DOCUMENTATION OF A MULTI-TOPIC QUESTIONNAIRE-BASED SURVEY ON SUSTAINABLE RESOURCE USE IN RURAL CHINA

XIANGPING JIA, GERTRUD BUCHENRIEDER

DISCUSSION PAPER NO. 102
2007
Xiangping Jia is a Chinese researcher at the Department of Rural Development Theory and Policy (490a), University of Hohenheim. His educational background is agricultural economics and institutional economics of development. His doctoral research is about credit rationing and institutional constraint, with micro evidence in rural China.

Address: Department of Rural Development Economics and Policy (490a)
University of Hohenheim
70593 Stuttgart
Germany

Phone: (+49) 711-45922556
Fax: (+49) 711-45923934
E-mail: jjoseph@uni-hohenheim.de

The paper was reviewed by Dr. Jana Fritzsch and Dr. Martin Petrick (Department 'External Environment for Agriculture and Policy Analysis') at the Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO).

Phone: (+49) (0) 345/29 28 100
Fax: (+49) (0) 345/29 28 199
E-mail: buchenrieder@iamo.de
Internet: http://www.iamo.de

Discussion Papers are interim reports on work of the Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO) and have received only limited reviews. Views or opinions expressed in them do not necessarily represent those of IAMO. Comments are welcome and should be addressed directly to the author(s).

The series Discussion Papers is edited by:
Prof. Dr. Alfons Balmann (IAMO)
Prof. Dr. Gertrud Buchenrieder (IAMO)
Prof. Dr. Thomas Glauben (IAMO)

ISSN 1438-2172
ABSTRACT

This Discussion Paper documents a survey conducted in rural China in 2005. A multi-topic survey funded by the Sino-German international Training Program "Sustainable Resource Use in North China", the project covers farm management, land property rights and rural credit access on the North China Plain.

From a credit standpoint, this paper covers questionnaire design, sample, survey implementation, data entry, a brief assessment of the overall experience, as well as lessons drawn.

JEL: C81, C93, D19

Keywords: Questionnaire, Survey, Credit access, Household model, Categorical variables, Stata.

ZUSAMMENFASSUNG

DOKUMENTATION EINER AUF UMFRAGEN BASIERENDEN ERHEBUNG IM LÄNDLICHEN CHINA: AUS DER PERSPEKTIVE DER VERFÜGBARKEIT VON KREDITEN


Hinsichtlich der Kreditverfügbarkeit wird die Ausgestaltung des Fragebogens, die Stichprobenerhebung, die Umfrageimplementierung, der Dateneintrag, die Ausreißer und Fehlwerte betrachtet sowie eine kurze Aussage zu den Erfahrungen und Erkenntnissen gemacht.

JEL: C81, C93, D19

Schlüsselwörter: Fragebogen, Erhebung, Kreditverfügbarkeit, Datenbearbeitung.
**CONTENTS**

Abstract .............................................................................................................................. 3

Zusammenfassung ............................................................................................................. 3

List of tables ................................................................................................................. 6
List of figures ...................................................................................................................... 6

Summary ............................................................................................................................ 7

1 Introduction ................................................................................................................... 7

2 Survey design ................................................................................................................. 7

3 Research questions ........................................................................................................ 8
   3.1 General background ................................................................................................................ 8
   3.2 Focused questions - Credit access ......................................................................................... 9

4 Questionnaire ................................................................................................................. 9
   4.1 Questionnaire structure ....................................................................................................... 9
   4.2 Cropping patterns .............................................................................................................. 11
   4.3 Roster module ....................................................................................................................... 12
   4.4 Credit modules ...................................................................................................................... 12
   4.5 Other modules .................................................................................................................. ....... 13
   4.6 Unit issue ..................................................................................................................... ........ 14

5 Sampling ......................................................................................................................... 15

6 Survey implementation ................................................................................................. 16

7 Data entry and data management ................................................................................ 17
   7.1 Overview of data entry ......................................................................................................... 17
   7.2 Categorical information and value label ............................................................................. 18

8 Experience and lessons .................................................................................................. 18

References .......................................................................................................................... 20
LIST OF TABLES
Table 1: Cropping calendar........................................................................................................ 11
Table 2: Composition of household revenue ............................................................................. 14
Table 3: Sampling profile ........................................................................................................ 15
Table 4: Survey cost components ........................................................................................ 17

LIST OF FIGURES
Figure 1: Planning questionnaire-based survey ........................................................................ 8
Figure 2: Questionnaire design ............................................................................................... 10
Figure 3: Survey map ............................................................................................................... 16
Figure 4: Flowchart of data entry ........................................................................................... 18
SUMMARY

This analysis is sponsored by the Sino-German Project "Sustainable Resource Use in North China". The field survey was conducted from October 2004 to July 2005, and covered sustainable farming systems, land and water property rights, and rural credit access on the North China Plain. In this multi-topic survey, the authors specifically focus on rural households’ credit access and its impacts on structural adjustment in China’s small-scale farming sector. Though the multi-topic survey was advantageous in some respects, the selective sampling technique still limits generalization of the results.

This paper aims to document the procedures undertaken for the survey. First, the design of the questionnaire-based survey is streamlined with a flowchart in Section 2. Then research questions are presented in Section 3. Section 5 describes the questionnaire’s structure. Sampling and survey implementation are reported in Section 5 and 6. Then, after having discussed the employed data entry procedure in Section 7, the general experience and lessons learned are shared in the last section.

1 INTRODUCTION

Data at the household level in developing countries provide a means for researchers to test theoretical hypotheses, and offer barometers for policymakers to assess public policies. Agricultural households, however, are more complex than other micro-agents because they are consumers, producers and laborers concurrently. This makes household surveys extremely challenging, not only because of funding but also due to institutional capacities. Development economics at the micro-level is, therefore, mired by a lack of suitable data and an inability to test hypotheses. Many policy-oriented surveys that cost huge sums of money fall short of expectations because of the ill-designed survey schemes and lack of transparency.

