

Collection of Survey Data on Contraception: An Evaluation of an Experiment in Peru

Noreen Goldman, Lorenzo Moreno, and Charles F. Westoff

The objective of this analysis is to compare two different approaches to the collection of information on contraceptive use. The data for this comparison are derived from the 1986 Demographic and Health Surveys (DHS) in Peru. Approximately 7,500 women were interviewed with the standard DHS questionnaire, whereas about 5,000 women received an "experimental questionnaire." The major difference between the questionnaires is the inclusion of a six-year monthly calendar in the experimental questionnaire that records pregnancies, contraceptive use, and postpartum information, in contrast to the more common tabular format of the standard questionnaire. The analysis demonstrates that although reports of contraceptive knowledge, ever-use, and current use are relatively robust to the variations in questionnaire design, estimates of past use are dependent on the survey instrument. Several different comparisons indicate that reporting of information on contraceptive histories in the experimental questionnaire is superior to that in the standard one. (STUDIES IN FAMILY PLANNING 1989; 20, 3: 147-157)

Over the past several decades, hundreds of large-scale surveys in developing countries have obtained data on contraceptive practice. In the decade between the mid-1970s and mid-1980s alone, 61 countries gathered fertility and family planning information through the World Fertility Surveys (WFS) and the Contraceptive Prevalence Surveys (CPS) (Population Information Program, 1985). By the end of 1989, the Demographic and Health Surveys (DHS) project will have provided similar information from 31 large-scale surveys worldwide. These surveys have become the primary source of data in developing countries for studying contraceptive behavior, in part because alternative sources of information such as family planning program service statistics usually cover only some of the population,

are restricted to particular family planning methods, and are often inaccurate (Laing, 1984; Jejeebhoy, 1988).

All of these surveys have had, among other objectives, the aim of estimating contraceptive prevalence at the time of the survey. Some, although far from all, have also attempted to collect information on previous use of contraception. The most common types of questions concerned with use prior to the survey have focused on ever-use, use between the last two births, and use since the most recent birth. However, most of the surveys have not obtained specific information on the timing and duration of contraceptive use, not even for a recent period prior to the survey (Laing, 1984).¹ In many cases, demographers argued that questions on the timing of use would place too much strain on a respondent's memory.

However, during the past 20 years, several attempts have been made to use calendars to collect fertility exposure information (including contraception) in surveys in both developed and developing countries. For example, in the United States, the 1975 National Fertility Study and the later National Surveys of Family Growth employed five-year and three-year monthly calendars, respectively. Selected surveys in Latin America dating back to the late 1960s incorpo-

Noreen Goldman, D.Sc. is Professor of Demography and Public Affairs and faculty associate, Office of Population Research, Princeton University, Princeton, NJ 08544, USA. Lorenzo Moreno, Ph.D. is a postdoctoral fellow, Demographic and Health Surveys, Institute for Resource Development/Westinghouse, Columbia, MD 21044, USA. Charles F. Westoff, Ph.D. is Director, Office of Population Research, Princeton University.

rated a 12-month "sexual activity table," and several surveys carried out during the past decade in the Philippines collected similar fertility exposure information in a monthly calendar format for a period of two to two-and-a-half years prior to interview (Gaslonde and Carrasco, 1982; Laing, 1984).²

Very few assessments of the reliability of data on contraception from fertility surveys have been carried out, even though standard types of demographic data from the same surveys have undergone rigorous evaluation (for example, United Nations, 1987; Goldman et al., 1985). In particular, virtually no attempts have been made to evaluate the quality of contraceptive data associated with the use of a calendar. Among the few studies on the reliability of contraceptive data, most have evaluated only reports of current use. Nevertheless, there is clear evidence that such evaluations are essential. For example, a recent assessment of the consistency of reporting of contraceptive use in three Korean national fertility surveys indicates large inter-survey differences in the completeness of reporting and demonstrates that reports of use for periods in the past are substantially less complete than reports of current use (Pebley et al., 1986). The authors stress the importance of good questionnaire design for eliciting accurate contraceptive histories.

