CHFS Quality Control in Data Collection

Keeping the quality of the data collected on the field becomes the top priority of the CHFS project. Much of our effort in quality control is put into in the following three aspects.

1. Interview training and community relations

We carried out a series of training sessions for our mapping technicians and interviewers. Through classroom lectures, case simulations, field practice, and evaluation and feedback, they were able to acquire the principles, methods and skills of sampling and interviewing. Such trainings are essential for avoiding various errors one may make in the fieldwork.

(1) Training

The tasks of onsite mapping and sampling require multiple skills, including a good sense of direction, the ability of drawing precisely, the operational capacity of computer and software, and on top of all that, the ethic of hardworking and persistence. All the mapping technicians were selected from the high-caliber undergraduate and graduate students of SWUFE and went through rigorous trainings in GIS, drawing, sampling, and software operation. On the site, they compare the base map provided in the system with what they observe in the reality, make corrections and add necessary information to the base map. For the purposes of CHFS, they also make distinctions between residential and non-residential buildings and add information to the electronic map regarding the number of buildings and households, empty buildings, unoccupied homes, etc.

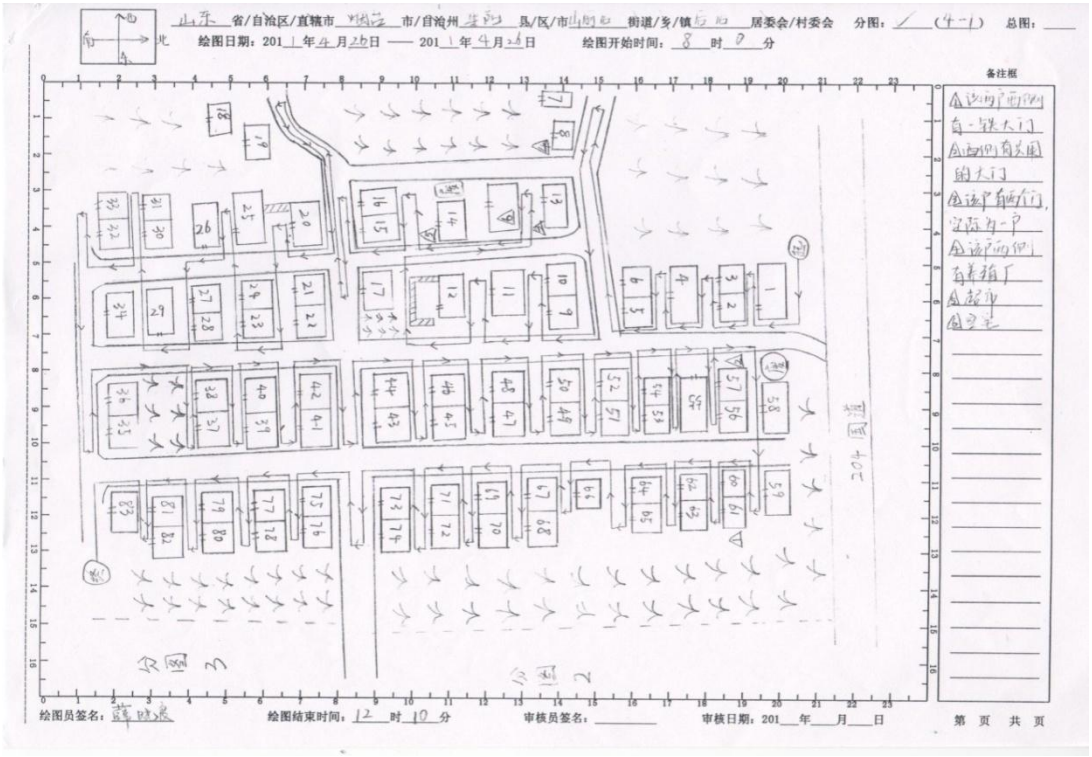

Figure 1 Map Drawing by Hand

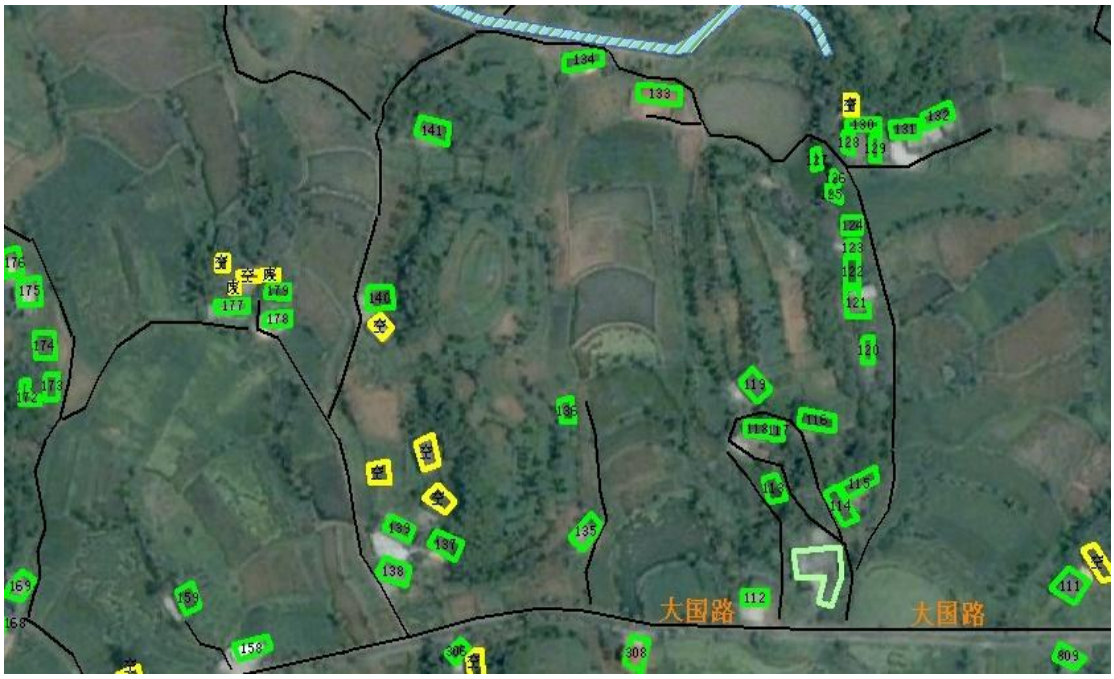

Figure 2 Maps Drawing by Computer

The interviewers of CHFS mainly consist of undergraduates and graduate students from SWUFE. They are well versed in economics, understand the questionnaire thoroughly, and are able to explain to and communicate with respondents effectively. They all received the following trainings before going to the field.

First, obtain interview skills. It includes how to identify qualified interviewees and build up trustful relationship, how to ask and explain questions in a precise and neutral way, how to deal with emergent issues during interview, and how to save and transmit data while keeping interviewees' information confidentially after interview.

Second, understand the questionnaire. We divided the interviewers into small groups, familiarized them with the questions, and made sure each of them understood the questions. We also employed PPT, video clips and other multimedia technologies to make the process lively and vivid. Interviewers conducted mock interview in the classroom, observed each other's performance, and discussed how to do the tasks in a better way.