Many of the individual surveys carried out by economists and anthropologists fill existing gaps and enrich empirical analysis. However, few efforts have been made to document various surveys and to make similar works more reproducible. Such knowledge is poorly communicated and scientists have to learning by doing. Motivated by such deficiencies, this paper will attempt to share our survey experiences and lessons in rural China.

2 SURVEY DESIGN

Managing a survey calls for systematic and holistic knowledge of questionnaire design, data collection, database management, and data cleansing. Errors in design and execution cannot be undone, and such errors occasionally reduce the reliability of the database and the derived results.

---

1 This project was co-funded by the German Research Foundation (DFG) and the Ministry of Education (MOE), China, between University of Hohenheim (Stuttgart) and China Agricultural University (Beijing) in 2004 (http://www.uni-hohenheim.de/rtgchina). More information of the subproject related to this article is available at http://rtgchina.uni-hohenheim.de/index-Dateien/sp33.htm. The questionnaire is also available at http://rtgchina.uni-hohenheim.de/irtg-survey05.pdf.

Cooperation partners were Stephan Piotrowski and Roland Barning from the University of Hohenheim, and Jiao Wang and Ji Ma from the China Agricultural University. Each member had individually-targeted research topics. The survey focused on rural households on the North China Plain and the sampling regions were purposively selected to meet multi-topic needs.

2 Two exceptions are the institutions which have been engaged for years in developing high-end household questionnaires in developing countries with a variety of topics. One institution is the Living Standards Measurement Study (LSMS) of the World Bank <http://www.worldbank.org/lsms/>. The other is the "virtual center" for fieldwork in Development Economics <http://sticerd.lse.ac.uk/fieldwork/>. 
results. As shown in Figure 1, the first step in survey design is to clarify the objectives, the hypotheses, and even the proposed empirical models. After selecting the study area, questionnaire design (at both the household and regional level) becomes the main focus and is a time-consuming task. Then, scientific sampling needs to be designed to ensure the generalization of the derived results. Before actually using the survey data, pre-estimation must be carried out to deal with missing values, identify outliers, and classify samples in the light of research topics.

**Figure 1:** Planning questionnaire-based surveys

---

3 **RESEARCH QUESTIONS**

3.1 **General background**

The North China Plain spans seven provinces (Hebei, Henan, Shandong, Anhui, part of Jiangsu Province, Beijing and Tianjing municipalities) with a total population of 200 million people, a 50% share of China’s total wheat production and 33% of its maize production. The North China Plain is regarded as China’s granary due to its special cropping patterns and geographical characters (Cheng, 2004).

Currently, China as a whole is responsible for 27% of global nitrogen fertilizer consumption, compared to 15% in 1978 (Han and Zheng, 2004). In the North China Plain, the average cropping index is 145%, compared to 120% at the national level (Cheng, 2004). The environmental sustainability in the North China Plain is therefore questionable due to the intensive cropping pattern, over-used irrigation water, and excessive application of chemical fertilizer, which results in the serious environmental problems of water availability and pollution, soil contamination and erosion. For too long, policy makers in developing countries have failed to recognize that institutional reforms are just as effective, when it comes to economic growth, as improvements in physical infrastructure. Thus, there is a growing international consensus that sustainability should be understood in physical, economic, environmental and fiscal dimensions. Each of these dimensions has underlying institutions that are crucial for their effectiveness.
3.2 Focused questions – Credit access

Credit is essential to development strategies in a variety of ways; not only does it promote seasonal agricultural investments and facilitate the adoption of new technologies, better access to credit also smooths consumption. In the presence of underdeveloped formal insurance systems in rural areas, credit cushions the risks to agricultural households. Credit access therefore occupies a central place in development strategies. As a consequence, we find ubiquitous government intervention in rural credit markets to redress urban-biased macroeconomic policies. Government-led credit institutions, earmarked credit programs and subsidized interest rates, together with other variants of intervention, are widely implemented in developing countries. The performance of government-driven rural finance institutions and programs, however, has fallen short of expectations and many of the credit programs have become a costly drain on government budgets.

With its limited natural resources, China deserves praise for feeding over one-fifth of the world’s population and lifting millions of people out of poverty. Agriculture, however, has been downplayed in the national economy in favour of industry. To overcome long-term depression in the rural economy and ease increasing political unrest due to mounting inequality, credit policies, often believed to be effective tools to provide money to investors, become appealing. Given the failure of government-driven rural credit programs in other developing countries, China must ask itself: What will the outcomes of the current intervention be? Is credit policy a candidate for promoting agricultural development? What are the appropriate interventions to redress informational problems that are rooted in the credit market?

Based on the above information, we envisioned a multi-topic survey which consists of farming activities, land property, irrigation water and rural credit.

4 QUESTIONNAIRE

Compared with an individual household survey that examines a specific topic, a multi-topic survey has the advantages of lower average costs per topic as well as the availability of more instrumental variables. These variables are particularly useful for overcoming problems of endogeneity, which widely exists in rural household modeling, especially when markets are missing and household behaviors are interlinked (VAKIS and SADOULET, 2004). In addition, a multi-topic survey provides the possibility of measuring a wide spectrum of variables and paves the way for true interdisciplinary research, which integrates different disciplines from the very beginning, rather than pooling various modules without interaction.

In this section, after taking a snapshot of questionnaire design, the referenced period, which is based on the cropping calendar in the survey area, is reported. Then, different modules (roster, credit access, and farm management) are presented vis-à-vis questionnaire design.

