
Collecting Community Histories to Study the Determinants of Demographic Behaviour: A Survey in Burkina Faso

Author(s): Bruno Schoumaker, Hubert Bonayi Dabire, Bilampoa Gnoumou-Thiombiano

Source: *Population (English Edition, 2002-)*, Vol. 61, No. 1/2 (Jan. - Apr., 2006), pp. 81-107

Published by: [Institut National d'Études Démographiques](#)

Stable URL: <http://www.jstor.org/stable/30042113>

Accessed: 19/09/2011 12:02

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Institut National d'Études Démographiques is collaborating with JSTOR to digitize, preserve and extend access to *Population (English Edition, 2002-)*.

Collecting Community Histories to Study the Determinants of Demographic Behaviour

A Survey in Burkina Faso

Bruno SCHOUMAKER*, Hubert BONAYI DABIRE**
and Bilampoa GNOUMOU-THIOMBIANO***

Multi-level event history analyses depend upon data that are rarely available. Alongside individual trajectories, indicators of change in the social, economic and demographic environment must also be reconstructed. In most cases, the context is described via individual characteristics and a small number of indicators gleaned from available administrative sources. To obtain a more accurate measure of changes in the context of individual life courses, a survey was conducted in Burkina Faso to collect community event history data that could be factored into subsequent analyses. In this article, Bruno SCHOUMAKER, Hubert BONAYI DABIRE and Bilampoa GNOUMOU-THIOMBIANO present a detailed description of this survey process, demonstrating its feasibility and suggesting a promising range of applications.

Over the past thirty years, many authors have stressed the importance of taking the characteristics of the local environment, or context⁽¹⁾, into account when explaining demographic behaviour. The general idea is that demographic behaviour is not only influenced by individual-level and household factors, but also depends upon the social and economic environment, sanitary conditions, etc. The demographic literature contains

* Institut de démographie, Université catholique de Louvain.

** Institut supérieur des sciences de la population, Université de Ouagadougou.

*** Department of Demography, Université de Montréal.

Translated by Julia Key

⁽¹⁾ The local context corresponds in this case to an individual's day-to-day "living space", where most social interaction takes place and where the availability, or otherwise, of services and infrastructures has the most impact. It usually corresponds to the village or neighbourhood (Schoumaker, 2001).

numerous examples of potential ways in which the local environment may influence fertility (Schoumaker, 1999; Casterline, 1987; Freedman, 1974), mortality (Frankenberg, 1995; DaVanzo, 1985) or, of greater interest here, migration (Lucas, 1997; Hugo, 1985; Bilsborrow, 1984).

In parallel with this growing interest in the influence of the local context on demographic behaviour, the World Fertility Survey (WFS), from the 1970s onwards, and Demographic and Health Surveys (DHS), from the mid-1980s, started to collect community data in localities where survey clusters for surveys of individuals were located. In total, around one third of the surveys carried out under these programmes involved the collection of community data, resulting in numerous analyses of the determinants of fertility and mortality, incorporating explanatory variables at different levels (Schoumaker, 1999; Tuo, 2002)⁽²⁾.

However, the community data gathered in these surveys have several weaknesses. Often considered to be of poor quality⁽³⁾ and primarily concerning health and family planning services, they include practically no retrospective data. While individual-level data on fertility and mortality collected in the WFS and DHS are based on the full birth histories of women, the community data generally only concern the characteristics of the localities at the time of the survey. This considerably limits the scope for carrying out longitudinal analyses incorporating both individual-level and community variables. However, since the mid-1990s, a number of operations to collect retrospective community data have been conducted as part of demographic surveys on developing countries (for Indonesia and Malaysia, see Frankenberg, 2000; for Nepal, see Axinn et al., 1997), paving the way for multi-level longitudinal analyses of demographic behaviour.

Yet even when retrospective community data are collected in these surveys, they only relate to the sample localities selected for the survey of individuals. This means that for individuals who once lived outside the sample area (migrants), there are no suitable community data to explain their behaviour prior to arrival in the survey locality. For analyses concerning mortality or fertility, this problem can to some extent be overcome by restricting analyses to those periods after the last migration for example (Schoumaker, 2001). However, this lack of retrospective data on non-sample areas has more serious consequences for the study of the determinants of migration. If one is interested in the contextual characteristics that prompted individuals to leave their former place of residence, some knowledge of the features of this place in the past is essential. Yet such data are not available in a standard retrospective community-level survey.

⁽²⁾ The World Bank surveys of living standards also frequently collect both demographic data on individuals and contextual data suitable for such analyses.

⁽³⁾ According to several authors, the collection of contextual data in the DHS surveys often appears to be a secondary task with "inadequate supervision of fieldwork and data quality" (Mensch et al., 1996, p.61).

One solution to this problem is to conduct a retrospective survey covering not only the localities sampled in a retrospective survey of individuals, but also the former places of residence of the persons surveyed. This potentially involves a much more extensive data collection operation, to the extent that some authors have ruled out an approach of this type (Findley, 1982). Such a survey was nevertheless carried out recently (2002) in Burkina Faso, as part of a research project on migration and urban integration, involving researchers from the Department of Demography at the Université de Montréal, the Institut supérieur des sciences de la population (ISSP, former UERD) at the University of Ouagadougou and the Center for Study and Research on Population and Development (CERPOD) in Bamako. This survey, designed to be linked to data from an event history survey on migration and urban integration carried out in 2000 (Poirier et al., 2001), is, to our knowledge, the first data collection operation of its type. It provides opportunities to examine new issues regarding the influence of environmental factors on migration, as well as on fertility, nuptiality, school enrolment, employment, etc.

This article outlines the main objectives of this survey, explores a number of methodological and practical issues arising at the design and planning stages, and presents some of the work based on the data. It first examines the reasons for gathering retrospective community data on the current and former places of residence of the individuals surveyed, and looks at various ways of collecting this data. It goes on to discuss the process of selecting sample localities and the practical organization of the survey, before turning to the methods used to collect community data (questionnaire, calendar) and a number of methodological issues relating to questionnaire design. The final part contains a brief outline of work based on these data and of some ongoing projects.

I. Context and objectives of the survey

In the year 2000, ISSP, CERPOD and the Department of Demography at the Université de Montréal carried out an event history survey of migration and urban integration in Burkina Faso, entitled EMIUB (Poirier et al., 2001). Although apparently similar to surveys of urban integration in other African countries (Antoine et al., 1999), this survey is national in scope, covering not only the inhabitants of the country's major towns and cities, but also those in rural areas. A total of 8,644 individuals aged 15-64 were questioned about their migration and occupational history since the age of six, their marital history and, for women, their full birth history. Although clearly focusing on the issues of migration and urban integration, the survey encompasses many other aspects of respondents' lives.

Despite the wealth of personal and household data gathered in this survey, it does not include detailed information on place of residence, other than basic location and whether or not the area was urban or rural⁽⁴⁾. Yet some knowledge of the characteristics of the towns or villages in which the individuals live is essential to explain spatial mobility (Lucas, 1997; Findley, 1987; Hugo, 1985; Bilsborrow, 1984; Oberai and Bilsborrow, 1984; Gardner, 1981). Such information is also important for the study of mortality, fertility, school enrolment or access to employment, to give just a few examples.

Without reviewing in detail the reasons behind migration, we will look briefly at a few examples of community characteristics likely to influence such movement. One of the factors most often cited, in both developed and developing countries, is the availability of paid work in or around the locality in question. The localities offering significant opportunities for paid employment are thought to both retain their population and to attract migrants (Hugo, 1985). Another key element, and a central issue for migration policies designed to reduce rural exodus, is the impact of infrastructures and public services in rural localities (Beauchemin and Schoumaker, 2005; Lucas, 2000; United Nations, 1995). The availability of land, the presence of natural resources (goldmines, etc.), and the existence of conflict are other factors liable to influence migration⁽⁵⁾.

Combining individual-level and community data to study migration

There are two ways of approaching the collection of community data to study migration (Findley, 1982): (1) investigating the locality characteristics that influence the decision to emigrate, and (2) examining the community characteristics that affect the choice of destination. In principle, the factors influencing these two aspects of the migration decision will be similar. Which of these two questions (or indeed both of them) one chooses to consider will nevertheless affect the organization of data collection. In the first case, the data must concern the sending localities, whereas in the second they relate to the potential destinations. Since the main objective of the survey in Burkina Faso was to determine the influence of community-level factors on the risk of emigrating, this article will focus on the

⁽⁴⁾ The name and location (*département* and *province*) of each locality in which individuals had resided for at least three consecutive months since the age of six were recorded. Using village census data, each locality was classified, for each year, as either urban or rural, according to the size of its population, with urban areas defined as those with more than 10,000 inhabitants (Beauchemin, et al., 2002).

⁽⁵⁾ Contextual data are also useful in explaining other demographic and social phenomena. They are vital to assess the impact of the presence of schools on school enrolment, the relationship between the level of local development and access to paid employment, the impact of health services on child mortality or fertility for example.

influence on emigration⁽⁶⁾ of the characteristics of sending localities. The final part will look briefly at how data from the survey may be used to shed light on factors influencing the choice of destination.