Third, the CAPI (Computer-assisted Personal Interviewing) system and the corresponding survey management system. Each interviewer is given a laptop with the CAPI system and the management system pre-installed. Sufficient amount of time was spent to teach the interviewers to become efficient users of the equipment as well as the software. They were also trained to make remarks and use various shortcuts during interviews.

Fourth, trial interviews on the field. Besides classroom trainings, we brought interviewers to conduct trial interviews in surrounding neighborhoods to test their acquisition extent of interview skills and questionnaire content. We provided extensive and detailed feedback to each of them to help them further improve the skills.

In summary, our 232 mapping technicians experienced five round training with average training time of 42 hours. Meanwhile, we trained 343 interviewers, each of whom received average 80-hour training. Each trainee was evaluated by rigorous standards at the end of training.

The unqualified trainee would be asked to take the training again or dismissed. The positions of field auditors are mainly held by the doctoral students in SWUFE. As those who shoulder important responsibilities of managing interviewers and the fieldwork, each field auditor received over 30 hours' rigorous and extensive training. They all became proficient users of the interview monitoring system, the interview assignment system, and the CAPI system. The rigorous training and evaluation procedures produced first-rate field auditors and interviewers, which built a solid foundation for collecting high-quality data at later stages.

(2) Community network

One of the major obstacles of household interview is to earn respondents' understanding and trust. To overcome this, our interviewers were often brought in and introduced to the respondents by the staff working for the selected residential committee/village. Their presence was proved extremely helpful when the respondent lacked enthusiasm to cooperate. The explanations and persuasions from the staff greatly decreased the refusal rate of the survey. The local branches of the People's Bank of China also helped tremendously in making arrangement and providing logistic assistance.

2. Quality control over interview

Based on the framework and design concept of the cutting-edge CAPI (Computer-assisted Personal Interviewing) system, CHFS develops a proprietary interview system and management platform. This integrated system provides a full package for computer-based household interview. This innovation can effectively decrease potential man-made non-sampling errors by presetting the range of possible answers, catching typing errors, and avoiding errors in skipping questions, etc. Meanwhile, it helps to keep the data confidential and make it accessible in real time. All these significantly improve the data quality.

