Study of family history
France's 1999 Family Survey

In the March 1999 population census, 380,000 men and women living in private dwellings filled out an additional schedule on the subject of their "family history," including questions on their origins, children, partnerships, and social history, as well as the languages (both national and regional) customarily spoken in their families. The same information was gathered from 6,600 people living in institutions, including 1,700 male prison inmates.

The 1999 Family Survey demographic questionnaire has been thoroughly redesigned, and is the first to be addressed to men as well as women. It will provide material for numerous studies, covering family histories in all their complexity. The analysis of the supplementary questions on languages and dialects will, moreover, yield eagerly awaited information on France's linguistic diversity and the dynamics of national integration in the course of the twentieth century.

Preliminary findings on inmates were published in April 2000. The final file will be made available to the processing group in December 2001 and to the general public in 2002.

For the sake of simplicity, our discussion of methodological issues will focus on private households.

There are two main sources of demographic information in France. The census provides data on the state of the population at a given date; vital records tell us about population changes as measured by births, marriages, and deaths. But both sources are too general to shed light on aspects of individual behavior such as number of children, birth timing, types of partnership, remarriage, and so on. The chief purpose of the census is to provide localized information—at a very detailed geographic level—notably to determine the resident (de jure) population of each municipality (commune), authenticated by government decree. The prime purpose of vital records is to define personal identity.

The Family Survey completes this information system. It has been coupled with the population census since 1954, making it one of INSEE's oldest sample surveys. It is mainly intended to track the emergence of new family forms, through a retrospective and biographical questionnaire that reconstructs the demographic history of generations. Each successive survey also devotes a set of questions to a specific, topical theme, such as female participation in the labor force in 1982 or child care in 1990. The subject chosen for 1999 was the inter-generational transmission of languages and dialects.

The survey is conducted on very simple principles. In a fraction of the "enumerator sectors" defined for census data gathering (one in fifty since 1962), the persons included in the field of the Family Survey receive a special questionnaire in addition to their individual schedule. Respondents are asked to fill out the special questionnaire themselves and return it to the enumerator with the other census forms. One of the advantages of this arrangement is that it enables us to remove from the individual schedule a number of demographic questions that are included in censuses of other countries or recommended by the United Nations but would be expensive to put to the entire population (number of children born, detailed marital status, dates of death of partner or children, etc). Reciprocally, the Family Survey uses information gathered in the census.

1 The size of the survey sample is dictated by the need to break down most of the data according to many criteria, such as age and generation.
A substantial revision, based on the work of the French National Institute for Demographic Studies (INED)

The 1999 survey was the occasion for a major overhaul, largely inspired by the work of the French National Institute for Demographic Studies (Institut National d'Études Démographiques: INED) on family situations and histories.

The questions on partnership and family situations (demographic section) were considerably revised. Special care was taken to distinguish between de facto and de jure situations, and to identify the individuals who have lived in several successive partnerships; new questions were added on adopted children and on the partner's children, as well as on grandchildren and great-grandchildren.

The survey field was expanded to two new categories: men and people aged over 64. For the institutional population, the field was extended to persons living in retirement homes and prison inmates.²

Men and women

Until 1990, the Family Survey was restricted to women—in keeping with a favorite habit of demographers. At the outset, it covered married women who had completed their fertility, i.e., aged 45-54. The field was later broadened to non-single women aged under 65, then to unmarried women. The 1990 survey covered 340,000 women aged 18-64.

Under this policy of continuous enlargement, the extension to men will yield reliable information on a relatively undocumented subject: paternity. Social contrasts are not identical for men and women. The frequency of never-married status rises with educational attainment for women, and decreases for men. Mean fertility-indicator values can diverge slightly between men and women (because of differential mortality and migrations). Most important, however, male behavior is more diversified than female behavior. There are more childless men or father with many children, owing to the higher frequency of never-marrieds and remarriages among men. Lastly, as children live more often with their mothers after a breakup of the parental union, they are more likely to live with a stepfather than with a stepmother. From the adult standpoint, residence with the partner's children is a growing phenomenon among men, but remains infrequent for women.