To study the factors influencing the risk of emigrating from a locality, as is the case here, it must therefore be possible to establish a link between individuals liable to emigrate and the characteristics of their place of residence. However, this proves problematic when it comes to collecting data on both individuals and their localities of origin, a difficulty inherent in attempts to combine individual-level and community-level information to study spatial mobility (Findley, 1982; Lucas, 2000). One can collect data on the places selected for the survey, although, by definition, the emigrants will no longer be there, and information on such individuals cannot be obtained directly. Alternatively, one can question individuals living in the locality about their former places of residence (migration history), but, as has already been pointed out, information on such places cannot in this case be gathered directly in the localities where the survey of individuals is conducted.

In the first case, one solution is to gather data on individuals by questioning households not only about members still present, but also about those who have emigrated in recent years (Findley, 1982; Lucas, 2000; Bilsborrow et al., 1987). This information, combined with community data gathered in the sample localities, serves to measure the influence of individual-level and community-level characteristics on the risks of emigrating. Such an approach has a practical advantage – community data is gathered in the localities selected for the survey of individuals – though it does have a number of drawbacks. First, the emigration of entire households (including one-person households) goes unrecorded, so the extent of migration may be considerably underestimated. Second, given that information on emigrants is obtained from other individuals, it will generally not be very comprehensive, and may well be unreliable. Finally, such data will relate only to recent migration, and cannot therefore be used for detailed longitudinal analyses.

The second of these two approaches was the one adopted for this project. Individuals were questioned about their migration history, complete or partial. Occupational and family life histories were also recorded where possible, thus permitting much more in-depth explanatory analyses than with cross-sectional data. At the individual level, this method makes it possible, for example, to assess the impact of changes in individuals' lives on migration risk. Similarly, the effect on emigration of changes to the local context can be examined. In this case, a community-level survey restricted to the sample localities selected for the survey of individuals no longer suffices, and it becomes necessary to collect retrospective data on the localities where the individuals once resided.

⁽⁶⁾ Insofar as one is concerned with how the characteristics of a sending locality affect emigration, these data also serve to study the influence of such characteristics on the risk of migrating abroad (among individuals who have returned).

Choosing a source of retrospective community data

How might community data be collected? There are several possible approaches, but let us first review the conditions to be satisfied. To carry out event history analyses of the determinants of migration, we need to know, for both past and present, not only the characteristics of the individuals themselves (marital status, employment status, etc.), but also those of the localities where they live. These data must therefore (1) be *retrospective*, and (2) ideally, cover *all localities in which these individuals once lived*. Broadly, there are three approaches to gathering this information. The first is to collect community data directly through the survey of individuals. Alternatively, secondary data gathered for other studies can be used. The third possibility, and the one described in this article, is to organize a survey on the localities identified in the individual life event histories.

The first approach provides basic information on the localities in which individuals once lived (Axinn et al., 1997). For example, if one is interested in the existence of primary schools, each individual might be asked questions such as: "When you left [name of locality], was there a primary school there?" This information could be supplemented by more retrospective data, for example: "When you settled in [name of locality], was there a primary school there?" Depending on the answers to these two questions, one could then ask when the school was first opened. The main advantage of such a method is that it does not entail visiting these past places of residence, thereby cutting project costs. Yet despite its apparent simplicity, it does have serious flaws. Firstly, it results in a considerably longer individual questionnaire, since in addition to their own event histories, individuals must recall information on the history of the localities where they once lived. Furthermore, if 50 individuals resided in the same locality over the same period, the same information will be gathered from 50 different sources. This needlessly adds to survey costs, and the overall quality of data collected in this way will also be lower than that obtained by surveying informants in the localities themselves (Axinn et al., 1997)⁽⁷⁾. Finally, there are certain subjects that cannot be discussed, or only with difficulty, in a survey of individuals. The availability of land and the existence of conflict are two such examples. This approach would therefore seem inappropriate for gathering retrospective community data of sufficient detail and quality, although a few simple questions on the local context could well be incorporated into the individual questionnaires. Consequently, the method was rejected for this particular project.

⁽⁷⁾ For example, it is unlikely that an isolated individual could give the precise date when a primary school was opened in the locality where he/she lived 25 years ago, particularly if he/she was very young at the time. It is also highly probable that the information will differ from one respondent to another, thus posing the problem of choosing one single value for that locality.

The second approach is to gather information from existing sources (Frankenberg, 2000; Findley, 1982). This has a cost advantage, since it relies on data already collected. These data must nevertheless fulfil the conditions outlined above: they should be retrospective, and, ideally, cover all localities in which the individuals surveyed once resided. This severely limits the number of potential sources, although some data do meet these criteria. In Burkina Faso, the Ministère de l'enseignement de base et de l'alphabétisation (MEBA, Ministry of Basic Education and Literacy) holds an exhaustive database listing the country's primary schools, with the dates they first opened and their location. Information on when dams were constructed or boreholes drilled is also available in exhaustive databases. These data were brought together, but it was not possible to obtain retrospective information on health centres and secondary schools, to give just two examples. In other words, although some relevant data are accessible, in the case of Burkina Faso – and most probably in many other African countries – the quantities are limited. A further potential weakness of such data is that they are not always of high quality (Frankenberg, 2000; Axinn et al., 1997). Finally, they are clearly not appropriate for information on conflict, poor harvests, etc.

The third approach is to collect the desired information by organizing a retrospective survey of the localities. This involves visiting the localities identified in the migration histories and questioning informants about subjects that may be relevant to the phenomena studied. This method offers substantial flexibility (Findley, 1982), and although limited by the retrospective nature of data collection, it provides a means to explore topics for which administrative sources or the survey of individuals are not appropriate (land-related issues, conflicts, etc.). The data obtained from such a survey may also be of higher quality than that from administrative sources. The major potential weaknesses of this approach are the degree of organization involved and the high cost. However, experience from the Burkina Faso survey showed that an operation of this type is feasible within a relatively short time-frame and at reasonable expense. The following sections deal with the main methodological and organizational issues arising in the design and planning of this survey.

II. Selecting the sample localities and organizing the survey

At the time of the 1996 census, there were approximately 8,500 localities in Burkina Faso, for a population of just over 10 million (INSD, 2000). One objective of the migration section of the event history survey was to record the migration history of each of the 8,644 individuals aged 15-64 included in the survey, identifying all the localities in which respondents had spent more than three consecutive months since the age of six. Prior to the survey, it had been difficult to estimate precisely the total number of localities for which community data would be required. Once the survey was completed, the findings were clear: over the course of their

lives, the 8,644 individuals surveyed had lived in 1,703 localities within Burkina Faso, one fifth of the total number of localities in the country⁽⁸⁾.

Figure 1 illustrates the cumulative number of localities identified in the migration event histories, moving back in time from the date of the survey. For the year 2000, only the 101 localities chosen for the survey of individuals (i.e. those in which the survey clusters were located) are listed. Adding together all the localities in which time was spent over the 10 years prior to the survey gives a total of almost 1,000, while during the 20 year period preceding the survey almost 1,400 localities had been lived in for three months or more by at least one person. In total, almost 1,700 different localities had been resided in at least once since 1950.

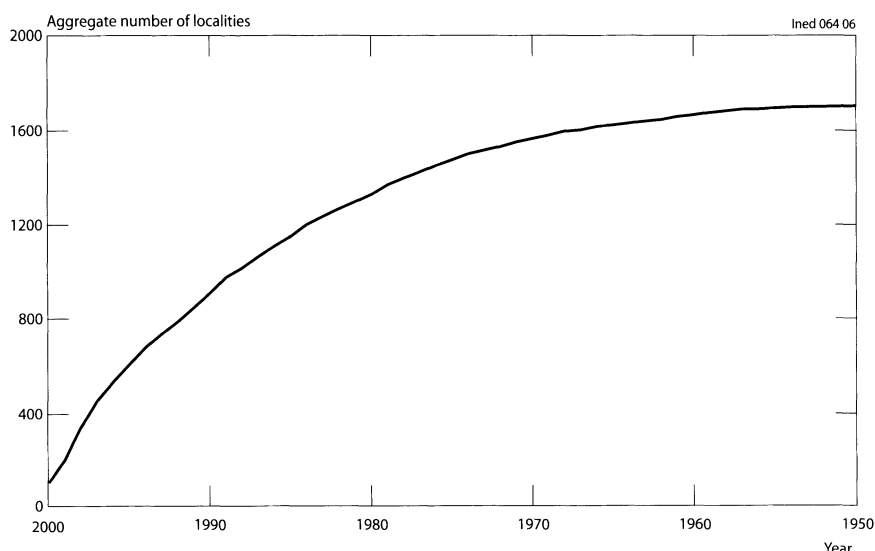


Figure 1. – Aggregate number of localities lived in by at least one individual in the event history survey, Burkina Faso, 1950-2000

Interpretation: The x-axis indicates the years, moving back in time from the survey date (April 2000) to 1950. The y-axis indicates, for each year, the aggregate number of localities in which at least one survey respondent reported having lived for at least three months. At the time of the survey, the number of localities is equal to the number of sample localities in the survey of individuals. Going back in time, the localities where individuals lived before settling in their current place of residence are progressively included in the total number of localities.