The objective of this analysis is to compare two very different approaches to the collection of information on contraceptive use. The data used for this comparison are derived from the Demographic and Health Surveys in Peru. In 1986, a large-scale experimental field evaluation of different approaches to the collection of demographic data was undertaken in Peru³ as part of the project. The overall goal of this undertaking was to resolve a number of methodological issues in survey research in connection with the measurement of levels and determinants of fertility, contraception, child health, and infant and child mortality. One major objective was to assess the potential of a calendar approach to the collection of demographic data, including vital events, the proximate determinants of fertility, and employment and residence histories. In order to address these issues, an "experimental" questionnaire was designed for use in conjunction with the standard questionnaire previously developed for the Demographic and Health Surveys.

As described in more detail below, there are many differences between the two questionnaires, ranging from simple rewordings of questions and different orderings of the same questions to completely distinct formats for eliciting certain types of information. The primary concern in this analysis is the comparability of estimates of knowledge and current and

past use of contraception as derived from the standard DHS survey and the experimental survey.⁴ Much of our focus is on a comparison of estimates of recent use—that is, contraceptive prevalence during the five-year period prior to the survey. This evaluation implicitly involves an assessment of the experimental calendar approach to collecting dates of use with the tabular format incorporated in the standard DHS survey. Our ultimate objective is to determine the comparative merits of different approaches for measuring contraceptive behavior so as to improve the quality and usefulness of future fertility and family planning surveys.

The Demographic and Health Surveys in Peru

Since the goal of this study is to ascertain differences in response that result from two different questionnaires, survey conditions for the experimental and standard survey were held constant as much as possible. A two-stage cluster sampling design was used to draw a national sample of dwellings. At the final stage of sample selection, a systematic subsample of one in three dwellings was assigned to the experimental questionnaire and the remainder was assigned to the standard questionnaire. All women aged 15–49 were interviewed in each selected dwelling. Approximately 7,500 women were interviewed in total: 4,997 with the standard (or core) questionnaire and 2,534 with the experimental questionnaire.

So as to eliminate possible interviewer effects, the same set of interviewers was used for both the standard and the experimental surveys. During June and July of 1986, interviewers received a three- to four-week intensive training course in both questionnaires, including a week of field testing. The interviews took place between September and December of 1986. In most cases, interviewers administered the experimental questionnaire on separate days from the standard questionnaire. Response rates for both surveys were about 95 percent. Comparisons of basic characteristics of the two samples, such as age, marital status, and parity (which were collected with the same questions in both surveys), indicate very similar distributions (Goldman et al., 1989).

About one-third of the content of the experimental and core questionnaires is identical and two-thirds is devoted to experimental variants. The major difference between the two questionnaires is the inclusion of a monthly calendar in the experimental questionnaire, which records pregnancies, contraceptive use,

reasons for contraceptive discontinuation, breastfeeding, postpartum amenorrhea, postpartum abstinence, women's employment, and place of residence for the period January 1981 to the month of interview (see Appendix). Similar information is either not collected at all in the core questionnaire or is obtained from a tabular format, with questions referring to the interval between successive births. The questionnaires also differ with regard to the nature of the fertility and infant mortality information (a full birth history in the core and a truncated history in the experimental questionnaire) and with respect to the collection of data on contraceptive knowledge and availability, future use of contraception, fertility preferences, current pregnancy status, and other health-related questions. These and other differences between the experimental and the core questionnaires are discussed in detail in Goldman et al. (1989).

It is interesting to note that the overall length of the interviews was almost the same for the two questionnaires: 30 and 31 minutes for the core and experimental surveys, respectively. It appears that the time saved by inclusion of a truncated (rather than a full) birth history in the experimental questionnaire was compensated for by the inclusion of other types of information (such as employment and residence) not collected in the core.