Tests performed since 1993 have shown that the male response rate was equal to the female rate, and that there were no notable differences in the quality of the responses. The only difficulty that emerged concerns young male bachelors, who do not always see the point of responding to a questionnaire on the family. It is to counter this phenomenon that INSEE has changed the name of the questionnaire from "Study of Families" to "Study of Family History": even if the respondent has not started a family of his (or her) own, everyone has a family history.

Barring some gender differences in the wording, the questionnaire is identical for men and women. Now in most cases the two members of a partnership have very similar family histories. To avoid asking for the same information twice, we have therefore disconnected the male and female samples. In a given dwelling, we have surveyed either the men or the women—which, in practice, means all men or all women in the household aged 18 or older.

No more age limit

The extension to the over-64s is justified by the retrospective character of the Family Survey. We can thus recapitulate French demographic history by going back as far as possible, beyond the baby boom. In the language section, we can go back even further, since the oldest respondents

² The Family Survey had already been extended since 1975 to women living in university residences or in female workers' hostels. In consequence, the only institutions still excluded in the 1999 survey were hospitals, barracks, and religious communities.
will report the language their parents spoke to them when they (the respondents) were five years old, i.e., on the eve of World War I.

The elimination of the upper age limit does, however, pose a number of risks that complicate the data-gathering process, such as selection bias due to mortality, defective memory, or physical or mental inability to fill out a questionnaire by oneself.

Children, step-children, grandchildren, great-grandchildren

As in earlier versions, the questionnaire begins with questions on the respondent's children, including adopted children. For each child, the respondent is asked to indicate the gender, date and place of birth, and, for adopted children, the date the child came to live in the household. Where appropriate, the ages at leaving the parental home and the dates of any deaths (with a special mention of stillborn children, which people make a point of reporting) enable us to document the periods in which the respondent has raised children. The questions on adopted children yield information never previously available in France: the social and demographic characteristics of the foster parents, and how various methods of adoption have spread since World War I (the child's place of birth allows a distinction between national adoption and international adoption).

The next set of questions concerns stepchildren, defined here as the children of a partner (or ex-partner) raised by the respondent. The questions are identical to the questions on children, making it possible, for study purposes, to examine children and raised stepchildren together.

Three further questions complete the parenting-history section: (1) other children taken in; (2) presence of grandchildren and great-grandchildren; (3) number and age of the eldest, which should finally enable us to find out when people become grandparents or great-grandparents (for the moment, this can be determined only by means of statistical simulations).

Partnerships, with or without marriage

The questions on partnership history concentrate on de facto unions, with or without marriage. After a question on current status, we ask for the dates of the start and termination of the first and latest partnerships, and, where applicable, dates of marriage, separation, divorce, and death of spouse. As recommended by the National Commission on Information Technology and Civil Liberties, (Commission Nationale de l’Informatique et des Libertés: CNIL) the survey questions are confined to no more than two partnerships, the first and the latest; past partnerships are restricted to those involving a cohabitation of at least six months.

The questions on the characteristics of the latest partner (date and place of birth, any marriage prior to this union, occupation) give an idea of the social status of partners whose relationships have broken up. For the purposes of demographic analysis, it would have been preferable for respondents who had entered into several partnerships to be asked about the first partner, but tests have shown the reluctance of people to give excessive details on a partner who has been "replaced" by another in the partnership history.

The questions on partnerships and on children and stepchildren yield a precise description of different family situations, notably "single-parent" and "reconstituted" families. Accordingly, they give equal treatment to consensual unions and married couples, including any children not living with the parent or step-parent. The retrospective nature of the questions also makes it possible to track the sequence of types of families in which respondents may have lived.

Social origins and history

As a complement to the standard questions on the number of children born to the respondent's mother, the respondent's birth order, and the occupations of the respondent's parents, new questions have been added on the parents' places of birth and on whether they are still alive. The
schedule also includes questions on the main stages of the respondent's occupational and social history: age at leaving school, at leaving parents' home, and at starting first job; breaks in labor-force participation; unemployment spells.

**Language section**

Questions on the languages spoken in the family (and beyond, with close relations) seek to describe France's linguistic diversity and the momentum toward national integration all during the twentieth century. This represents a major innovation, as French official statistics had never addressed the topic on a national scale. Thanks to the data collected—and eagerly awaited by historians and linguists—INSEE will at last be able to establish a basic level of comparability with other countries in the field of languages spoken and their family transmission.