Source: EMIUB, 2000.

⁽⁸⁾ This figure excludes localities of residence abroad, in which case only the name of the country was noted. Of the 26,626 periods of residence recorded for the 8,644 respondents, 2,863 were outside the country, accounting for around 11% of the total number of periods. There were more such periods for men (14%) than for women (8%). The majority of this international migration was to Côte d'Ivoire (76% of foreign stays), with the rest shared between Ghana (7%), Mali (6%), Niger (5%) and other African countries. Spells of residence outside Africa accounted for no more than 2% of all periods spent abroad.

Carrying out a survey of 1,700 localities is an enormous operation, and, in view of the limited resources available, the idea of selecting a subgroup was soon suggested. Yet such an approach inevitably results in missing data for certain periods of residence identified in the event history survey. What criteria should be used in selecting these localities? Intuitively, it would seem logical to focus the survey on localities where several spells of residence were recorded: given that many of the localities are only mentioned once or twice in the event histories, it would have been impracticable to survey each one. This more focused approach, which was the one we adopted, has the advantage of *minimizing the proportion of missing data for a given number of localities surveyed*. Other criteria were nevertheless considered. It would have been possible, for example, to select all the localities of residence of respondents over the 10 years prior to the study. This would have provided us with a complete set of data for a given period. However, as Figure 1 shows, doing so would have meant including more than 1,000 localities for those 10 years alone, still a sizeable number for a limited period of coverage.

The approach taken was therefore to select the subgroup localities according to the number of periods of residence recorded for each of them. Figure 2 illustrates the advantage of such an approach, showing the relationship between, on the y-axis, the proportion of localities surveyed – out of a total of 1,703 areas identified – and, on the x-axis, the percentage of periods of residence for which community data are missing (the total being 23,763 spells of residence within the country for the 8,644 respondents). If all localities are surveyed, this proportion is zero of course. Taking only those localities in which at least two periods of residence were recorded halves the number of localities to be surveyed (from 1,703 to 890), yet the corresponding proportion of missing community data is only 3.5% of all time periods (bottom curve), and 5.5% when time spent in Ouagadougou and Bobo Dioulasso is excluded (top curve). Restricting the survey further to areas in which three or more periods were recorded reduces the number of localities to 600, with no more than 6% missing data for all spells of residence (10% excluding Ouagadougou and Bobo Dioulasso). The number of villages to be surveyed may, of course, be reduced further, but the percentage of missing data then rises more rapidly as the number of villages decreases. Cutting the sample size from 900 to 600 villages, for example, incurs a “cost” of approximately 4% missing data, whereas halving it from 600 to 300 results in a further 7% data loss. Selecting only villages with at least three periods of residence was partly an arbitrary decision, and influenced essentially by the budget available and the amount of missing data deemed reasonable. Data for localities with fewer than three periods were then imputed using a multiple random imputation process⁽⁹⁾.

⁽⁹⁾ For analyses coupling individual-level and contextual data, the missing contextual data are imputed using a random *hot-deck* imputation method. The imputation procedure is based on two variables: the province where the village is located (a total of 45 provinces) and the locality size (below 5,000 or between 5,000 and 10,000 inhabitants — there are no missing data for towns). The localities covered by the contextual survey are classified according to these two variables, and the missing values for a village are replaced by the values of a randomly selected village among those having the same values for these two variables, i.e. a village of the same size and in the same province. A multiple imputation, in which the results of several models estimated with different imputed values are combined, is used to correct the standard deviations of the regression coefficients, which are under-estimated in the simple imputation procedures (Allison, 2002).

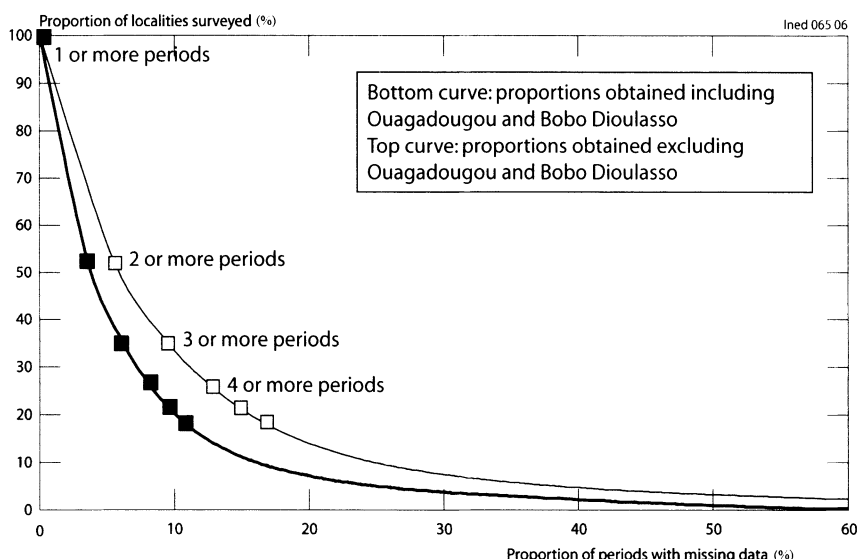


Figure 2. – Relation between the proportion of localities surveyed and the proportion of missing community data among total periods of residence in Burkina Faso, 1950-2000

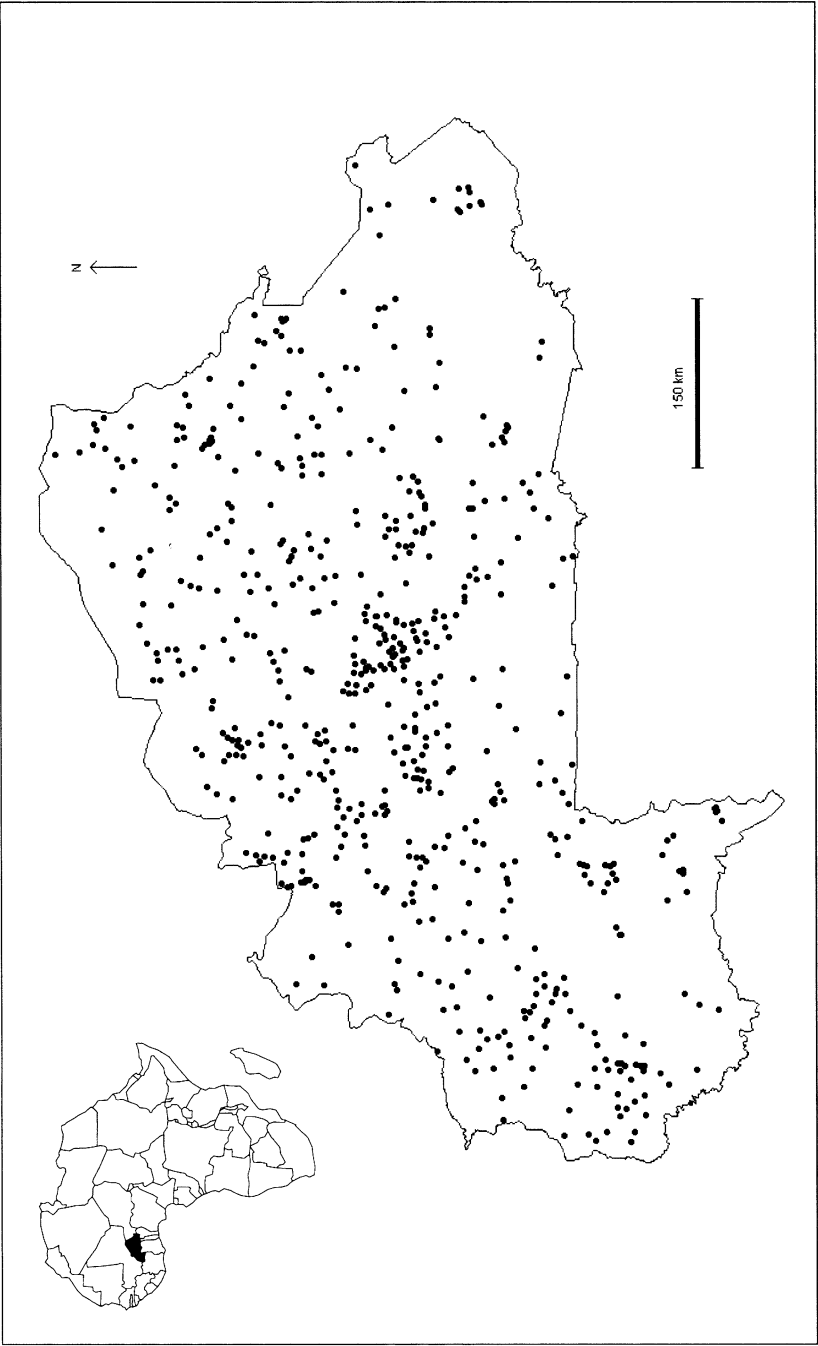
Interpretation: This graph compares different locality selection scenarios for the community-level survey, each scenario being distinguished by the minimum number of periods of residence required. In the first scenario (1 or more periods), all the localities identified in the survey of individuals are surveyed. In this case, 100% of localities are included in the survey and the community data are available for all periods of residence reported in the survey of individuals (0% missing data). In the second scenario (2 or more periods), the localities where a single period of residence were reported are excluded from the community-level survey. Slightly more than 50% of localities are included and the proportion of periods of residence for which community data is missing is below 10%. The lower curve indicates the percentage of missing data including all periods of residence in Ouagadougou and Bobo Dioulasso, the two main cities, for which it is assumed that all data is available. The upper curve indicates the proportion of missing data excluding the periods of residence in Ouagadougou and Bobo Dioulasso.

