**Data Files for " Arranged Cohabitation among Chinese Muslims"**

This dataset encompasses three distinct sets of data analyzed in the study, namely the survey data on cohabitation in Gansu, China, the census data on underage births and premarital births in China, and the survey data on underage births in China.

This data is available at yuxie.com and Princeton DataSpace.

**Survey Data on Cohabitation in Gansu, China**

The first part of the dataset comprises the analysis shown in Figure 2 and Table 1, which is collected from the 2007 Developmental Idealism Survey in Gansu, China.

We append the data from the survey. The analytic data at the micro level can be found in [Gansu\_DI\_2007.dta] and the questionnaire in [CHINA\_2007\_QUESTIONNAIRE\_ENGLISH.pdf], sufficing the replication of the results presented in Figure 2 and Table 1. We disclose the relevant variables used in the research, including whether the respondent had ever cohabited, age at first marriage, ethnicity, years of schooling, personal annual income, self-rated social class, whether the respondent lived in an urban community, age, gender, whether the respondent had ever married, a composite scale measuring endorsement of modern family behaviors, and the frequency of religious service attendance.

Our study object is to compare and understand the prevalence of cohabitation between Han and Chinese Muslims, so our sample is restricted to Han and two Muslim groups, namely, Hui and Dongxiang, which leaves us with 529 Han, 47 Hui, and 36 Dongxiang.

Data columns description of [Gansu\_DI\_2007.dta]:

* evcohab: indicator for cohabitation experience
* agefirma: age at first marriage
* ethnicity: official Chinese ethnic designation
* edu: years of schooling
* lninc: logged personal annual income
* socclass: self-rated social class
* urban: indicator for urban residence
* age: personal age
* gender: sex
* evmar: indicator for marriage experience
* famval: index for family values
* attd4: fequency of religious services attendance

**Census Data on Underage Births and Premarital Births**

The second part of the datasets provides information used in Table 2 and Supplementary Materials 7, involving the China census data in 2000.

The National Bureau of Statistics conducts the census in China, collecting data from all households in the country. We used a 1% sample of the 2000 census data extracted from the Integrated Public Use Microdata Series prepared by the Minnesota Population Center. We aimed to use this dataset to compare underage births and premarital between Han and the ten Muslim ethnic groups. Percentage of first birth under legal age was calculated as the proportion of women who had their first child before turning 20, among those who had ever given birth and who were 50 years old or younger in 2000, as they were subject to the restricted minimum marriage age. Percentage of birth before marriage was calculated, for the same women respondents, as the percentage of women whose first child was born prior to the year of first marriage. Northwest region includes Gansu, Qinghai, Ningxia, and Xinjiang, where most Chinese Muslims live. The analysis is restricted to the married women at or below age 50 from Han or Muslim ethnic groups, yielding an analytic sample of 2,322,247 [China\_census\_2000.dta].

Data columns description of [China\_census\_2000.dta]:

* ilbirth: indicator for birth under minimum legal marriage age
* pmbirth: indicator for premarital birth
* ethniccn: official Chinese ethnic designation
* prov: province in China
* prov26: indicator for Gansu Province
* prov27: indicator for Qinghai Province
* prov28: indicator for Ningxia Hui Autonomous Region
* prov29: indicator for Xinjiang Uyghur Autonomous Region

**Survey Data on Underage Births**

The third part of the datasets provides information used in Supplementary Materials 7, involving the 2010 wave of the CFPS.

The China Family Panel Studies (CFPS), conducted by Peking University, is a nationally representative, longitudinal, comprehensive, and biennial social survey started in 2010. We aimed to use this dataset to demonstrate consistent patterns of underage childbirths as shown in the aforementioned census-based results. Percentage of first birth under legal age was calculated as the percentage of respondents who had their child before age 20 for women and 22 for men among respondents who had ever had children. We further restricted the sample to those who were 60 years old or younger in 2010, as they were subject to the restricted minimum marriage age, which leaves us with a sample of 19,855 Han, 173 Hui, and 29 Dongxiang [CFPS\_2010.dta].

Data columns description of [CFPS\_2010.dta]:

* ethnicity: official Chinese ethnic designation
* agefirbirth: age at first birth
* ilebirth: Indicator for birth under minimum legal marriage age
* gansu: indicator for Gansu Province

**Data Publisher**

The 2007 Developmental Idealism Survey in Gansu, China is part of the joint conceptualization and data collection efforts of the Developmental Idealism Studies (https://developmentalidealism.org/) program directed by Arland Thornton at the University of Michigan. Census data was conducted by the National Bureau of Statistics of China. We use a 1% sample of the 2000 census data extracted from the Integrated Public Use Microdata Series at the Minnesota Population Center. China Family Panel Studies (CFPS) is conducted by the Institute of Social Science, Peking University.