4.1 Questionnaire structure

A questionnaire is a medium which is used to save information for future analysis (DOPPLER, 2005). After clarifying focused hypotheses, topical modules are drafted. These modules are then compiled into a consistent questionnaire. Arriving at a questionnaire’s final version is an iterative process, sometimes involving several pre-tests to adapt the questions to local institutions and terminology (see Figure 2 for an outline of questionnaire design).

The font and format of a questionnaire are important as well, because a good format minimizes potential mistakes during field survey and data entry. Although the number of pages that a questionnaire consists of should not be used as criteria for the time required for the
interview\textsuperscript{3}, interviewers and interviewees might be frustrated by similar questions, for instance questions relating to credit from formal and informal financial institutions. In addition, thumbing through pages increases the possibilities of error making. Therefore, a concise and compact questionnaire facilitates the field survey, error checking and data entry.

\textbf{Figure 2: Questionnaire design}

\begin{figure}
\centering
\includegraphics[width=0.8\textwidth]{questionnaire_design.png}
\caption{Questionnaire design}
\end{figure}

Several types of questionnaires were applied in this study: Questionnaires at the household and community level, with the latter including expert questionnaires for County Agricultural Bureaus, heads of villages, and Rural Credit Cooperatives at the township level. The household questionnaire combined five modules: (1) meta data, basic information of the rural household and a roster; (2) land property, land transfer and land reallocation; (3) farming activities (i.e., sowing, irrigating, fertilizing, weeding, harvesting and livestock raising) and self-run non-farming activities; (4) household balance and initial assets; (5) formal and informal credit. In the beginning, the questionnaires were customized individually within three different topics (land tenure, farming management and credit access). Then, all the modules were integrated and restructured into a single draft questionnaire (for the household as well as expert questionnaires). After translating the modules into Chinese, two pre-tests were carried out in order to adapt the questionnaires to local circumstances.

Another interesting aspect of this survey was the use of a separate and portable codebook. An alternative approach is to print response codes in the box where the question appears (as with LSMS). Using a separate code book allows more valued (measurable) answers, and thus permits exhausting all mutual and exclusive possibilities (GROS\textit{H et al.}, 2000: 65). Nevertheless, our lesson was that it is virtually impossible to think of all the possible answers in advance. A qualifier like "other (specify____)" would have been useful for recoding. Less thoughtful coding results in non-responses and might be mistakenly treated as missing in data analysis afterwards.

\textsuperscript{3} Due to the modular structure of the questionnaires, entire pages or modules might apply only to a handful of households, which implies that these can be skipped for the other households.
4.2 Cropping patterns

If questions are retrospective, a well-defined recall period for each questionnaire module or even question is necessary (Deaton, 1997). Generally it holds that the longer the recall period, the greater the danger of recall errors. The reference scale in this analysis was selected according to the cropping calendar due to the recognition that the decision-making of households with regard to production, consumption, savings, and credits are all made consistent with the cropping calendar.

The recall period in the study at hand was one year (September 2003 to October 2004), during which winter wheat and summer maize were planted on a rotating cycle. Though cropping patterns in the survey region differ from household to household, three systems are dominant and distinctive in terms of intensity of soil and water usage.

The first type, which is traditional in the North China Plain, is the rotating system between summer maize and winter wheat, which was sown in September 2003, the starting point in the referenced period. Summer maize was then sown in July 2004, immediately after the winter wheat harvest.

The second pattern is the mono-cropping of cotton. After plots were left fallow, vegetables were planted in the winter season in 2003, then cotton was sown in April 2004 and harvested in September 2004.

The third type, the most intensive and complicated one, is characterized as the combination of inter-cropping and the rotation of wheat and cotton. After harvesting at the last cropping season in September 2003, winter wheat was sown, with enough room left between the rows of wheat so that cotton could be sown there in April 2004. After wheat was harvested in July 2004, other vegetables or fruits were immediately planted on the furrows left by the harvested winter wheat. The shadowing effect between wheat and cotton is minimized to the lowest degree. The coping pattern in this type is remarkably land and labor intensive.

The rotating pattern is the most dominant and nearly 67% of rural households applied this type. Compared with the application ratio of about 21% for the inter-cropping pattern, 51% of rural households applied mono-cropping, under which plots can be left fallow for some time. However, different cropping patterns could be applied on different plots to balance the farming intensities; approximately 80% of inter-cropping farmers reported mono cropping.

Table 1: Cropping calendar

<table>
<thead>
<tr>
<th></th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>Aug</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td></td>
<td></td>
<td></td>
<td>H</td>
<td></td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maize</td>
<td></td>
<td></td>
<td></td>
<td>S</td>
<td></td>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cotton</td>
<td>S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: 'H' stands for harvest and 'S' stands for sowing. According to our database, wheat, maize and cotton are the three dominant crops in the sampling area; 95%, 81% and 48% of households reported planting these respective crops.

4 The rotating farming pattern is also called double cropping.
4.3 Roster module

Metadata, such as the name of interviewers and the date of interview, enable researchers to interpret the survey data and evaluate different survey procedures. Module 1 of the household questionnaire provides the administrative affiliation of households. Though in most cases nature villages (zi ran cun) are identical to administrative villages (xing zhen cun), exceptions are not rare, especially in mountainous areas where nature villages are formed by natural typology.

The roster section (Section 2.2), is one of the most significant and indispensable parts of the Agricultural Household Model (AHM). Here, one concept must be addressed. What is the "household"? How to define them? SADOULET and DE JANVRY (1995: 144) put it as follows:

"Identifying the decision-making unit which sets the strategy concerning the generation of income and the use of this income for consumption and reproduction".