Of primary concern in this analysis are the differences in the questionnaires with regard to information on contraceptive use. The third section⁵ of both questionnaires is devoted to the collection of information on contraception. In the first part of this section, data are collected on contraceptive knowledge, ever-use, availability, and acceptability. Questions on knowledge and ever-use are essentially the same in the two questionnaires: the respondent is first asked to mention spontaneously any method she knows; the interviewer subsequently reads a description of each method and asks the respondent whether she has heard about the method and whether she has ever used it. However, the questionnaires differ with regard to the ordering of methods. In the core questionnaire, the ordering proceeds, in general, from more to less effective methods: pill, IUD, injections, vaginal methods, condom, sterilization, rhythm, and withdrawal. By contrast, in the experimental questionnaire, the order is basically reversed: rhythm, withdrawal, condom, sterilization, injections, vaginal methods, IUD, and pill.⁶

The next part of both questionnaires is primarily concerned with information on current use of contraception and use within a recent period prior to the survey. Both questionnaires first obtain information

with regard to the current method and its duration of use.⁷ The remainder of this section differs completely between the questionnaires. In the standard questionnaire, information, including the timing of use, is obtained next with regard to the previous method (that is, the method preceding the current method) that was used subsequent to the last birth or marriage (that is, in the open interval). For women not currently using a method, information on type of method and duration of use is obtained only for the last method used in the open interval. Subsequently, in the standard questionnaire, information on use is collected in a tabular format for the interval preceding each birth since January 1981. The questionnaire allows for the coding of up to two methods within an interval; however, duration of use is reported only for the last method in an interval.

The experimental questionnaire is based on an entirely different approach to obtaining information on recent contraceptive use. After inquiring about current use and entering it into the calendar, interviewers administering the experimental questionnaire used the 72-month calendar (see Appendix) to probe for all previous segments of use between 1981 and interview date. Interviewers were trained to use information already coded in the calendar to aid the respondent's recall: months of pregnancy and birth had already been recorded in the calendar as of this stage of the interview. Months of contraceptive use were coded in the first column (including a code of "0" for nonuse) so that each month of this column contained one and only one code—a code for pregnancy, birth, nonuse, or use of a particular method (or a specified combination of methods).⁸

Comparisons of Knowledge, Ever-Use, and Current Use

In order to assess the importance of the order in which contraceptive methods are presented to the respondent, we compared estimates of knowledge of each method as derived from the two questionnaires. Table 1 contains the percentages of women who know about each method, both spontaneously and after hearing the description read by the interviewer; the former estimates, of course, should not be expected to differ between questionnaires. Both sets of estimates are very similar. For only two methods do the estimates disagree by more than three percentage points. Only for rhythm and injection are the differences statistically significant (at a 5 percent level). The order in which the methods are given does not appear to have a large affect on the resulting estimates of knowledge.

Estimates of the percent of ever-married women

Table 1 Knowledge of contraception by method, all women, core and experimental questionnaires, Peru

Method*	Percent who heard of method			
	Yes, spontaneous		Yes, probed	
	Core	Experimental	Core	Experimental
Pill	50.9	49.9	24.4	25.7
IUD	37.3	34.9	31.1	33.1
Injections	27.9	27.2	36.0	40.0
Diaphragm, foam, jelly	13.7	14.2	28.0	30.3
Condom	9.8	9.8	37.6	39.3
Female sterilization	8.8	8.7	63.6	64.9
Male sterilization	1.2	1.4	24.0	26.6
Rhythm	22.1	23.3	49.0	43.6
Withdrawal	2.6	2.9	39.1	37.0
(Number of women)	(4,997)	(2,534)		

*The ordering of methods is as follows: pill, IUD, injections, diaphragm, condom, female sterilization, male sterilization, rhythm, and withdrawal in the core survey; rhythm, withdrawal, condom, male sterilization, female sterilization, injection, diaphragm, IUD, and pill in the experimental survey.

who have ever used each of the contraceptive methods are presented in Table 2. The two surveys yield virtually identical estimates: 63.6 percent and 63.4 percent have ever used any method of contraception, as derived from the core and the experimental surveys, respectively.⁹ The estimates are very similar for each of the specific methods as well; none of the differences is statistically significant.

Estimates of current contraceptive use are shown in Table 3 for currently married women. Once again, estimates from the two surveys are in almost perfect agreement: the percentages of women using any method at the time of the survey equal 45.8 and 45.2 percent in the core and experimental surveys, respectively. Estimates of current use for each of the specific methods are not statistically different between the two questionnaires.¹⁰

There are, however, certain ambiguities with regard to simultaneous use of more than one method. After obtaining information on the current method, interviewers in the core questionnaire asked respondents whether they "regularly use any other method during the same month." It appears that a substantial

Table 2 Ever-use of contraception by method, ever-married women, core and experimental questionnaires, Peru

Method	Percent ever using method	
	Core	Experimental
Pill	21.7	21.1
IUD	10.9	10.9
Injections	9.0	9.8
Diaphragm, foam, jelly	7.9	8.0
Condom	9.4	9.4
Female sterilization	5.8	6.8
Male sterilization	0.0	0.2
Rhythm	38.5	36.6
Withdrawal	18.3	17.5
Any method	63.6	63.4
(Number of women)	(3,237)	(1,679)

Table 3 Current use of contraception by method, currently married women, core and experimental questionnaires, Peru

Method	Percent currently using method	
	Core	Experimental
Any method	45.8	45.2
Pill	6.5	5.8
IUD	7.4	7.6
Injections	1.4	1.1
Diaphragm, foam, jelly	0.9	1.1
Condom	0.7	0.7 ^a
Sterilization	6.2	7.2
Rhythm	17.7	17.5 ^b
Withdrawal	3.6	2.7
Other	1.5	1.4
No method	54.2	54.8
Total	100.0	100.0
(Number of women)	(2,899)	(1,493)

^aIncludes reported combinations of rhythm and condom (0.2) and condom and withdrawal (0.1). ^bIncludes reported combination of rhythm and withdrawal (1.9).

proportion of women acknowledge such multiple use: for example, 15 percent of current users (excluding sterilized couples) acknowledged using more than one method during the same month; not surprisingly, 60 percent of these multiple users reported their current method as rhythm, and 13 percent as withdrawal.¹¹ Although the experimental questionnaire did not specifically ask for multiple use, interviewers were trained to probe for such use and the coding of methods in the calendar included three combinations: rhythm and condom, rhythm and withdrawal, and condom and withdrawal. Six percent of current users (excluding sterilized couples) acknowledged using one of these three combinations; the same three combinations totaled to 9 percent of current users in the core. Note that whereas the core survey obtained information on simultaneous method use only for current users, the experimental survey allowed for the above-mentioned combinations for the entire period covered by the calendar.

Comparisons of Previous Use

The above comparisons suggest that the estimates of contraceptive knowledge, ever-use, and current use derived from the two surveys are in close agreement. Of particular interest in this analysis is the extent to which estimates of previous use are also consistent between surveys.

The first indication of a discrepancy between the surveys is the reported duration of use for the method used at the time of the survey. Although the wording of the specific question is similar¹² in the two surveys, the coding of the response is different. In the core survey, the response is coded in terms of the number of

months and/or the number of years, including a special code to indicate the method that the women used since the last birth. In the experimental survey, all responses are coded in terms of number of months, with a special code for 96 months or longer. Subsequently, only in the experimental questionnaire were interviewers instructed to enter the months of use of the current method into the calendar, with each month of consecutive use receiving the appropriate method code.

The net result of these differences is much more heaping of reported durations of use in the core survey. In the experimental survey, there is very little tendency for respondents to overreport rounded durations such as 6, 12, and 24 months, whereas a high proportion of durations are reported as such in the core. The extent of heaping on selected durations for the segment of current use is shown on the left-hand side of Table 4. The fact that the heaping is especially high for 24, 36, and 48 months in the core suggests that respondents (or interviewers) simply coded an integral number of years. In fact, over one-quarter of responses to the duration of current use were reported as years only. The absence of heaping in the experimental survey is undoubtedly due in large part to the use of a calendar that may have altered interviewer behavior in several ways. For example, interviewers may have verified reported durations in terms of calendar months; and, interviewers could not have accepted reported durations if such durations lead to an overlapping of use with pregnancy. On aggregate, the heaping appears to have produced slightly longer durations of reported use in the core survey: mean durations of current use of 39.4 and 38.6 months and median durations of current use of 23.4 and 21.5 months in the core and experimental surveys, respectively. The

right-hand side of Table 4 indicates that a large degree of heaping occurred in the core survey within closed birth intervals as well, whereas very little heaping is present in the calendar.