Source: EMIUB, 2000.

A total of 600 areas were thus selected for survey. Of these, 599 were successfully surveyed, with just one proving impossible to locate. They included 48 towns (defined as localities with more than 10,000 inhabitants), and the provincial capitals of the country's 45 provinces were all included in the survey (with the exception of Ouagadougou and Bobo Dioulasso).

Survey organization and data collection

One of the challenges posed by this survey was that of organizing the data collection. While experience from other retrospective community-level surveys provided a benchmark for questionnaire design, this did not prove the case when it came to organizing the survey on the



Map . – Location of the 599 community-level survey areas in Burkina Faso

ground. The large number of areas to be surveyed (600) and the fact that they were widely dispersed throughout the country (see Map) represented a sizeable difficulty, and required that interviewers be given a considerable degree of autonomy.

The approach taken was to divide the 600 localities among 13 experienced interviewers (seven men and six women), capable of working independently in difficult conditions⁽¹⁰⁾. Each interviewer was given a list of localities to survey according to their linguistic knowledge. They were provided with maps on which the localities had been marked, and a moped enabling them to travel between villages⁽¹¹⁾. The route taken by each interviewer had been established in advance, although they were given a certain degree of autonomy to ensure that the work was organized effectively. A further problem was that of interviewer supervision, as it was impossible to monitor their work continuously. It was decided that the three researchers in charge of the survey should travel throughout the country to meet each of the interviewers in turn. This proved an effective solution, given that the interviewers were experienced and reliable.

As with most surveys of localities, the respondents were interviewed in groups. In contrast to surveys of individuals, it is widely recognized that interviewing several people together improves data quality (Frankenberg, 2000; Axinn et al., 1997). Some respondents will be better informed than others with regard to certain points in the questionnaire, and discussion among them can enhance the quality of the information obtained. However, it is neither useful nor desirable to multiply the number of respondents unnecessarily. Preliminary tests established that a large group size (ten or more) hampered the interview process. For the purposes of this survey, around six respondents was the recommended number, with a strict minimum of three per locality.

The first stage of the interview process generally involved meeting the village representative⁽¹²⁾ or the headman, who undertook to find suitable respondents. Although the interviewers had been briefed about the type of respondent to be selected, the choice was largely left to the inhabitants themselves, and particularly to the village representative. In general, it was recommended that the village representative and/or the headman be present at the interview, and that a wide range of profiles in terms of age, sex and ethnic

⁽¹⁰⁾ Two major criteria applied when selecting interviewers were their experience and their knowledge of languages, since each was required to cover around 50 localities in several linguistic areas. Like most interviewers in Burkina Faso and many other developing countries, they were not professionals. Some were university students, while others were looking for stable employment. But they all had experience as interviewers in several major data collection operations. Some, for example, had taken part in the DHS survey in 1999, others in the event history survey of migration in 2000.

⁽¹¹⁾ The interviewers covered, on average, around 40 kilometres a day. This fieldwork stage lasted approximately 40 days altogether, with an average of 1.2 interviews per interviewer and per day. The average duration of each interview was a little over two hours.

⁽¹²⁾ The village representative reports to the government, i.e. he represents the village in administrative matters. The headman is the traditional village authority.

group be represented. In a large majority of localities, the village representatives and headmen took part in the survey (at least one of the two was present in 84% of localities), and the breakdown by age and ethnic group was generally satisfactory⁽¹³⁾. However, it proved more difficult to involve women in this survey. Only 12% of villages had at least one female respondent.

In general, finding respondents posed very few problems. The survey was carried out in the dry season when there is less agricultural work, and this clearly increased the availability of respondents. In most cases, the local authorities had been informed by the prefect of the *département* that a survey was to be carried out in their locality, and this also expedited the process. The idea of tracing the history of the localities, while intriguing certain participants, generally received a very favourable response, and a warm and friendly atmosphere prevailed in most villages. Although some sections of the questionnaire inevitably provoked discussion regarding the “exact” date of particular events, on the whole there was little notable dissension among respondents. The only questions that sometimes posed problems were those concerning conflicts within the village or with other villages, with certain headmen categorically refusing to discuss these issues.

III. Data collection tools: questionnaire and calendar

Two tools were used to collect the data: a contextual questionnaire and a village calendar. All information was recorded on the questionnaire, the calendar being purely a means of improving data quality⁽¹⁴⁾.

The questionnaire

The use of closed questions is essential for collecting data that is comparable over a large number of localities (Bilsborrow, 1984). Consequently, the questionnaire used in this survey was highly structured, with a large majority of closed questions. Most of the information gathered is retrospective and dates back as far as 1960⁽¹⁵⁾. The questionnaire content and structure were based on a number of other retrospective community-level surveys, among them the Family Life Surveys in Indonesia (Frankenberg, 2000; Frankenberg et al., 1995) and a survey carried out in Nepal (Axinn et al., 1997). The questionnaire comprises nine major sections in total, with an additional section for basic information about the village and the respondents (Table 1). The chosen

⁽¹³⁾ In each locality, the characteristics of the most active participants, with a maximum of six per locality, were recorded at the end of the interview.

⁽¹⁴⁾ Audio recordings of the interviews were also made in around 30 localities, and are currently being processed.

⁽¹⁵⁾ The year 1960 was chosen for two reasons: (1) the vast majority of migratory events in the survey of individuals took place after 1960, and (2) it was a landmark year in the country's history (independence), and thus an easy reference point for respondents.

topics encompassed most of the potential determinants of migration identified in the literature that could feasibly be covered in a retrospective survey⁽¹⁶⁾. It also covers many topics of relevance for explanatory analyses of other demographic and social variables (fertility, school enrolment, etc.).

TABLE 1.— SECTIONS OF THE COMMUNITY-LEVEL SURVEY QUESTIONNAIRE

Section 1	<i>Administrative history</i>
Section 2	<i>Physical characteristics and potential of the village</i>
Section 3	<i>Educational infrastructure</i> A. Primary schools B. Secondary schools
Section 4	<i>Health services and pharmacies</i> A. Healthcare facilities B. Pharmacies
Section 5	<i>Settlement, conflicts, endemics, drought</i> A. History of settlement B. History of conflict in the village C. Endemics D. Harvests and droughts
Section 6	<i>Economic activity, agriculture and land issues</i> A. Main activities B. Agricultural activity C. Agricultural technology and practices D. Quality and accessibility of land
Section 7	<i>Paid employment</i>
Section 8	<i>Development projects, groups and associations</i> A. Development projects B. Village groups C. Migrant associations
Section 9	<i>Public facilities, consumer goods, accessibility</i> A. Public and leisure facilities B. Markets C. Consumer goods D. Accessibility and transport links

It was decided that a single questionnaire should be used for all the localities surveyed, be they villages with a few hundred inhabitants or small or medium-sized towns⁽¹⁷⁾. There were several reasons for this, the

⁽¹⁶⁾ Information on migratory networks, potentially important for explaining migration, was not gathered directly, although two sections of the questionnaire cover this question more indirectly. Section 8C concerns the existence of associations of persons from the locality (i.e. migrants associations), and section 5A looks at the history of settlement, including information on the origin of immigrants and the destinations of emigrants from the locality.

⁽¹⁷⁾ According to data from the 1996 census, the localities surveyed had populations ranging from 42 to 72,490 inhabitants, with a median size of 2,179 and an interquartile range of between 1,112 and 4,305 inhabitants. Around 8% of the localities surveyed had 10,000 inhabitants or more, and 12 had over 20,000.

main one being that these localities form a continuum of urbanization, and there was no clear criterion for deciding which localities should be given a rural questionnaire and which an urban one. This question was further complicated by fact that the data collected were retrospective, and that certain areas which were urban at the time of the survey had been quite clearly rural some years before. For most of the questionnaire sections (infrastructure, services, etc.), using the same questionnaire for both towns and small villages did not pose a particular problem⁽¹⁸⁾. Only questions relating to farming were problematic in the most urbanized areas. Retaining this section did, however, make it possible to collect information on urban agriculture, which remains an important activity in Burkina Faso.

Collecting retrospective community data entails certain difficulties that are not encountered in cross-sectional data collection. In view of the rarity of this type of procedure, we will outline here a number of these problems, along with the solutions adopted. Although by no means exhaustive, this section uses concrete examples to illustrate some of the difficulties and constraints encountered.