The questionnaire deliberately gives equal status to regional languages and national languages, as the latter are often dialects, while the former can be the extensions of neighboring countries' national languages into border areas.

Concerning immigrants, a vital item of explanatory information is not directly gathered in the survey but will be extracted from the individual census schedules (see below): the date of arrival in France. Absent this datum, we would run the risk of attributing to a cultural effect a behavior that mainly depends on the length of residence in France.

Another decisive information item, which applies to everyone, is the desire of respondents who do not live in their region or country of origin to return there some day. This subject is covered in the last question of the survey.

**Family Survey and census**

Thanks to the Family Survey's linkage to the census, we can survey a large sample at a very low marginal cost. As in earlier surveys, and to keep data-collection procedures simple, INSEE has chosen to cover entire "enumerator sectors,"³ of which some are totally dedicated to the survey of males and others to the survey of females. As in previous years, the survey was carried out in one of every fifty dwellings.

Along with the census material, the participating enumerators distributed the following items to each sampled household: a number of Family Survey schedules equal to the number of adult household members of the designated gender, a leaflet describing the aims of the survey, the findings of previous surveys, a justification of the retrospective questions, and instructions for filling out the schedules.

The same method—namely, a respondent-administered supplementary schedule—was used for the Daily Life and Health (Vie Quotidienne et Santé: VQS) survey. This was also combined with the 1999 census, and its findings were used to construct the sample for the Handicaps-Disabilities-Dependency (Handicaps-Incapacités-Dépendance: HID) survey of persons living in private dwellings. The samples of the Family and VQS surveys were picked in a coordinated manner so as to prevent the same households from being subjected to both surveys.

**Guaranteeing confidentiality**

The Family Survey is, indirectly, a nominal survey (i.e., of named individuals), since the identifiers used for the census (district number, rank of building in district, rank of dwelling in building) appear on the schedules gathered. However, the schedules do not contain any names or addresses. To

---

³ For the census, the collection unit is the district (of which there are 330,000 in all). Each of the 115,000 enumerators hired by municipal governments to conduct the census was in charge of one or more districts. On average, an enumerator sector comprised a population of about 225 households and 500 people.
ensure even greater confidentiality, the responses have been sent directly to INSEE by the census census supervisors (délégués)\(^4\) rather than being forwarded via the town halls.

**Use of census information**

The Family Survey data base will not be restricted to the information contained in the survey schedules. It will also incorporate most of the information derived from the census. The matching will be performed using the census identifiers, date of birth, and gender.

This information gathered prior to the survey itself is, of course, very useful. It tells us about the household composition, which enables us in particular to compare the children born to the respondent with the children who are still the respondent's dependents, and to determine all the persons residing with the respondent. Most important, the information provides the main explanatory variables for processing the data: legal marital status, place of residence, educational attainment, four-digit socio-occupational category, nationality indicator, date of arrival in France for immigrants—all these items being available for the respondent as well as all other members of the household. In fact, apart from the names of businesses and some dwelling-description items, the Family Survey draws on almost all the census information. That is why all the households included in the field of this survey will be incorporated into the "supplementary" processing field for the population census.\(^5\)

In addition, the census data yield information on persons who have refused to fill out the survey schedule.\(^6\)

**The sample**

With a coverage of one in fifty dwellings, and allowing for the forecast non-response rate of 10%, the sample size could have been set at 470,000 adult females or 430,000 adult males. INSEE wanted to avoid a situation where the survey's extension to men would halve the female portion of the sample, which—at all events—remained the focus of analysis. On the other hand, the male sample had to be large enough to allow detailed analysis of both the demographic data and the language data.

We therefore adopted an intermediate solution. The sample's total size was set at 445,000 people. The gender ratio is not 1:1 but 3:5, i.e., 167,000 men and 278,000 women. In addition, the sample had to be designed to achieve representativeness at the regional level (again, this was a first). To this end, we chose average sampling rates of 1/170 for men and 1/110 for women. Of necessity, the minimum numbers were set at 5,000 men and 8,500 women in regions where these sampling rates led to smaller numbers being chosen (except in Corsica and the Limousin). In some areas, we found it necessary—for the purpose of processing the language section of the survey—to increase the size of the samples obtained on the basis of these rules, especially for men. The areas concerned are the North ("Nord"), Alsace, Corsica, the Catalan country, the Basque country, and Brittany.