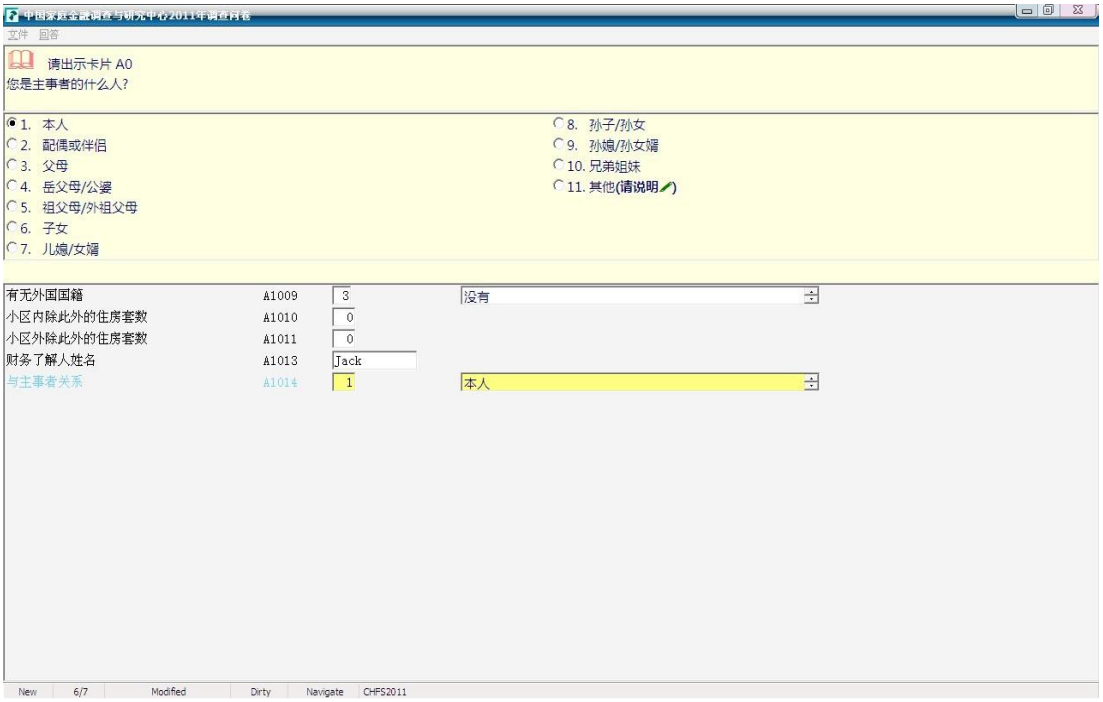

Figure 3 Questionnaire System of CAPI

Besides the CAPI system, we have also designed a comprehensive system for quality control to keep man-made errors to the minimum level. The system consists of the following aspects:

(1) Stringent management on sampled cases

The management, allocation and replacement of sampled cases during the fieldwork are all recorded by the computer system. To a great extent, this measure prevents interviewers from replacing difficult cases with easier ones at will. This sample management system combined with the CAPI is the key to ensure the randomness and representativeness of data collected.

The main functions of the system include sample construction, sample allocation, sample replacement, sample maintenance, sample tracking, personnel monitoring, double-checking sample allocation and execution, etc. In doing so, we follow the following steps:

First, we construct the sample data based on the third stage sampling which provides information on the assigned numbers for households and base maps, and household postal addresses, etc.

Second, we input the information of interviewers and fieldwork auditors and team them up.

Third, we establish rules for allocating interview cases and assign task to fieldwork auditors and interviewers.

Four, we collect real time survey data and examine the data quality in some preliminary ways.

We also keep record about the time of sampled cases being assigned and interviewed, and the sample replacement information.

Figure 4 The Survey Management System

(2) A detailed survey management system

The survey management system plays four functions: (1) to receive data from the system designed for managing and tracking sampled cases; (2) to get in touch with the selected household and collect relevant data; (3) to conduct survey data collection through the Blaise questionnaire; (4)

to transmit the data back to the server.

The information we collect at this stage includes not only the interview data, but also information about how the interview is initiated, duration of the encounter with selected households, outcome of the encounter, way of interaction, surrounding environment, appointment time, and the way of appointment. We also receive information about the exact time when a particular question is asked, how long it takes to answer the question, sequence of the questions and answers, records for the use of keyboard or mouse, and the time of finished cases being transmitted to the server, etc.

Through data synchronization, the computers used in the interview can connect to the server using VPN and transmit the data back the server. This enables the researchers on the backstage to perform real time check and analysis of the data.

3. Checking the data

When interviewers are on the field to interact with selected households, we ask the interviewers to record relevant information, such as time spent in knocking the door, reaction of the local guide or interviewees. Such information helps us to understand possible reasons for refusals and provides information for future actions. It also deters interviewers from replacing the selected cases at will.

The CAPI system offers a recording function. Among many things, it records all conversations during interviews and all movements of the keyboard and the mouse. In other words, the paradata of the interview process is well recorded. With internet connection, these data are transmitted to the server, which makes real-time monitoring possible. In doing so, we have done the following:

(1) We selected a certain portion of finished interviews, listened to the recordings and double-checked with the interviewees by phone.

(2) We used statistical tools to check data quality. For the outliers, we listened to the recording and replay the computer screens of the interviews. When necessary, we called the interviewees by phone or revisited them in person.

(3) We conducted call-back interviews to all the interviewed households.

Those efforts help to smooth the fieldwork process and improve the data quality as well.