Though there are no simple rules or universal principles to follow, three approaches can be used to define rural households in China: First, the people who live in or share the same abode or hearth; second, the family members who are registered in the "Local Resident Registration System" (called hukou); third, the family members who live in a house for some time (as well as others who have lived there for some time) but place their income at the disposal of the family head. While the first definition is too theoretical and ambiguous to be applied, the second excludes those people who are registered in rural areas but work outside of them. Given the important role of off-farm activities, the roster section pays substantial attention to displaying the detailed profile of household members.

The importance of the roster section is not only pragmatic, but also based on the theory of household models. In the case of a missing market, the "separability" (or recursive) relationship between consumption and production in neo-classical theory is undermined (SINGH et al., 1986). The factors that influence consumption, such as size and structure of families, also enter into the production function (BENJAMIN, 1992). Land, labor and capital markets in rural China are largely missing, but when they are present, they differ significantly. First, the existence of the "Local Resident Registration System" and the "Residence Permissive Certification System" in urban regions limit the rural migrant workers’ ability to both find jobs and to settle down. Second, though the reallocation system functions partially as a land market (ZHAO, 2002), the absence of an efficient rural land tenancy market exacerbates the situation of imperfect markets (CARTER and YAO, 2002; ROZELLE and LI, 1998). Lastly, compared with the labor and land market, the rural financial market is even more restrained due to the monopoly enjoyed by formal financial institutions and limitations from central government (HE, 1999; 2001). In the presence of market failure, demographic variables such as the size of households, the composition of family labor and the education of the family head etc., are significant in product decisions, and the property of "separability" between consumption and production no longer holds (BENJAMIN, 1992; CARTER and YAO, 2002; SADOULET and DE JANVRY, 1995; SINGH et al., 1986). Accordingly, the roster section is not only of practical use, but can also shed light on theoretical considerations.

4.4 Credit modules

Since the goal of decision-making involving credit is to optimize utility inter-temporally, it makes sense to include credit history into the questionnaire. The tables in the credit module were therefore designed in a matrix form, with credit records horizontally placed and the index
of credit listed vertically. Informal credit, which is pervasive in rural China, was designed in the same way.

The credit module is placed in the lattermost position of the household questionnaire because questions regarding incomes and credits are considered sensitive. Modules at the beginning are meant to "develop a rapport with the household members and reduce loss" to the utmost extent (GROSH and GLEWWE, 2000: 49).

Categorical information is used intensively in the credit module. Compared with quantitative questions, qualitative (or even open) questions offer researchers a better chance for observations and conversation (UDRY, 2003). Categorical variables translate the qualitative questions to measurable information.

4.5 Other modules

Modules 2 and 3 primarily describe farming activities. Land property rights are recorded plot by plot in Module 2.3. Module 3 covers sowing, irrigation, fertilizing, weeding, harvesting and marketing, livestock raising, and self-run business. Farm income can be generated by aggregating the gross income from cropping activities (grain and cash crop) and agricultural sidelines (forest, livestock and aquaculture). Income, especially off-farm income, has always been notoriously underreported in household surveys. In our experience with this study, of those rural households reporting more than one off-farm laborer in the referenced year, 49% of them claimed no off-farm income. The underestimation of reported annual off-farm income reinforces the importance of auxiliary data on off-farm labor in the Roster Module. Though underreporting is by no means avoided totally, the questions in the roster section are less sensitive, and more flexible and friendly.

In Module 7 of the questionnaire, other income information is recorded in a monthly (or annually) aggregated manner. Total income can be computed by summing up different items. Though such a recalling approach presumably results in less precise figures, the costs for expanding the questionnaire are prohibitive due to the length of constraint in the multi-topic survey. An alternative approach for estimating total income using relative terms, i.e., the percentage shares of various income sources including agricultural production, was applied by Poland (PETRICK, 2001) in a survey of its rural credit.

In Table 2, household income is grouped into six categories, which is largely consistent with official Chinese statistics. One exception is self-employment, which has not been captured by official statistics. Most farm households consume at least part of their production at home and some of their labor resources are used directly for home production. (BECKER, 1965), using Z-goods, which constitute household-produced products and home time, extends the concept of utility of households in his early literature. Not directly related to productive activities, home time includes cooking, rearing the children and some socialization work. Only a small fraction of home time is "time off", particularly for women (SADOULET and DE JANVRY, 1995). In our survey, home time is not fully covered owing to its daunting complexity. As a challenging but less intensive topic, at least in the case of rural China, this field invites further research.

In Module 8, liquidity and expenditure are recorded. Since this study is by and large production-oriented, living standards were designated somewhat roughly (living expenses were requested for monthly and yearly periods). We must concede, and this is surprising, that data in this questionnaire module are of poor quality.
Table 2: Composition of household revenue

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Agricultural income</strong></td>
<td></td>
</tr>
<tr>
<td>Farming revenue</td>
<td></td>
</tr>
<tr>
<td>Income from agricultural sidelines</td>
<td></td>
</tr>
<tr>
<td><strong>Household business income</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Wage income</strong></td>
<td></td>
</tr>
<tr>
<td>Wage income from local enterprises, organizations and other local undertakings</td>
<td></td>
</tr>
<tr>
<td>Wage income from urban off-farm activities</td>
<td></td>
</tr>
<tr>
<td><strong>Property income</strong></td>
<td></td>
</tr>
<tr>
<td>Land rent</td>
<td></td>
</tr>
<tr>
<td>Machinery, buildings and other assets renting income</td>
<td></td>
</tr>
<tr>
<td>Interest Income</td>
<td></td>
</tr>
<tr>
<td><strong>Transfer income</strong></td>
<td></td>
</tr>
<tr>
<td>Government subsidy</td>
<td></td>
</tr>
<tr>
<td>Tax exemption</td>
<td></td>
</tr>
<tr>
<td>Remittance from household members</td>
<td></td>
</tr>
<tr>
<td><strong>Self-employment</strong></td>
<td></td>
</tr>
</tbody>
</table>

Source: Own presentation.

4.6 Unit issues

We did not allow respondents to choose units by themselves; we had them pre-fixed. However, this approach is adventurous because the respondents may always round off the figure when converting. And different places might use different measurement units, especially in rural China, where the economic and institutional circumstances are quite region-specific. Fixed units can only be applied if survey regions are targeted and pre-tests are fully prepared.

The biggest concern regarding the use of units in this study is how to term the interest rate. Compared with the internationally-calculated interest rates in *percentage (annual interest rate)*, *fen* and *li* are widely used in rural China, with the former denoting monthly interest per 100 *yuan*, and the latter being one-tenth of the former. In our experience, though some rural households responded well to the interest rate, there were still some households who were not aware of the precise interest rate charged. For the sake of consistency, neither annual interest rate nor the rural Chinese system was applied in the credit section. Instead, the precise monthly interest expenditure per 100 *yuan* was employed. Though conversion was involved, it is easier to delegate this job to the enumerators, rather than leaving the problem in the field.
5 SAMPLING

The research region in this study focuses on the North China Plain (see Figure 3) with five counties selected, each comprising two townships, four villages and 68 households in a dichotomously equal way (Table 3). Neither random sampling nor selective sampling was achieved due to the failure of attaining data to calculate strata information. A three-stage sampling technique was instead applied.

First, the officially reported coverage of micro credit for rural households and the acreage of arable land per capita were used as criteria to generate township strata dichotomously. Stratification can be fulfilled in terms of economic magnitude, geographic characteristics, or even ethnicity. Townships (xiang zhen) in rural China, the entities between counties and villages, have been the basic rural administrative units since the early 1990s, though they are still politically affiliated to counties. Additionally, because the RCC County Union (RCCU) at the county level is only nominally superior to township RCCs, we targeted the township as the stratum unit.

Secondly, villages were selected in terms of distance (from the municipal townships) and road infrastructure. Third, because it was difficult to attain the names of all households in the villages, a random walk was applied to sample households. We believe that the derived results are applicable to a wider regional spectrum.

Although it is evident that multi-topic surveys have many merits, they may suffer from sampling problems. There are other data users and a range of research topics, each of which demand a different design. Several criteria had to be balanced to reach a final agreement. Didactically, we tried to adhere to the principle 'keep it simple', cited from DEATON (1997: 18) who excerpts from CASLEY and LURY (1981). The problems created by complicated design often 'offset the theoretical benefits conferred'. Therefore we suggest, when applying a more advanced sampling technique, that one makes sure to exhaust all relevant information in order to compute weights, probabilities and inflation factors. Otherwise, the non-sampling errors are more problematic.

### Table 3: Sampling profile

<table>
<thead>
<tr>
<th>Province</th>
<th>County</th>
<th>Township</th>
<th>Village</th>
<th>Household</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hebei</td>
<td>Quzhou</td>
<td>2</td>
<td>4</td>
<td>67</td>
</tr>
<tr>
<td>Henan</td>
<td>Kaifeng</td>
<td>2</td>
<td>4</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Yanjin</td>
<td>2</td>
<td>4</td>
<td>66</td>
</tr>
<tr>
<td>Shandong</td>
<td>Liangshan</td>
<td>2</td>
<td>4</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Huiming</td>
<td>2</td>
<td>4</td>
<td>68</td>
</tr>
</tbody>
</table>

6 SURVEY IMPLEMENTATION

Two pre-tests were carried out, one in the suburbs of Beijing and the other in the sampling regions to adapt the questionnaire to local circumstances. Thereafter, the questionnaire was updated and calibrated. Before the field survey, most hired enumerators were given three days of training. One lesson we learned is that the training should consist of two dimensions, the questionnaire itself and the required skills (or the art) to elicit questions during the field survey itself. It would have been better if the enumerators were offered a chance to rehearse the questionnaire in the target regions. The combination of training during the pre-test and questionnaire calibration would help unearth many problems. However, higher costs prevented us from doing this.

The final survey was conducted in April, 2005, when rural households were less occupied by farming activities. The survey was carried out in only two weeks because the enumerators were full-time students and were unavailable for long-term surveys. One experience from this survey is that the relationship between the quality of the collected data and the duration of the field survey is represented by an inverted U-shape. As the duration of a survey increases, enumerators are more familiar with questioning and are more likely to improve the quality of survey data. Nevertheless, drawn-out surveys tire enumerators, resulting in increased odds of making errors.

Although it varies from person to person, the average survey time was 90 minutes and the figure decreases with the progress of survey. All survey-related expenses amounted to approximately €4,400, with a unit cost €13 per household questionnaire. Compared with the average costs of the World Bank LSMS of around $150 to $250 per respondent, and the similar survey

5 Most were Master’s students from China Agricultural University in Beijing.
carried out in Poland of around €39 per questionnaire, this figure is remarkably low, especially when the multi-topic nature of the questionnaire is taken into consideration (DEATON, 1997: 40; PETRICK, 2001: 15). The low payments to the enumerators because of our small budget, however, discouraged them from fully working on questioning. The unit cost of the questionnaire, though important and necessary, should not be the determining criteria for evaluation. Hiring students (or local people) in developing countries is meanwhile a chance, though not obligatory to researchers, to disseminate knowledge and information, and to empower local junior researchers.

Table 4: Survey cost components

<table>
<thead>
<tr>
<th>Items</th>
<th>Outlays in Euro</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traveling</td>
<td>9,700</td>
</tr>
<tr>
<td>Food</td>
<td>4,627</td>
</tr>
<tr>
<td>Accommodation</td>
<td>7,314</td>
</tr>
<tr>
<td>Gifts</td>
<td>7,199</td>
</tr>
<tr>
<td>Communication</td>
<td>1,310</td>
</tr>
<tr>
<td>Others</td>
<td>717</td>
</tr>
<tr>
<td>Copies</td>
<td>580</td>
</tr>
<tr>
<td>Salary for interviewers</td>
<td>13,440</td>
</tr>
<tr>
<td>Total</td>
<td>44,887</td>
</tr>
</tbody>
</table>


7 DATA ENTRY AND DATA MANAGEMENT

7.1 Overview of data entry

Figure 4 displays the flowchart of data entry for a questionnaire-based survey. Programming the data entry software at the time of questionnaire design is strongly recommended because many unexpected errors, which are less likely to be corrected after the field survey, can be detected and remedied at this stage. The data entry software EpiData was used in this study. Compared with inputting data directly in statistical tools, EpiData has overwhelming advantages in error detection and control. The CHECK file, which is configured after programming the data file based on the questionnaire, provides several excellent tools: a) basic entry validation during the data entry process (type or range to a number of specified values), b) value labels, c) control of the flow in the data entry process (e.g. automatic jumps from one entry field to another field or forcing an entry in a field). In addition, the user-customized layout, the compatible configuration with Stata, SPSS and SAS, and the enhanced documentation function all provide this software with a great many advantages for empirical analysis, especially for those multi-topic research projects with large sized datasets.

---

6 EpiData is a free software, released by EpiData Association in Denmark. For more information, visit <http://www.epidata.dk/index.htm>.
7.2 Categorical information and value label

Categorical variables, which carry integral values but explain their meaning with textual counterparts, were used extensively in this study. For data entry of categorical variables, it is very important to value labels in such a way that variables carry value labels automatically once they are exported to various statistical or econometric tools. But there is one problem to watch out for when using EpiData to define categorical variables. In EpiData, the field 'Boolean', which denotes logic operators with the options True or False (e.g. whether or not a household applied credit in 2004) denotes two logic categories. Unfortunately, once 'Boolean' variables are exported into Stata, they turn out to be a 'string' and do not identify the valued information. Though mapping string variables to categorical ones in Stata is in no way impossible, it is troublesome if a large amount of variables are mistakenly valued.

8 EXPERIENCE AND LESSONS

Overall, the data collection conducted in Spring 2005 was a well-organized survey and the questionnaire was moderately suitable. In addition, the cooperation between different disciplines advances our knowledge of households’ behavior in rural China. Though enjoying a handful of advantages, multi-topic surveys face the problem of diverse sampling criteria. Interests of individual topics need to be balanced and compromised. True 'inter-discipline' (or cross-discipline) calls for deep integration right from the beginning, instead of simply pooling various topics together, which should be pursued persistently for future studies (KANBUR, 2002). The cross-disciplinary study perhaps needs to demonstrate how the different disciplines are integrated with well-defined and focused hypothesis.

Another salient problem is the structure of the questionnaire. While cutting some presumably insignificant questions due to the length of the questionnaire, the price for giving up possibly enlightening information is very high. Our suggestion is that the length should not be a determining factor when judging a questionnaire. If the length is too daunting to be finished in one-time survey, the enumerators can adapt their field survey techniques to the length and the content of the questionnaire; researchers might carry out the unfinished survey several days later.
or any time they believe both sides are not tired, provided that the survey itself is not time-sensitive to multiple sessions.

No data collection can be error-free. Nevertheless, if potential traps are anticipated, errors can be reduced to a large extent, granting the database higher quality. Some additional experience and lessons are presented below, though they seem to be didactical. Hopefully they are helpful to those who will collect data in rural China or those who intend to use data collected in rural China.

- Questionnaires should be compatible with local economic and institutional terminology. For rural household surveys, the agro-structure and corresponding cropping calendar are used as the reference.
- The selection of survey time is important. It should not be at odds with the busy season of farming.
- Compile a survey manual in which the summaries and instructions of questions are well written for enumerators. Remark upon possibly ambiguous questions. Use this manual for training.
- Programming the data entry process should precede the field survey.
- Code categorical variables with great care. When data entry is programmed, value the labels of variables correctly because it benefits the data analysis afterwards.
- In field work, keep away any interested onlookers, such as other family members and neighbors; this makes it possible to administer the more sensitive portions of the questionnaire with greater privacy (GROSH et al., 2000).
- Use the best-informed respondent in the household. The heads of households, always male in rural China, are the ideal person unless some modules are particularly specific to other family members.
- Verbalize the questions in a uniform manner and use spoken (or local) language. Adapt to local economic and institutional contexts.
- The pre-test should cover various types and agro-ecological regions to foresee potential discrepancies. Further, before the full field pre-test, there should be one (or a round of) rehearsal pre-test. By doing this, the efficiency of the full-sized pre-test can be increased greatly because, in most cases, "nearly half of the problems will show up in the first 10 households interviewed in pre-test" (GROSH et al., 2000: 58).
- Fully prepare for training and have all selected enumerators participate in the pre-test. Provide effective incentives to enumerators, whether it be good payment or other trade-in-kind.
REFERENCES


HE, G. (1999): Restriction and deepening of rural financial market from aspect of rural credit behaviour in China, China Rural Economy, 10.


DISCUSSION PAPERS
OF THE LEIBNIZ INSTITUTE OF AGRICULTURAL DEVELOPMENT
IN CENTRAL AND EASTERN EUROPE (IAMO)

No. 1 Frohberg, K., Hartmann, M. (1997):
Promoting CEA agricultural exports through association agreements with the EU
– Why is it not working? –

No. 2 Frohberg, K., Hartmann, M. (1997):
Comparing measures of competitiveness: Examples for agriculture in the Central
European Associates

No. 3 Oganietz, W. R., Glauch, L. (1997):
Migration durch EU-Integration? Folgen für den ländlichen Raum

No. 4 Weingarten, P. (1997):
Agri-environmental policy in Germany – Soil and water conversation –

No. 5 Kopsidis, M. (1997):
Marktintegration und landwirtschaftliche Entwicklung: Lehren aus der Wirtschafts-
geschichte und Entwicklungsökonomie für den russischen Getreidemarkt im Trans-
formationsprozeß

No. 6 Pieniadz, A. (1997):
Der Transformationsprozeß in der polnischen Ernährungsindustrie von 1989 bis
1995

No. 7 Oganietz, W. R. (1997):
Vermindern Transferzahlungen den Konflikt zwischen Gewinnern und Verlierern
in einer sich transformierenden Volkswirtschaft?

Difficulties in the privatization and reorganization of the agricultural enterprises
in Russia

No. 9 Girgzdiene, V., Hartmann, M., Kuodys, A., Rudolph, D., Vaikutis, V.,
Restructuring the Lithuanian food industry: Problems and perspectives

Restructuring the Latvian food industry: Problems and perspectives

No. 12 Bergschmidt, A., Hartmann, M. (1998): Agricultural trade policies and trade relations in transition economies

No. 13 Elsner, K., Hartmann, M. (1998): Convergence of food consumption patterns between Eastern and Western Europe


No. 17 Shvytov, I. (1998): Agriculturally induced environmental problems in Russia


No. 20 Wildermuth, A. (1999): What kind of crop insurance for Russia?


No. 22 Jasjko, D., Hartmann, M., Miglavs, A., Wandel, J. (1999): Industrial organisation of the food industry in Latvia: Results of an expert survey in the dairy and milling branches


No. 24 Petrick, M., Ditges, C. M. (2000): Risk in agriculture as impediment to rural lending – The case of North-western Kazakhstan
No. 25  **POGANIETZ, W. R.** (2000):  
Russian agri-food sector: 16 months after the breakdown of the monetary system

No. 26  **WEBER, G., WAHL, O., MEINLSCHMIDT, E.** (2000):  
Auswirkungen einer EU-Osterweiterung im Bereich der Agrarpolitik auf den EU-Haushalt  
(Steht nicht mehr zur Verfügung – Aktualisierte Version DP 42)

No. 27  **WAHL, O., WEBER, G., FROHBERG, K.** (2000):  
Documentation of the Central and Eastern European Countries Agricultural Simulation Model (CEEC-ASIM Version 1.0)

No. 28  **PETrick, M.** (2000):  
Land reform in Moldova: How viable are emerging peasant farms? An assessment referring to a recent World Bank study

No. 29  **WEINGARTEN, P.** (2000):  

No. 30  **BROSIG, S.** (2000):  
A model of household type specific food demand behaviour in Hungary

No. 31  **UVAROVSKY, V., VOIGT, P.** (2000):  
Russia’s agriculture: Eight years in transition – Convergence or divergence of regional efficiency

No. 32  **SCHULZE, E., TILLACK, P., GERASIN, S.** (2001):  
Eigentumsverhältnisse, Rentabilität und Schulden landwirtschaftlicher Großbetriebe im Gebiet Wolgograd

No. 33  **KIELYTE, J.** (2001):  
Strukturwandel im baltischen Lebensmittelhandel

No. 34  **ШУЛЬЦЕ, Э., ТИЛЛАК, П., ГЕРАСИН, С.** (2001):  
Отношения собственности, рентабельность и долги крупных сельскохозяйственных предприятий в Волгogradской области

No. 35  **FROHBERG, K., HARTMANN, M.** (2002):  
Konsequenzen der Integration im Agrar- und Ernährungssektor zwischen Beitrittsländern und EU-15

No. 36  **PETrick, M.** (2001):  
Documentation of the Poland farm survey 2000

No. 37  **PETrick, M., SPYCHALSKI, G., ŚWITLYK, M., TYRAN, E.** (2001):  
Poland's agriculture: Serious competitor or Europe's Poorhouse? Survey results on farm performance in selected Polish voivodships and a comparison with German farms


No. 40 ШУЛЬЦЕ, Э., ТИЛЛАК, П., ПАТЛАССОВ, О. (2002): Факторы, влияющие на прибыль и рентабельность крупных сельскохозяйственных предприятий в Омской области в России


No. 43 PETRICK, M. (2002): Farm investment, credit rationing, and public credit policy in Poland – A microeconometric analysis –


No. 46 PETRICK, M., LATRUFFE, L. (2003): Credit access and borrowing costs in Poland’s agricultural credit market: A hedonic pricing approach


No. 49 ЛИССИТСА, А., БАБИЧЕВА, Т. (2003): Теоретические основы анализа продуктивности и эффективности сельскохозяйственных предприятий

No. 50 ЛИССИТСА, А., БАБИЧЕВА, Т. (2003): Анализ Оболочки Данных (DEA) – Современная методика определения эффективности производства

No. 51 ЛИССИТСА, А., ОДЕННИГ, М., БАБИЧЕВА, Т. (2003): 10 лет экономических преобразований в сельском хозяйстве Украины – Анализ эффективности и продуктивности предприятий
Russian Agrarsektor im Aufschwung? Eine Analyse der technischen und Skalen-Effizienz der Agrarunternehmen

No. 53 Valentino, V. (2003):
Social capital, transition in agriculture, and economic organisation: A theoretical perspective

No. 54 Borkowski, A. (2003):
Machtverteilung im Ministerrat nach dem Vertrag von Nizza und den Konvents- vorschlägen in einer erweiterten Europäischen Union

Cost of compliance with the acquis communautaire in the Hungarian dairy sector

Quantitative analysis of the impacts of Croatia’s agricultural trade policy on the agri-food sector

No. 57 Бокушева, Р., Хайдельбах, О. (2004):
Актуальные аспекты страхования в сельском хозяйстве