In practice, we began by selecting a representative sample of districts in each region or fraction of a region (when the survey included a "languages" extension). The representative sample was chosen by a systematic random draw after a classification of neighborhoods by socio-demographic characteristics. The classification method was developed by Nicole Tabard. The census supervisors' zones containing these districts have been included in the survey sample. Within each

\(^4\) I.e., the direct superiors of enumerators


\(^6\) Because of the private and retrospective character of the information requested from respondents, INSEE has not wished to make the Family Survey compulsory. The survey has been certified as "in the public interest" by the National Council on Statistical Information (Conseil National de l'Information Statistique: CNIS) and has been approved by the National Commission on Information Technology and Civil Liberties (Commission Nationale de l'Informatique et des Libertés: CNIL).
zone, the enumerator sectors for the male survey or the female survey have been chosen with specific inclusion probabilities for each region or fraction thereof.

The participation rate was slightly lower than expected: in the end, 380,000 schedules were collected, of which 145,000 from men and 235,000 from women. The non-compliance rate can thus be provisionally estimated at 14%, but the final figure will not be known until the exact census data have been factored in. We will then be able to calculate the non-compliance rate by gender and age, marital status, socio-occupational category, household size, etc., and develop adjustment coefficients.

Numerous opportunities for data editing...

How credible are statements gathered in a respondent-administered questionnaire? The cornerstone of any official statistical survey is trust. The basic assumption is that respondents describe their situation truthfully. But that is not enough, and the questionnaires must incorporate internal verification procedures for corroborating and editing the data gathered.

The Family Survey covers three event timings: that of births, that of partnerships, and that of social history. The first timing is complete; the second is a simpler version of the detailed chronology that one can compile in demographic surveys carried out by field officers; the third is summary. But the crucial issue is the coordination of the three timings. Above all else, we must be able to identify the marital and social context prevailing when children arrive in the family.

Many information items can be edited internally by comparison with overall responses. For example, the number of partnerships can be measured from male responses or female responses. Likewise, the number of great-grandparents can be deduced from the direct responses to the question “do you have great-grandchildren?” or estimated from the number of persons who report having grandchildren and living parents, taking into account the number of brothers and sisters of these respondents.

In addition to these internal edits, we can compare the survey findings with information from other sources, such as the annual number of marriages and births in the vital records. Despite their being complicated by migrations, deaths, and the fact that some men may not have recognized their children, these edits will nevertheless allow a fairly strict monitoring of the quality of the results obtained.

...including the language section

The same goes for the evolution in language distribution. The survey asks all generations to list the languages received from the preceding generation and those handed down to the following generation. The behavior of a given generation will thus be described from three angles: by itself and by the two surrounding generations. We can thus contain a risk emphasized by the linguists and dialectologists whom we have consulted, namely, that language prevalence is often understated or overstated. The inter-generational loss rate may exhibit a dubious slope—either too steep or too shallow—but the “slope of slopes” for all generations combined will be reliable. It will even allow an estimation of reporting bias.

An eclectic processing group

The data gathered will furnish material for many studies, covering the individual demographic events described in the schedules and the family histories in all their complexity. The starting point will be the comparison of the timings covered in the survey. Given the sample size, these studies will be able to focus on the current situation and evolution of demographic patterns in each social group or region. Longer-term analyses will also be possible, drawing on the findings of earlier Family Surveys. Beyond that, there are international comparisons: demography is particularly well-suited to this type of exercise, and many countries will obtain broadly similar data from their future censuses.
Most phases of the survey’s preparation involved extensive consultation with INED researchers. For the section on languages—to be processed in partnership with the "Language-Patterns Observatory" (Observatoire des Pratiques Linguistiques) recently set up by the French Ministry of Culture—many other organizations had also been asked to help: the Executive Council for the French-Speaking World (Haut Conseil de la Francophonie) and the Social-Action Fund for National Languages (Fonds d’Action Sociale pour les Langues Nationales), and the cultural-affairs or language-affairs delegations for regional languages (such delegations operate in about fifteen French départements).