*Availability of services in time and space:
the example of primary schools*

The questionnaire section concerning the availability of primary schools illustrates, with a simple example, some of the particular problems involved in collecting retrospective data on infrastructure and services (Figure 3). The aim of this section was to establish, for each year from 1960 onwards, (1) the number of primary schools in the locality, and (2) the distance to the nearest primary school at any point in time. The type of school (public, private) was also recorded.

The basic principle of this section, and indeed most of the sections relating to infrastructure and services (healthcare facilities, roads and transport infrastructure, etc.), was to begin with the situation in the locality at the time of the survey. If one or more schools were available (Q300), establishing the dates when they were first opened made it possible to go back in time and determine the date when the oldest school was built (Q302). Then, if the oldest school in the locality had been opened after 1960, information was collected on the schools attended prior to its opening (Q315b-Q320). Respondents were asked: "Before the oldest school opened, where was the school attended by most of the village children?"⁽¹⁹⁾ and the distance to this school was then recorded. Depending on

⁽¹⁸⁾ The main difficulty was the large number of schools, healthcare facilities, etc. in some medium-sized towns. To lighten the workload where this was the case, we provided interviewers with information from administrative sources (lists of schools, etc.), which they were required to check and complete if necessary. Administrative data available in the towns was also sometimes used.

⁽¹⁹⁾ This approach, based on collecting information about the school attended by most of the children, was preferred to one based on the school that was simply the nearest. The questionnaire also included a series of questions designed to check whether the school most attended was also the nearest.

Q300. Combien d'écoles primaires y a-t-il actuellement dans le village ? (si 0, → Q307)										
RELEVEZ LES INFORMATIONS SUR TOUTES LES ECOLES PRIMAIRES DANS LE VILLAGE EN COMMENCANT PAR LA PLUS ANCIENNE.										
N°	Q301. Quel est le nom de l'école ?	Q302. En quelle année a-t-elle été ouverte ?	Q303. De quel type d'école s'agit-il ?				Q304. Combien de périodes d'interruption prolongée (>1 an) y a-t-il eu dans son fonctionnement depuis sa création (après 1960) ? (SI NE SAIT PAS, NOTEZ 99)			
1		a. Avant 1960 b. Après 1960 :	Classique publique..... 1				N°	Classique publique..... 1		
			Classique privée..... 2					Q305. Début interruption		
			Satellite..... 3					Q306. Fin interruption		
			Franco-arabe / medersa.. 4					1		
2		a. Avant 1960 b. Après 1960 :	Classique publique..... 1				N°	Classique publique..... 1		
			Classique privée..... 2					Q305. Début interruption		
			Satellite..... 3					Q306. Fin interruption		
			Franco-arabe / medersa.. 4					1		
3		a. Avant 1960 b. Après 1960 :	Classique publique..... 1				N°	Classique publique..... 1		
			Classique privée..... 2					Q305. Début interruption		
			Satellite..... 3					Q306. Fin interruption		
			Franco-arabe / medersa.. 4					1		
Q307. Combien d'écoles primaires y a-t-il eu dans le village depuis 1960 et qui n'existent plus actuellement ?										
-S'IL N'Y A JAMAIS EU D'ECOLE DANS LE VILLAGE (Q300=0 ET Q307=0), PASSEZ A Q315.										
-S'IL Y A ACTUELLEMENT AU MOINS UNE ECOLE (Q300>0) ET QU'IL N'Y EN A PAS EU DANS LE PASSÉ (Q307=0), PASSEZ A Q315b SI LA DATE D'OUVERTURE DE L'ECOLE LA PLUS ANCIENNE DU VILLAGE EST APRES 1960. SI LA DATE D'OUVERTURE EST AVANT 1960, PASSEZ DIRECTEMENT A LA SECTION 3.B.										
RELEVEZ LES INFORMATIONS SUR LES ECOLES PRIMAIRES AYANT EXISTE DEPUIS 1960 EN COMMENCANT PAR LA PLUS ANCIENNE.										
N°	Q308. Quel est le nom de l'école ?	Q309. En quelle année a-t-elle été ouverte ?	Q310. En quelle année a-t-elle été fermée ?	Q311. De quel type d'école s'agit-il ?				Q312. Combien de périodes d'interruption prolongée (>1 an) y a-t-il eu dans son fonctionnement depuis sa création jusqu'à sa fermeture (après 1960) ? (SI NE SAIT PAS, NOTEZ 99)		
1	a. Avant 1960 b. Après 1960 :			Classique publique..... 1				N°	Classique publique..... 1	
				Classique privée..... 2					Q313. Début interruption	
				Satellite..... 3					Q314. Fin interruption	
				Franco-arabe / medersa.. 4					1	

Figure 3. – Section of the questionnaire on primary schools

POSEZ LA QUESTION Q315 UNIQUEMENT SI AUCUNE ECOLE N'EXISTE ACTUELLEMENT DANS LE VILLAGE. S'IL Y A UNE ECOLE DANS LE VILLAGE, PASSEZ A Q315b. -SI L'ANNEE D'OUVERTURE DE L'ECOLE (Q318) EST AVANT 1960, TERMINEZ DE COLLECTER LES INFORMATIONS SUR CETTE ECOLE ET PASSEZ ENSUITE A Q321.									
N°	Q315. Les enfants du villages vont-ils à l'école primaire ?	Q316. Où se trouve l'école primaire la plus fréquentée par les enfants du village ?	Q317. A quelle distance se trouve-t-elle ?	Q318. En quelle année a-t-elle été ouverte ?	Q319. De quel type d'école s'agit-il ?	Q320. S'agit-il de l'école primaire la plus proche ?			
1	Oui 1 Non 2 (→PRENEZ LES INFORMATIONS SUR L'ECOLE LA PLUS PROCHE)	Nom école : Localité : Dép : Prov :	[] [] [] km	a. Avant 1960 b. Après 1960 [] [] []	Classique publique 1 Classique privée 2 Satellite 3 Franco-arabe / medersa ... 4	Oui 1 Non 2			
	Q315b. Avant l'ouverture de l'école la plus ancienne fréquentée, les enfants du villages allaient-ils à l'école primaire ?	Q316b. Où se trouvait l'école la plus fréquentée par les enfants du village ?	Q317. A. quelle distance se trouvait-elle ?	Q318. En quelle année a-t-elle été ouverte ?	Q319. De quel type d'école s'agit-il ?	Q320. S'agissait-il de l'école primaire la plus proche ?			
2	Oui 1 Non 2 (→PRENEZ LES INFORMATIONS SUR L'ECOLE LA PLUS PROCHE)	Nom école : Localité : Dép : Prov :	[] [] [] km	a. Avant 1960 b. Après 1960 [] [] []	Classique publique 1 Classique privée 2 Satellite 3 Franco-arabe / medersa ... 4	Oui 1 Non 2			
3	Oui 1 Non 2 (→PRENEZ LES INFORMATIONS SUR L'ECOLE LA PLUS PROCHE)	Nom école : Localité : Dép : Prov :	[] [] [] km	a. Avant 1960 b. Après 1960 [] [] []	Classique publique 1 Classique privée 2 Satellite 3 Franco-arabe / medersa ... 4	Oui 1 Non 2			
POSEZ LES QUESTIONS Q321 A Q327 POUR TOUTES LES ECOLES EN DEHORS DU VILLAGE (LIGNES CI-DESSUS) QUI NE SONT PAS OU N'ETAIENT PAS LES PLUS PROCHES DU VILLAGE (Q320=2). SI TOUTES LES ECOLES CITEES ETAIENT LES PLUS PROCHE DU VILLAGE (Q320=1), PASSEZ A LA SECTION 3.B.									
N°	Q321. NOTEZ LE NUMERO DE LA LIGNE CONCERNEE	Q322. Où se trouve/trouvait l'école la plus proche ?	Q323. A quelle distance se trouve-t-elle ?	Q325. En quelle année a-t-elle été ouverte ?	Q326. De quel type d'école s'agit-il ?	Q327. Pour quelle(s) raisons les enfants du village ne vont / n'allaient-ils pas à cette école ? (EN CLAIR)			
1	N° de ligne []	Nom école : Localité : Dép : Prov :	[] [] [] km	a. Avant 1960 b. Après 1960 [] [] []	Classique publique 1 Classique privée 2 Satellite 3 Franco-arabe / medersa ... 4				

Figure 3 (cont'd). – Section of the questionnaire on primary schools

when the school was established, the same question was repeated until the school most attended predated 1960⁽²⁰⁾. Similarly, if there were no primary schools in the locality at the time of the survey, the school that most children attended was identified (Q315). Again, depending on when it was opened, questions were then asked about the schools attended before this, until the school most attended predated 1960.

To obtain comprehensive information on the availability of schools throughout the survey period, two further series of questions were asked: one set concerning times when these schools were temporarily closed (Q304-Q306), and a second set on schools that were open in the past but no longer existed at the time of the survey (Q308-Q314). In the end, these two elements proved to be of little relevance. Very few schools (below 1%) had been temporarily shut down, and the information provided is not vital to most of the analyses. The question relating to schools that had previously existed but had closed by the time of the survey could also have been omitted, given the small number of such cases recorded. However, this recommendation would not have been appropriate for all services. Many pharmacies, for example, had existed in the past but were no longer open at the time of the survey. Omitting these questions would have significantly underestimated the availability of pharmacies in previous years.