Although these results suggest better reporting of use in the experimental survey, it is not necessarily the case that the unheaped responses in the experimental questionnaire are more accurate than the heaped ones in the core.¹³ Hence, it is important to evaluate the relative completeness and accuracy of reports of previous contraceptive use by other criteria. Our evaluation is necessarily restricted to comparisons of aggregate estimates of use derived from the survey data. Although the goal of such an evaluation should be the determination of accuracy, we do not have any independent measures of contraceptive use that are demonstrably better than those derived from recent surveys. For example, service statistics are grossly incomplete and would be entirely inappropriate for measuring use in a country such as Peru where traditional contraceptive methods dominate.

The objectives of our aggregate comparisons are two-fold: first, to compare estimates of use, as of successive dates, between the two DHS surveys; and second, to compare estimates of current use reported in an earlier survey with estimates of use reconstructed from the DHS data for the date of the earlier survey. Although such calculations of aggregate consistency do not conclusively reveal the sources of discrepancy, we can usually be confident that reports of current use (from the earlier survey) are more complete than the reconstructed estimates in the later survey, which are derived from reported dates of use (Pebley et al., 1986).

Reconstruction of the distribution of contraceptive use as of dates prior to the survey is a straightforward calculation from the experimental data, since the calendar allows the analyst to determine use status as of any month between interview and January 1981. However, the same calculation cannot be readily carried out from the standard survey because dates of use are not provided for all segments of use: that is, only durations of use are reported for segments of use in closed intervals and for the episode of use preceding the current method in the open interval. Hence, we undertook the creation of a calendar from the data reported in the standard survey.

The actual steps involved in the creation of this simulated calendar are too lengthy and tedious to present here, but they are described in detail elsewhere (Goldman et al., 1989). The goal of the procedure was to use the reported information from the core questionnaire to create a contraceptive and birth history in

Table 4 Index of heaping on particular durations of contraceptive use, for current use and use of last method in closed intervals, ever-married women, core and experimental questionnaires, Peru

Duration (months)	Current use		Use in closed interval*	
	Core	Experimental	Core	Experimental
6	1.4	1.1	2.4	1.4
12	4.1	1.4	7.9	1.0
24	11.7	0.6	13.3	0.8
36	9.1	1.7	10.9	1.6
48	10.1	0.9	5.3	0.0

Note: The index of heaping is equal to the number at the reported duration divided by the average number at the two consecutive durations on either side. For example, the index for six months equals the number of segments with duration of six months divided by the number of segments with durations of four, five, seven, and eight months divided by four.

*In order to make the estimates from the core and experimental surveys comparable, this calculation includes only those closed intervals that began subsequent to January 1981.

the same format as the first two columns of the calendar in the experimental questionnaire. Data on the dates of pregnancy and birth and months of use for the current method could be directly entered into such a calendar from the information provided in the standard questionnaire. For last segments of use that were reported to have resulted in contraceptive failure, the dates of use could be determined from the reported duration and the date of the ensuing pregnancy. However, for the remaining last segments of use (as well as for segments of use that preceded current use in the open interval), starting dates of use had to be imputed.¹⁴ Since information on duration of use was not collected for the next-to-last methods in closed intervals, these episodes of use were not coded in the simulated core calendar. Although one might be concerned with the use of imputed data for the purposes of evaluation, it appears as if the results are not very sensitive to the particular imputation scheme. Experimentation with various imputation schemes has demonstrated that estimates of prevalence and, especially, estimates of failure and discontinuation are much more sensitive to reported durations of use than to reports of the specific timing of use within a narrow period.¹⁵

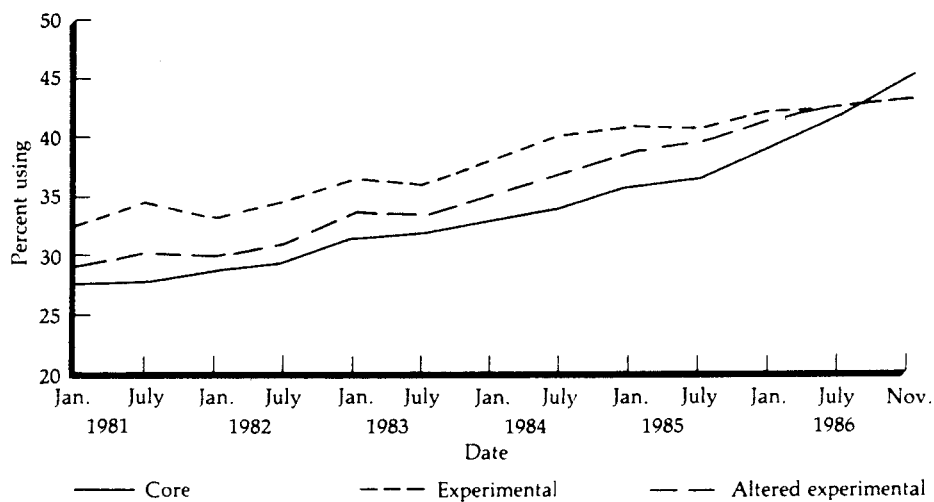
Trends in Contraceptive Use, 1981–86

Based on the resulting simulated calendar from the core questionnaire and the actual calendar in the experimental questionnaire, trends in contraceptive use for the period 1981 to 1986 were estimated. Figure 1

shows the resulting percentages of ever-married women 15–44 using a contraceptive method as of each of the selected dates.¹⁶ The graph illustrates clearly the fact that although the surveys yield similar estimates of current use, they produce substantially different estimates of prevalence for the recent past. It is interesting to note that the estimates diverge as one moves backward in time, but only up to about 18 months prior to survey. From this point back in time to January 1981, the estimates based on the experimental survey remain about five percentage points (or about 15 percent of use) higher than those based on the core. If we assume that women are unlikely to report use that did not occur, the pattern of discrepancy suggests that the core questionnaire failed to capture completely episodes of contraceptive use. (The altered experimental survey shown in Figure 1 is discussed below.)

The relative shortcomings of the core as compared with the experimental survey are apparent from a comparison of estimates of use reconstructed from the DHS surveys with estimates of current use from the 1981 Contraceptive Prevalence Survey in Peru. Estimates of use reconstructed as of the date of the 1981 CPS are shown in Table 5 and are compared with those reported in the CPS. The estimates confirm the superior contraceptive history data collected in the experimental survey, but indicate that even this estimate (34.6 percent of ever-married women aged 15–44 using a method) is significantly below the value of 38.1 percent reported in the CPS. Unfortunately, there are no other independent estimates of contraceptive prevalence in Peru for the 1981–86 period.

Figure 1 Use of any contraceptive method among ever-married women aged 15–44, as reported in the core versus experimental questionnaire (including altered experimental survey), Peru, 1986



Analysis of other survey data on contraceptive use have indicated that episodes of use of ineffective methods are generally reported less completely than those of modern methods (for example, Pebley et al., 1986; Laing, 1984). In Figure 2, estimated trends in the prevalence of modern methods (pill and IUD) are compared with those of rhythm, the most common method used in Peru. The data suggest that the under-reporting in the core survey relative to the experimental survey occurred to a similar extent (in percentage

terms) with regard to both types of methods. However, a comparison of both DHS surveys with reported prevalence as of the CPS date suggests that, in fact, the experimental survey obtained complete reporting of the most effective methods (pill and IUD) for a date more than five years prior to the survey. Reporting was less complete with regard to withdrawal and other methods. By contrast, estimates derived from the core are considerably below those from the CPS for all methods.

Table 5 Reconstruction of percent of ever-married women 15-44 practicing contraception, by method, as of the date of the 1981 Contraceptive Prevalence Survey, Peru

Method	Contraceptive Prevalence Survey	Demographic and Health Survey	
		Experimental	Core
Any method	38.1*	34.6	27.8**
Pill	4.8	4.9	4.1
IUD	3.9	3.8	3.1
Injection, diaphragm, condom	3.8	2.7*	2.5*
Rhythm	14.3	13.9	11.0**
Withdrawal	3.6	2.4	1.9
Sterilization	4.4	5.3	4.1
Other	3.3	1.4*	1.1*

Note: The National Contraceptive Prevalence Survey took place between August and December of 1981. Statistical tests are based on the assumption of simple random samples in each survey.

*Estimate is significantly different (at a 1 percent level) between the DHS and the CPS.

**Estimate is significantly different between the DHS core and experimental questionnaires.

†If those women are included who answer negatively to the question on current use of contraception, but affirmatively to a probe question on whether they used contraception in the past month, this figure would increase to 39.2 percent.

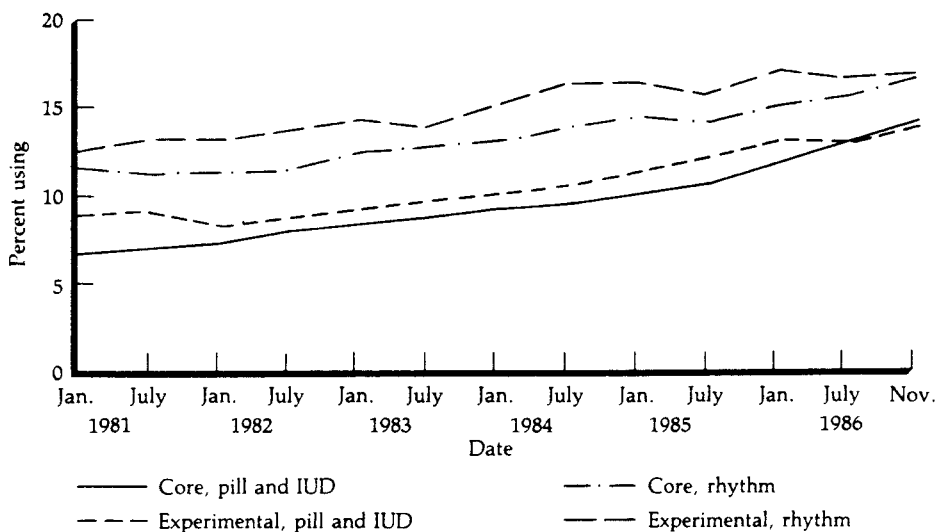
‡This percent would increase to only 28.5 if reports of second methods within closed intervals were attributed to all remaining months of nonuse within these intervals.

Overall, the above comparisons suggest that reports of prior contraceptive use are considerably more complete in the experimental survey than in the standard DHS survey. It is important to determine the ways in which the questionnaire design of the experimental survey improved the reporting of contraceptive use. Undoubtedly, one very important advantage of the calendar was that it allowed for reports of multiple segments of use within an interval. Even though the overall level of contraceptive use is relatively modest in Peru (particularly in contrast to levels in other Latin American countries), a substantial proportion of women use more than one method within a birth interval. For example, in the experimental survey, approximately 20 percent of users report use of more than one method in the open interval and about 15 percent in closed intervals.

Explaining the Differences in Prevalence Estimates

To what extent are these shortcomings of the core questionnaire a consequence of the fact that it ob-

Figure 2 Use of modern contraceptive methods and rhythm among ever-married women aged 15-44, as reported in the core versus experimental questionnaire, Peru, 1986



tained very incomplete information with regard to multiple use within an interval? We attempted to answer this question in two ways: First, we looked at responses in the standard survey on the penultimate method in each closed interval—recall that no information was obtained with regard to the duration of use of these methods. We modified our simulated calendar from the core questionnaire to recode all intervals in which respondents reported use of a previous method: essentially, all remaining months of non-use were altered to be months of use of the previous method. The net effects on estimates of prevalence for the period 1981 to 1986 turned out to be trivial. This surprising result is most likely due to a combination of errors in the core, which involve overestimates of the duration of use of the last method as well as possible misreports of the length of the birth interval.