In sum, we have worked with a wide variety of partners. The decision has therefore been taken to set up a "survey processing group" at INED that will be in charge of making the data available and organizing studies outside INSEE. The group has an open-door policy. Readers who would like to take part in the survey processing can contact Francine Cassan (+33 1 41 17 54 22, francine.cassan@insee.fr) or Cecile Lefèvre (+33 1 56 06 20 98, lefevre@ined.fr).

**Processing timetable**

Before we can produce reliable findings, we will naturally need to take into account not only the sampling-plan weights but also (1) the adjustments for partial and total non-responses, based on census data, and (2) the assessment of the robustness of the findings.

In June 2000, we will have a preliminary cleaned and weighted file, whose weighting will be based only on the sampling plan, and whose content will be limited to the information collected in the survey questionnaire. The file will be used to produce preliminary findings.

In December 2000 a preliminary study file will be released. It will incorporate the variables included in the "main" processing of the census (i.e., almost all the variables except for "occupations and socio-occupational categories" and "households-families"), the variables of study on individuals and children, and the information gathered from the survey’s language section. This file will be weighted to reflect non-responses by gender, age, and household size (preliminary weighting).

A memorandum on the reliability of the findings at the national level, after an external edit, will be published in June 2001. Lastly, in December 2001, the complete file—including the "supplementary processing" variables of the census, the variables of study on households, and the final weighting—will be compiled and made available to the processing group. It will be released to the public in the course of 2002.

Francine Cassan, François Héran, and Laurent Toulemon

François Héran is Director of INED.
Laurent Toulemon is head of the Demographic Studies Division at the INSEE head office.
Francine Cassan is on the Division’s staff.
The legal statuses captured in the census and in vital records do not take family histories into account. For example, the census does not distinguish between married persons and remarried persons; it does not precisely identify unmarried couples or the relationships between adults and children in a given household. Vital records, by contrast, distinguish between "legitimate" and "illegitimate" children but tell us nothing about the type of relationship between parents of children born outside marriage (consensual union, duration of cohabitation).

To track these phenomena, INED conducted a special survey on family situations in 1986, using a sample of 5,000 based on the INSEE Labor-Force Survey. Households comprising persons in "atypical" situations were over-represented. These "atypical" persons included ever-married adults living alone, adults cohabiting without being married, children of whom the household reference person is not the parent, and children with only one parent residing in the household. The schedule contained retrospective questions on all partnerships (with or without marriage), on the dates of birth and change of residence of all of the respondent's children (whether living in the household or not), and on all children living in the household who were not the respondent's children.

The INED survey allowed a description of the de facto family situations of adults and children. It showed that "family histories," even if they diverged from the patterns legalized by marriage, were not unspeakable "family secrets." On the contrary, respondents willingly answered questions on their family status and histories—questions that they regarded as down-to-earth and relevant to their personal situation. However, it suffers from two limitations: first, despite the efficient sampling plan, the sample's small size precludes obtaining precise information on the more uncommon situations; second, it excludes people aged over 50, which rules out a description of long-term trends.

A second survey of the same kind was conducted in 1994 as part of the European Fertility and Family Surveys (FFSs). This 23-country operation gave rise to more than seventy international-comparison projects—now in progress—based on standardized data files. For information on these studies, see http://www.unece.org/ead/pau/f_home1.htm.
Surveys of institutional population

People living in university residences or workers' hostels were included in special samples. Like people living in private dwellings, they filled out a supplementary schedule during the population census. In all, 1,900 schedules were collected.

Two separate operations were conducted to handle the survey’s extension to people living in retirement homes and prison inmates.

Retirement homes: a simpler schedule, filled out during another survey

France’s 6,000 nursing homes and retirement homes accommodate about 400,000 people, of whom half are over 85 years old, and three-quarters of whom are women. Tests carried out in retirement homes showed the need to improve data collection in two ways: by simplifying the schedule and providing assistance for respondents interviewed by the enumerators. This was liable to complicate the collection task for enumerators. INSEE therefore decided to conduct the "retirement homes" part of the survey not during the census, but at the end of the interviews for the Handicaps-Disabilities-Dependency (Handicaps-Incapacités-Dépendance: HID) survey, carried out in late 1998, notably among 5,000 residents of retirement homes. If the person consented and was capable of responding without assistance, the field officer continued the interview with selected questions from the simplified version of the Family History Study (i.e., the questions not already included in the HID form). This procedure yielded more than 3,000 completed schedules.