No. 58 Derlitzki, R., Schulze, E. (2004):
Georg Max Ludwig Derlitzki (1889-1958)

No. 59 Vöneki, E. (2004):
Zur Bewertung des Ungarischen SAPARD-Programms unter besonderer Berücksichtigung der Investitionen im Milchsektor

No. 59 Чимпоееш, Д., Шульце, Э. (2004):
Основные экономические проблемы сельского хозяйства Молдовы

No. 61 Baum, S., Weingarten, P. (2004):
Interregionale Disparitäten und Entwicklung ländlicher Räume als regionalpolitische Herausforderung für die neuen EU-Mitgliedstaaten

Can econometric analysis make (agricultural) economics a hard science? Critical remarks and implications for economic methodology

No. 63 Sauer, J. (2004):
Rural water suppliers and efficiency – Empirical evidence from East and West Germany

No. 64 Petrick, M., Balmann, A. (2004):
Beiträge des 2. Doktorandenworkshops zur Agrarentwicklung in Mittel- und Osteuropa 2004

No. 65 Bojnic, S., Hartmann, M. (2004):
Agricultural and food trade in Central and Eastern Europe: The case of Slovenian intra-industry trade
No. 66  GLITSCH, K., EERITS, A. (2004):  
Der slowakische Markt für Milch und Milchprodukte – Vom Beginn der Transformation bis zum EU-Beitritt

No. 67  FISCHER, C. (2004):  
Assessing Kosovo’s horticultural potential – The market for fruit and vegetables on the balkans

Competitiveness of milk and wine production and processing in Albania

No. 69  ШТАНГЕ, Г., ЛИССИТСА, А. (2004):  
Аграрный сектор России на подъеме?! Анализ технической эффективности аграрных предприятий

No. 70  SAUER, J. (2004):  
Die Ökonomie der (Ländlichen) Wasserversorgung

No. 71  HAPPE, K., BALMANN, A., KELLERMANN, K. (2004):  
The Agricultural Policy Simulator (Agripolis) – An agent-based model to study structural change in agriculture (Version 1.0)

No. 72  BAUM, S., TRAPP, CH., WEINGARTEN, P. (2004):  
Typology of rural areas in the Central and Eastern European EU new member states

No. 73  PETRICK, M. (2004):  
Governing structural change and externalities in agriculture: Toward a normative institutional economics of rural development

No. 74  RODIONOVA, O., SCHULZE, E., UERKOV, E., KARPOVA, G. (2004):  
Zur Besteuerung von Agrarholdings in Russland

No. 75  HEIDELBACH, O., BOKUSHEVA, R., KUSSAYINOV, T. (2004):  
Which type of crop insurance for Kazakhstan? – Empirical results

No. 76  BOKUSHEVA, R. (2004):  
Crop insurance in transition: A qualitative and quantitative assessment of insurance products

No. 77  RAMANOVICH, M., LAJTOS, I. (2004):  
Milchproduktion und -verarbeitung in Weißrussland: Eine Analyse der Wettbewerbsfähigkeit

No. 78  LUKA, O., LEV KOVYCH, I. (2004):  
Intra-industry trade in agricultural and food products: The case of Ukraine

No. 79  EINAX, CH., LISSITS A, A., P ARKHOMENKO, S. (2005):  
Getreideproduktion in der Ukraine – Eine komparative Analyse von Produktionskosten

No. 80  ИВАХНЕНКО, О., ЛИССИТСА, А. (2005):  
Информационно-консультационная служба в аграрно-промышленном комплексе России на примере Омской области
<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>Röth, A., Lissitsa, A. (2005):</td>
<td>Аграрный сектор Восточной Германии в переходном периоде – Исходная ситуация, развитие и основные проблемы</td>
</tr>
<tr>
<td>85</td>
<td>Melnychuk, V., Parhomenko, C., Lissitsa, A. (2005):</td>
<td>Процесс формирования рынка сельскохозяйственных земель в Украине</td>
</tr>
<tr>
<td>91</td>
<td>Чимпоеш, Д., Шульце, Э. (2006):</td>
<td>Экономическое состояние сельскохозяйственных предприятий Республики Молдова</td>
</tr>
</tbody>
</table>
No. 95 GRAMZOW, A. (2006):
Endogenous initiatives as a chance to improve rural livelihood? Results of a case study in Báltów, South-eastern Poland

No. 96 DUHFUES, T., BUCHENRIEDER, G., FISCHER, I. (2006):
Social capital and rural development: Literature review and current state of the art

Social capital among agricultural producers in the Czech Republic: Its impact on economic performance

No. 98 BOKUSHEVA, R., BUCHENRIEDER, G. (2006):
Contributions to the 4th Young Scientists Workshop on agricultural development in Central and Eastern Europe – YSW-2006

Zur Wettbewerbsfähigkeit der weißrussischen Milchwirtschaft: Eine Anwendung des Porterschen Diamanten

No. 100 GRAMZOW, A. (2006):
Doświadczenie oddolnych inicjatyw rozwoju regionalnego oraz perspektywy dla programu leader+ w regionie Doliny Strugu w Polsce

Partnerstwo Lokalne jako inkubator rozwoju terenów wiejskich: Przypadek Debrzna, północno-zachodnia Polska

No. 102 XIANGPING, J., BUCHENRIEDER, G. (2007):
Documentation of a multi-topic questionnaire-based survey on sustainable resource use in rural China

Die Discussion Papers sind erhältlich beim Leibniz-Institut für Agrarentwicklung in Mittel- und Osteuropa (IAMO) oder im Internet unter http://www.iamo.de.
The Discussion Papers can be ordered from the Leibniz Institute of Agricultural Development in Central and Eastern Europe (IAMO). Use our download facility at http://www.iamo.de.