Several sections (secondary schools, healthcare facilities, pharmacies, markets, roads, etc.) follow the same principle as that on primary schools, although a number of complications arise in some of them. One such problem concerns the change in status of certain services. For example, a secondary school that is currently a *lycée* (i.e. embracing the whole of secondary education) may well have been opened as a *collège* (covering only the first few years of secondary schooling). To establish whether there was a *lycée* in the locality for each year of the survey period, an additional question on the year the school changed from a *collège* to a *lycée* must be incorporated. A similar problem arises in the case of health services, although with further complications. The questionnaire identifies five different types of healthcare facility at the time of the survey: dispensaries (clinics), health and community centres (HCC), medical centres (MC), medical centres with operating facilities (MCO) and hospitals. The problem here is that not only might these services have undergone a change in status – from HCC to MC, for example – but the very way in which they are classified has also varied over time: what is now an MC was not known as such in the past. To obtain reliable information on these changes, it would have been necessary to carry out a survey of the healthcare facilities themselves (Frankenberg, 2000; Axinn et al., 1997), in addition to surveying the villagers. In view of budgetary constraints, however, this approach was rejected, and in the end, only the status of the particular facility at the time of the survey was recorded.

⁽²⁰⁾ Axinn et al. (1997) adopted a similar technique, albeit limited to a maximum of three schools. The advantage of the approach taken here is that it provides comprehensive data on the period after 1960.

Occurrence of events in the locality: the example of poor harvests

The survey was also designed to gather information about the occurrence of particular events. For example, the questionnaire contained a section to identify years when harvests were extremely poor. Such years, associated largely with drought, are often seen as a factor that prompts individuals to leave their village, either temporarily or permanently (Henry et al., 2004; Findley, 1994). In this case, the interviewers asked respondents to identify every year, since 1960, when the harvest had been very poor, i.e. well below what farmers considered as normal. Questions were then asked about the causes of each of these poor harvests (lack of rain, insects, etc.) and about their consequences (sale of property, migration, etc.).

Two basic problems were encountered in collecting these data. Firstly, the notion of what constitutes a “very poor harvest” may vary, not only from one village to the next but also over time. Figure 4 illustrates this point. It shows, for each year, the number of localities where respondents declared an extremely poor harvest. The two major droughts of the mid-1970s and 1980s appear quite clearly. More unexpected, however, is the significant increase in the number of localities reporting a very poor harvest at the end of the 1990s. This is partly because the harvests in the few years leading up to the survey were indeed of poor quality. Yet it may well also be due, in part, to the fact that the perception of what constitutes a bad harvest depends on when it occurred, with recent years probably more strongly imprinted in respondents’ memories⁽²¹⁾.

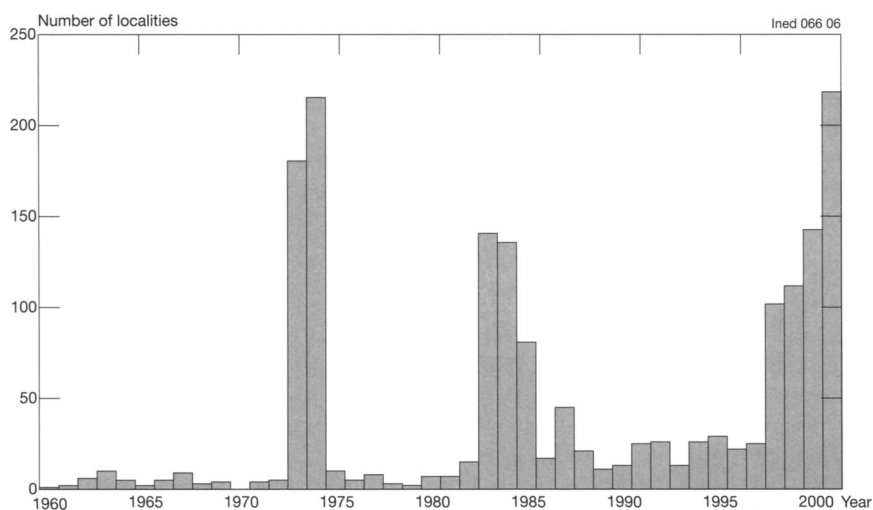


Figure 4.— Number of localities where respondents reported a very poor harvest, by year, Burkina Faso, 1960–2001

Source: Community-level survey, 2002.

⁽²¹⁾ Certain respondents may also have been tempted to describe the recent situation as particularly unfavourable, possibly in the hope of receiving aid.

A further – and classic – problem encountered in collecting these data was that of accurately identifying years when the harvest was poor. As will be seen below, a calendar was devised to improve the quality of chronological information gathered in the survey. This did not, of course, solve all the dating problems and some of the information remains approximate. However, an analysis of the poor harvest data indicated that, despite the data collection difficulties, they are of satisfactory quality⁽²²⁾.

Gradual introduction of certain practices: the example of the plough

The previous examples all relate to particular events: the opening of a school or a health centre, the occurrence of a poor harvest, etc. Yet the survey was also concerned with phenomena that develop more gradually, and here the issue of precise dates proves far less clear-cut. This is the case for certain agricultural practices, such as the use of ploughs or of chemical fertilizers for example. Should such retrospective information simply be overlooked? If not, what method should be used to gather data on the development of these practices?

One relatively simple approach involves asking a question such as: “In what year were ploughs first used by families in the village?” Even if the date given is approximate, it will give a good idea of how long the practice has existed in the locality. However, this approach is rather crude, and provides no indication of whether the practice is widespread. It was also suggested that respondents might be asked to estimate, for different periods of time, the proportion of families using ploughs, with questions such as: “What proportion of families in the village currently use a plough?”, “What proportion of families used a plough 10 years ago?”, etc. This approach was evaluated at the pre-testing stage, but proved unworkable. In an attempt to form at least some idea of how the practice developed in the village, a combination of three questions was finally decided upon: (1) “If ploughs are used in the village, how long has this been the case?”, (2) “Is this practice currently very widespread (>75% of households), fairly widespread (>30%), or not very widespread (<30%)?”, and finally (3) “If the practice is very widespread, since what year has this been the case?” Of course, no exact dates can be obtained with such questions, and this method is clearly a rough one. It does, however, for any given year, place each locality in one of three categories: those where ploughs are not used, those in which they are rarely used or only fairly widespread, and those where they are extensively used. The same type of information was gathered for various agricultural practices (fertilizers,

⁽²²⁾ The data quality analysis was based on comparing the dates given in the survey as years of poor harvest with rainfall data from meteorological stations for the period 1960-98 (New et al., 2000). This involved calculating the probability of a poor harvest in a given year and a given locality as a function of pluviometric variables for the area in which the locality was located. The models show a high degree of correlation between these variables and the probability of an extremely poor harvest being declared by respondents (see Schoumaker et al., 2004).

tractors, irrigation, etc.). These different elements may, for example, be combined to produce an index of agricultural development that varies over time (Beauchemin et al., 2004).

The village calendar

Given the importance of dates in this survey, and in view of the difficulties experienced by respondents in recalling the precise dates of certain events, a “village calendar” was devised. This was based upon the one used for a similar survey in Nepal (Axinn et al., 1997), and also resembles the calendars used in event history surveys (Freedman et al., 1988; Antoine et al., 1987). At the start of the interview, respondents were asked to recall key events in the life of the village (death of the headman, election of the representative, drilling of the first borehole, etc.). The dates of these events could generally be established with accuracy, and they were then written on the calendar, on the line corresponding to the year they occurred. A number of important national events had also been marked in. This series of milestones, with their correct dates marked prior to the survey itself, then served as points of reference to pinpoint more accurately the dates of events recalled thereafter. Although the calendar did prove useful, some respondents found it hard to recollect key events spontaneously, and it was often difficult to obtain information on more than four or five of them at the start of the interview.

IV. Results and related research projects

This section contains selected examples of work, published or scheduled, based on the combined analysis of retrospective individual-level and community data obtained in these surveys. As noted earlier, the main purpose of collecting these retrospective community data was to study the relationship between the characteristics of a locality and the risks of emigrating. The first example examined here deals with this same question. These data are also useful in other areas of research, some examples of which are suggested below.

The impact of local development on migration to cities

In a recent study, C. Beauchemin and B. Schoumaker (2005) investigated the impact of local development on migration to Burkina Faso’s two main cities (Ouagadougou and Bobo Dioulasso). This study starts out from the fact that many developing countries have introduced policies to reduce migration to their cities (United Nations, 1998), often founded upon the belief that improving the quality of life and economic conditions in rural areas reduces migration. Although this notion appears reasonable at first sight,

there are few empirical studies to support it, due to a lack of suitable data. The combination of life event histories and retrospective community data constitutes a unique source of information for examining the issue at national level.