Survey of prisoners: an entirely new operation

As the census included prison inmates, INSEE decided to include them in the survey as well. Prison demographics are a very little-known subject: France has a prison population of 50,000, 4% of whom are women. There is reason to believe that its profile resembles that of the most disadvantaged categories, but at the moment we have no comprehensive study to corroborate this assumption. The prison administration gave its unqualified support to a test run—which turned out to be highly positive—and then to the survey itself. After the test, the survey was limited to males, and the schedules slightly modified to ensure that the interview would be appropriate to the situation of incarcerated males, but without inserting specific questions on the prison environment. The schedules were filled out in the course of interviews conducted in total privacy, and were forwarded to INSEE without being seen by prison personnel. The survey was undertaken in 28 institutions and 1,700 schedules were completed.
Languages and dialects: a long wait

The issue of languages spoken in families was addressed in the nineteenth century, when Victor Duruy, Minister of Education under the emperor Napoleon III, requested elementary-school teachers and municipal registrars to enumerate the French-speaking and non-French-speaking children in their municipalities (1863). Attention was drawn to the data by Jacques Revel and Michel de Certeau (Une politique de la langue: la Révolution française et les patois, Paris, 1975), but they were first analyzed in 1977, by the U.S. historian Eugen Weber. The figures are highly aggregated (municipalities are classified into three categories: French spoken by all children, by some, or by none).

For the more recent periods, we have regional or local series compiled from INSEE surveys or inquiries commissioned by the regions, départements or community groups. Because of the different methods used, however, it is impossible to compute national estimates from these figures.

The language section of the 1999 Family Survey will therefore yield long-awaited information. The project originated in a 1992 INSEE survey on "Family Education Commitments" conducted on a small sample of 5,300 parents of schoolchildren. The 1992 survey, which included questions on languages, demonstrated that it was possible to ask French people for information on the subject and calculate the rate of transition to French in a generation, for all of the country's national and regional languages. Another survey performed in 1992 on "Geographic Mobility and Social Integration of Immigrants" (Mobilité Géographique et Insertion Sociale des Immigrés: MGIS) confirmed the relevance of these questions for measuring the pace of cultural integration of families of foreign origin.

1. A few years later, the War Ministry, at Duruy's request, extended the survey to conscripts.
2. As an example, the issue was covered only once within the context of the population census—by INSEE's regional office in Alsace, in 1962.
The Family Survey in French overseas départements: another protocol, but one that preserves comparability with the survey in metropolitan France

The Family Survey combined with the census covers only metropolitan France (mainland + Corsica), but is comparable with the fuller Family Surveys introduced in the French overseas départements (DOMs).

At the request of INSEE's regional offices in the West Indies (Antilles) and Reunion Island, a specific Family Survey was designed for the DOMs. The operation was carried out along the lines of a classic household survey unconnected to the census. The first survey took place in Reunion in November-December 1997. There are plans to conduct it in the French West Indies as well.

Administered by a network of field officers, the DOM questionnaire, which runs to some thirty pages, is far more detailed than the metropolitan one. While compliant with the international guidelines set by the world program of Demographic and Health Surveys (DHSs), it is consistent with the 1986 and 1994 INED-INSEE surveys on family situations.

The operational details lie outside the scope of this presentation, but one major point is worth emphasizing: the DOM questionnaire includes the entire metropolitan schedule. The DOM form was designed to allow the possibility of publishing the findings for all of France by combining DOM data with metropolitan data.
References

1982 Family Survey


1986 INED-INSEE survey on family situations


1990 Family Survey

The tables of the 1982 survey volumes were updated respectively in:


Maréchal, M. - "Cycle de vie et milieu social selon l'enquête Famille de 1990," INSEE Résultats, no. 580-581; Démographie-société, 63-64.

These volumes also contain a bibliography of studies based on the 1990 survey.

Classification of neighborhoods based on information from 1990 census


1992 Family Education Commitments Survey (questions on languages)


1994 INED-INSEE survey on family situations and employment


1999 Study of Family History

The survey's very first findings, on male prisoners, were published in:
