to the intervention across the villages, direct comparisons of data are therefore not only difficult, but also inappropriate. But satisfaction with the interventions was overall very high and the Centers continue to the present in two counties.

**Baseline survey**

This was conducted in village and township schools in three counties in Zhejiang in 2015 and two counties in Guizhou in 2016. Sschool attendance across all five counties to age 16 is around 98%, so schools provide a highly appropriate sampling frame. The dataset comprises data for children aged 8 to 16 including left behind children (LBC living with neither parent) SC (children living with one parent) and RC children living with both parents.

The bulk of the questionnaires were completed by children themselves with final questions (details of left behind status) completed face-to-face by researchers. All children agreed to participate, with the first page of the questionnaire comprising consent. However, we are aware that to refuse to participate in the school setting would be very rare.

Data includes sociodemographic, economic, psychosocial, emotional and behavioural variables using the Strengths and Difficulties Questionnaire. There are data for 3796 children in total: 1553 in Zhejiang and 2243 in Guizhou, 1553 children living with neither parent, 866 living with one parent and 1378 living with both parents.

The alphabetical columns of the dataset match the alphabetical codes in the questionnaire. Additional column codes are provided to ensure variables can be identified when the data are used in Statistical Packages. Names of counties and villages are in Chinese.

**Post-intervention Surveys**

These were carried-out between early 2017 and late 2018 among left behind children in the village schools to determine the impact (if any) on the children attending the Children’s Centers. A version of the original survey tool was used, but with added questions about use of the Centers, which were delivered verbally by researchers. Children who did not attend the Centers more than 30% of time served as the control group. The names of the Children’s Centers are in Chinese, and do identify the villages. The sample was 435 in Zhejiang and 484 in Guizhou.

**Post-intervention Interviews**

The interview components of the surveys were not recorded. But interviews with children (n=18) and parents (n=22) at Children’s Centers were recorded. The cosio-demographic background of each informant is provided in a separate document in the Zip file. Some adult informants did not provide consent but notes were taken. The aim of the interviews was to determine satisfaction with the Centers and suggestions for improvement. Interviews were conducted with other stakeholders (Village leaders, Women’s Federation representatives) but none were recorded at the request of informants.

If any researcher has interest in this dataset, we are very happy to provide more information.