Life event history analysis models are used notably to assess – while controlling for individual-level variables – the impact of the characteristics of rural localities on the probability of migrating to one of the two cities mentioned above. The main community variables selected for this study concern the availability of infrastructure (roads, electricity, telephone, boreholes), public facilities (schools, healthcare), business services (markets, cinemas, bars) and income-generating activity (large firms, wage employment, cash crop farming, etc.). Overall, this research shows that individuals living in more developed areas are more likely to migrate to the city, clearly contradicting the hypothesis underlying migration policies. More specifically, local facilities and well-developed infrastructures either have no effect on migration, or actually drive inhabitants away. Only certain income-generating activities, such as the presence of a large firm, seem to curb emigration.

The effect of development in sending localities on return migration

In the previous example, the central issue was that of assessing whether the best-equipped rural localities tend to retain their residents. A related question, which may be addressed with these data, is to investigate whether individuals leaving relatively developed areas are more likely to return there than those leaving villages that are less well-equipped. This issue was recently explored in the case of international migration (Beauchemin et al., 2004), though it could equally be examined for internal migration.

The study in question focuses on men aged 15-64 who had migrated abroad and who returned to Burkina Faso during the period 1970-99, and of whom around 75% returned to their locality of origin. Data from event history and community-level surveys offer new ways of analysing the factors that influence migrants' decisions to return to their place of origin or to settle elsewhere, since the community-level survey provides information on the characteristics of the sending locality, even when the emigrant is no longer living there. The extent to which its various features influence migrants' choice of destination when they return can thus be assessed.

One of the study findings was that community-level factors have less influence on the choice of return destination than individual-level factors (education, occupation, etc.). The results also show, echoing those of the study on rural out-migration mentioned above, that migrants from the most developed areas are between two and three times less likely to return there, being instead strongly inclined to head for the country's main towns and cities.

Other ongoing projects

Various studies based on a combination of individual-level and community data are in progress or at the planning stage. Some of these projects, on migration and other demographic variables, are mentioned below to illustrate the range of possible analytical approaches.

One current project focuses on the choice of destination in internal migration. While data were collected for the above-mentioned studies to assess the influence of sending locality characteristics on emigration, they might equally be used, with a number of assumptions, to examine the effect of community characteristics on the choice of destination. This research is based on the use of random utility models, and community data are used here to produce a sample of possible destinations (and their characteristics), including not only those chosen by migrants but also a group of localities not selected. One of the preliminary findings is that the presence of income-generating activities in secondary towns serves to draw in migrants, whereas this does not appear to be the case in rural areas.

Other projects concern the influence of community characteristics, among them healthcare facilities, on fertility and mortality (Schoumaker, 2005; Dos Santos-Beauchamp and Henry, 2004). Although such issues are more classically analysed using DHS data, the studies presented here reduce the problem posed by the lack of community data relating to the migrants' past. These data, can, for example, be used to assess the extent to which the assumptions made in the absence of retrospective data on localities outside the survey areas influence the analysis results.

Finally, community data may also be used independently of individual-level data, either for descriptive purposes or more explanatory analyses⁽²³⁾, to examine factors influencing the development of certain services and infrastructures for example. Preliminary analyses have shown that the building of primary schools depends upon factors such as accessibility, population size and the proximity of other schools (Schoumaker et al., 2004). An ongoing research project on the impact of migrants' associations on their sending localities suggests that those whose emigrants have set up associations in towns and cities are more likely to obtain a health centre than villages elsewhere (Beauchemin and Schoumaker, 2006).

Conclusion

This article provides a broad overview of the rationale behind the retrospective community-level survey carried out recently in Burkina

⁽²³⁾ Since the survey areas do not constitute a representative sample of all localities nationwide (large areas being overrepresented in the group), the sample must be corrected accordingly. This involves calculating post-stratification weightings such that the marginal distributions of the sample according to certain characteristics (province, administrative status, population size) correspond exactly to those of the country as a whole.

Faso, along with its characteristics and potential uses. The survey was in several respects the first of its kind, and offers a number of useful insights.

In particular, it confirms that *it is possible to collect high-quality retrospective community information* from respondents. Survey data quality is closely assessed as and when the data are analysed, so the process is not yet complete. However, initial assessments have shown the data to be reliable, despite certain typical problems (tendency towards 'round' dates, etc.). The survey also indicates that *community information on a large number of localities can be collected at reasonable cost and within a reasonable time-frame*. We were able to obtain retrospective community data on almost 600 localities identified from the migration histories collected in a survey of individuals, with the fieldwork taking less than six weeks to complete. Over and above these two main conclusions, the survey offers a number of lessons for similar operations in the future, be it in terms of the type of information that can be collected, the questionnaire design, or the practical organization of work in the field.

As we have briefly discussed, there are many possible uses for these data. This survey breaks new ground by linking individual-level event histories and community data. Although the main objective here was to use both individual-level and community-level factors to explain spatial mobility, other aspects of behaviour (fertility, mortality, marriage, school enrolment, employment, etc.) could equally well be analysed in this way. In short, these data provide a means to perform multi-level event history analyses for which suitable data are very rarely available (Axinn et al., 1997; Courgeau, 1999). By identifying the effects of community variables (roads, schools, electricity, etc.) on phenomena such as migration, mortality, fertility or schooling, these data can assist in identifying and evaluating appropriate policies. Over and above their purely scientific interest, the data also have potential applications in policy making.

Acknowledgements: This survey, a collaborative venture by the Department of Demography at the Université de Montréal, the former UERD at the University of Ouagadougou and the CERPOD in Bamako, was funded by the Mellon Foundation and the Canadian International Development Agency. The authors gratefully acknowledge the support of Jean Poirier, Victor Piché and Gaël Lejeune from the Department of Demography at the University of Montreal, whose idea first prompted the survey, and who were involved in designing the questionnaire.

We are grateful to Cris Beauchemin and Sabine Henry, as well as two anonymous readers, for their comments on the first draft of this article. We would also like to thank our interviewers for their hard work and enthusiasm. Finally, we extend our thanks to the thousands of willing and patient respondents who were kind enough to take part in this survey, and, in some cases, to welcome the interviewers into their homes.

REFERENCES

- ALLISON P., 2002, *Missing Data*, Thousand Oaks, Sage.
- ANTOINE P., C. BONVALET, D. COURGEAU, F. DUREAU, É. LELIÈVRE (eds.), 1999, *Biographies d'enquêtes. Bilan de 14 collectes*, Collection Méthodes et savoirs, Paris, Ined/IRD.
- ANTOINE P., X. BRY, P. DIOUF, 1987, "La fiche 'AGEVEN' : un outil pour la collecte des données rétrospectives", *Techniques d'enquête*, 13(2), pp. 173-181.
- AXINN W., J. BARBER, D. GHIMIRE, 1997, "The neighborhood history calendar: A data collection method designed for dynamic multilevel modeling", *Sociological Methodology*, 27, pp. 355-392.
- BARBER J., S. MURPHY, W. AXINN, 2000, "Discrete time multilevel survival analysis", *Sociological Methodology*, 30, pp. 201-235.
- BEAUCHEMIN C., E. BEAUCHEMIN, G. LE JEUNE, 2002, "Tabvilles Burkina Faso : rapport de présentation", *EMIUB Survey. Technical analysis document*, 2002-01, Montreal, Department of Demography, University of Montreal.
- BEAUCHEMIN C., B. SCHOUMAKER, 2005, "Migration to cities in Burkina Faso. Does the level of development in sending areas matter?", *World Development*, (33)7, pp. 1129-1152.
- BEAUCHEMIN C., B. SCHOUMAKER, 2006, "Migrant associations as actors in local development? A national event history analysis in Burkina Faso", Paper presented at the annual meeting of the *Population Association of America*, Los Angeles.
- BEAUCHEMIN C., B. SCHOUMAKER, S. HENRY, 2004, "Côte d'Ivoire—Burkina Faso (1970-2000) : une étude rétrospective des déterminants individuels et contextuels du retour", AIDELF congress on international migration, Budapest.
- BILSBORROW R., 1984, "The need for and design of a community-level questionnaire", in R. Bilborrow, A. Oberai and G. Standing (eds.), *Migration Surveys in Low Income Countries: Guidelines for Survey and Questionnaire Design*, London & Sidney, Croom Helm, pp. 407-486.
- BILSBORROW R., T. MCDEVITT, S. KOSSOUDJI, R. FULLER, 1987, "The impact of origin community characteristics on rural-urban out-migration in a developing country", *Demography*, 24(2), pp. 191-210.
- CASTERLINE J., 1987, "The collection and analysis of community data", in J. Cleland and C. Scott (eds.), *The World Fertility Survey: An Assessment*, London, Oxford University Press, pp. 882-905.
- COURGEAU D., 1999, "De l'intérêt des analyses multi-niveaux pour l'explication en démographie", in D. Tabutin, C. Gourbin, G. Masuy-Stroobant and B. Schoumaker (eds.), *Théories, paradigmes et courants explicatifs en démographie*, Louvain-la-Neuve/Paris, Academia/L'Harmattan, pp. 93-116.
- DA VANZO J., 1985, "Measuring community variables for household health and demographic surveys in developing countries", *Rand Working Paper*, P-7099, Santa Monica, Rand Corporation.
- DOS SANTOS-BEAUCHAMP S., S. Henry, 2004, "L'influence des variables communautaires sur la survie des enfants en Afrique sub-saharienne. Une analyse biographique multi-niveaux au Burkina Faso", Abstract of a paper presented at the 25th International Population Conference, Tours 2005, IUSSP.
- FINDLEY S., 1982, "Methods of linking community-level variables with migration survey data", in United Nations (ed.), *National Migration Surveys. X. Guidelines for Analyses*, New York, Economic and Social Commission for Asia and the Pacific, United Nations, pp. 276-300.
- FINDLEY S., 1987, *Rural Development and Migration. A Study of Family Choices in the Philippines*, Boulder, Westview Press.
- FINDLEY S., 1994, "Does drought increase migration? A study of migration from rural Mali during the 1983-1985 drought", *International Migration Review*, 28 (3), pp. 539-553.
- FRANKENBERG E., 1995, "The impact of health infrastructure on infant mortality in Indonesia: A fixed effects approach to evaluating health services", *Health Transitions Review*, 5 (2), p. 143-162.
- FRANKENBERG E., 2000, "Community and price data", in M. Grosh and P. Glewwe (eds.), *Designing Household Survey Questionnaires for Developing Countries. Lessons from 15 Years of the Living Standard Measurement Study*, Washington D.C., World Bank, pp. 315-338.

- FRANKENBERG E., L. KAROLY, P. GERTLER, C. PETERSON, D. WESLEY, 1995, *The 1993 Indonesian Family Life Survey: Overview and Field Report*, DRU-1195/1, Santa Monica, Rand.
- FREEDMAN D., A. THORNTON, D. CAMBURN, D. ALWIN, L. YOUNG-DEMARCO, 1988, "The life history calendar: A technique for collecting retrospective data", in C. Clogg (ed.), *Sociological Methodology*, San Francisco, Jossey-Bass, pp. 37-68.
- FREEDMAN R., 1974, "Community-level data in fertility surveys", *WFS Occasional Paper*, 8, The Hague, International Statistical Institute.
- GARDNER R., 1981, "Macrolevel influences on the migration decision process", in G. De Jong and R. Gardner (eds.), *Migration Decision Making: Multidisciplinary Approaches to Microlevel Studies in Developed and Developing Countries*, New York, Pergamon Press, pp. 59-87.
- HENRY S., B. SCHOUMAKER, C. BEAUCHEMIN, 2004, "The impact of rainfall on the first out-migration: A multi-level event-history analysis in Burkina Faso", *Population and Environment*, 25(5), pp. 423-460.
- HUGO G., 1985, "Investigating community-level effects on population movement", in J. Casterline (ed.), *The Collection and Analysis of Community Data*, Voorburg, International Statistical Institute, pp. 157-175.
- INSD, 2000, *Analyse des résultats du recensement général de la population et de l'habitation de 1996*, Ouagadougou, Institut national de la statistique et de la démographie.
- LUCAS R., 1997, "Internal migration in developing countries", in M. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics*, Amsterdam, Elsevier Science Publishers, pp. 721-798.
- LUCAS R., 2000, "Migration", in M. Grosh and P. Glewwe (eds.), *Designing Household Survey Questionnaires for Developing Countries. Lessons from 15 Years of the Living Standard Measurement Study*, Washington D.C., World Bank, pp. 49-81.
- MENSCH B., M. ARENDS-KUENNING, A. JAIN, 1996, "The impact of the quality of family planning services on contraceptive use in Peru", *Studies in Family Planning*, 27(2), pp. 59-75.
- NEW M., M. HULME, P. JONES, 2000, "Representing 20th century space-time climate variability. II: Development of 1901-1996 monthly terrestrial climate fields", *Journal of Climate*, 13, pp. 2217-2238.
- OBERAÏ A., R. BILSBORROW, 1984, "Theoretical perspectives on migration", in R. Bilsborrow, A. Oberaï and G. Standing (eds.), *Migration Surveys in Low Income Countries: Guidelines for Survey and Questionnaire Design*, London & Sidney, Croom Helm, pp. 14-30.
- POIRIER J., V. PICHÉ, G. LE JEUNE, B. DABIRE, H. R. WANE, 2001, "Projet d'étude des stratégies de reproduction des populations sahéliennes à partir de l'enquête 'Dynamique migratoire, insertion urbaine et environnement au Burkina Faso'", *Cahiers Québécois de démographie*, 30(2), pp. 289-309.
- SCHOUMAKER B., 1999, "Analyse multi-niveaux et explication de la fécondité dans les pays du Sud", in D. Tabutin, C. Gourbin, G. Masuy-Stroobant and B. Schoumaker (eds.), *Théories, paradigmes et courants explicatifs en démographie*, Louvain-la-Neuve, Academic/L'Harmattan, pp. 331-357.
- SCHOUMAKER B., 2001, *Analyses multi-niveaux de la fécondité. Théories, modèles et applications au Maroc rural*, Doctoral thesis, Institut de démographie, Louvain-la-Neuve, Université catholique de Louvain.
- SCHOUMAKER B., DABIRÉ B., GNOUMOU-THIOMBIANO B., 2004, "La collecte de biographies communautaires pour l'analyse des déterminants contextuels des comportements démographiques. L'expérience d'une enquête au Burkina Faso", *Les travaux de l'UERD*, no. 14, Ouagadougou, Unité d'enseignement et de recherche en démographie, 38 p.
- SCHOUMAKER B., 2005, "Local context and marital fertility in Burkina Faso", Paper presented at the 25th International Population Conference, Tours 2005, IUSSP.
- TUO D. S., 2002, *Analyse multi-niveaux et mortalité des enfants dans les pays du Sud. Une revue de littérature*, Post-graduate dissertation, Institut de démographie, Louvain-la-Neuve, Université catholique de Louvain.
- UNITED NATIONS, 1995, *Report of the International Conference on Population and Development : Cairo, 5-13 September 1994*, New York, United Nations.
- UNITED NATIONS, 1998, *National Population Policies*, New York, United Nations.

SCHOUMAKER, Bruno, BONAYI-DABIRE, Hubert, GNOUMOU-THIOMBIANO, Bilampoa – **Collecting Community Histories to Study the Determinants of Demographic Behaviour**

This article presents a retrospective survey of localities, designed to be linked to an event history survey of individuals focusing on migration in Burkina Faso. During the survey, carried out in 2002, data on around ten broad topics were collected from respondents in almost 600 localities. After outlining various possible methods for collecting community data to study migration, the article considers a number of issues in turn: selection of the survey localities, organization of the survey, and data collection methods. The final part summarizes the results of several studies based on this data, and gives a brief outline of ongoing projects. One of the conclusions drawn is that the approach taken is a cost-effective means of collecting, within a reasonable timeframe, a wealth of high-quality community data for multi-level longitudinal analyses of demographic behaviour.

SCHOUMAKER, Bruno, BONAYI-DABIRE, Hubert, GNOUMOU-THIOMBIANO, Bilampoa – **Collecter des biographies contextuelles pour étudier les déterminants des comportements démographiques**

Cet article présente une enquête rétrospective sur des localités, conçue pour être reliée à une enquête biographique sur les migrations au Burkina Faso. L'enquête, réalisée en 2002, a collecté des données sur une dizaine de grands thèmes auprès d'informateurs dans près de 600 localités. Après avoir présenté différentes approches possibles de la collecte de données contextuelles pour étudier la migration, nous traitons successivement de plusieurs questions: la sélection des localités de l'enquête, son organisation et la présentation des outils de collecte. Les résultats de quelques travaux reposant sur ces données et des projets en cours sont abordés dans une dernière partie. Nous montrons que l'approche adoptée permet de collecter des données contextuelles riches et de bonne qualité à un coût et dans des délais raisonnables, permettant des analyses longitudinales multiniveaux des comportements démographiques.

SCHOUMAKER, Bruno, BONAYI-DABIRE, Hubert, GNOUMOU-THIOMBIANO, Bilampoa – **Colectar biografías contextuales para el estudio de los determinantes de los comportamientos demográficos. La experiencia de una encuesta realizada en Burkina Faso**

Este artículo presenta una encuesta retrospectiva sobre localidades, concebida para unirla a los datos de una encuesta biográfica sobre las migraciones en Burkina Faso. La encuesta, llevada a cabo en el 2002, recogió información sobre diez grandes temas a través de agentes en unas 600 localidades. Después de presentar los diferentes métodos disponibles para la colecta de datos contextuales en el estudio de las migraciones, tratamos las cuestiones siguientes: la selección de localidades incluidas en la encuesta, su organización y la presentación de los útiles de colecta. En la última parte del artículo comentamos los resultados de algunos trabajos y proyectos en curso basados en estos datos. Mostramos que el método utilizado permite coleccionar datos contextuales ricos y de buena calidad a un coste y un plazos razonables, haciendo así posibles los análisis longitudinales multinivel de los comportamientos demográficos.