Second, we created a trial calculation from the experimental survey, which eliminated all but the last reported segment of use within each birth interval (with the exception of allowing two segments of use for current users). In other words, we created a calendar from the experimental survey that replicated the type of information collected (without a calendar) in the core. Estimates of prevalence for 1981 to 1986 from this altered calendar (see Figure 1) indicate that reports of multiple use in the experimental survey account for about half of the difference between it and the core survey. These results suggest that, although substantial switching of methods occurs within birth intervals in Peru, modifying the core questionnaire to include reported durations for two methods per interval would not be an effective way to improve use estimates.¹⁷ This analysis cannot be used to determine whether a more elaborate restructuring and enhancement of the core questionnaire would yield estimates as good as, or better than, the calendar design.

What explains the remaining differences between estimates of prevalence from the two surveys? Part of the difference is accounted for by slightly higher proportions of intervals with any use in the experimental survey and part is accounted for by slightly higher durations of use of the last method in closed birth intervals.¹⁸ The net effect of these differences is that 28.9 percent of all months in the simulated calendar from the core survey are coded with use as compared with 32.7 percent of all months in the experimental calendar.¹⁹

How large is the impact of these differences in reported prevalence on estimates of contraceptive failure and discontinuation? An analysis presented elsewhere (Goldman et al., 1988b) indicates that, in general, omission of use in the standard survey results in

estimates of failure that are somewhat too high. However, the order of magnitude of the differences is relatively modest: for example, first year use-failure rates for all methods combined are 19.2 and 16.2 per 100 women as derived from the core and the experimental questionnaires, respectively; the corresponding estimates equal 26.1 and 18.7 for the rhythm method and 7.0 and 6.3 for the pill.²⁰ Although these differences are not trivial, the ordering of methods by efficacy is the same from the two surveys and both sets of estimates indicate the expected differentials by age and other covariates. In contrast, estimates of discontinuation are substantially higher from the experimental questionnaire. For example, the percent of women discontinuing use within the first year (for reasons other than failure or a planned pregnancy) equal 19.1 and 34.2 from the core and the experimental questionnaires, respectively. Undoubtedly, the fact that interviewers were able to collect multiple segments of use within each birth interval led to the higher rates from the experimental survey.

Conclusions

In summary, the analysis described above has demonstrated that although reports of contraceptive knowledge, ever-use, and current use are relatively robust to the variations in questionnaire design employed in the DHS surveys, estimates of past use are dependent on the survey instrument. Several different comparisons have indicated that reporting of information on contraceptive histories in the experimental survey is superior to that in the core. For example, reported durations of use are not heaped, estimates of prevalence for dates prior to the survey consistently exceed those from the core, and estimates of prevalence for 1981 are considerably closer to those reported in the CPS than are estimates derived from the core. In fact, for certain modern methods, estimates derived from the experimental calendar are in agreement with those reported in the CPS. The fact that the calendar easily incorporates multiple segments of use within an interval, and allows the interviewer to reconcile dates of use with other events, particularly pregnancies, is in large part responsible for the more complete reporting of contraceptive use in this survey.

There are certain advantages to the experimental calendar that are not evident from the analysis presented above. The fact that all dates of pregnancy and use were entered into the same column of the calendar eliminates the possibilities of many types of potential inconsistencies. These errors could and did occur in the core questionnaire. For example, nearly 20 percent

of closed intervals with reported use had a reported duration of use of the last method that exceeded the length of the interval; about one-third of these exceeded the length of the interval by three months or more and over 10 percent exceeded the interval by at least one year. Not infrequently, women also reported using a previous method for these same intervals in which they reported a duration of use of the last method exceeding the length of the interval.

Another advantage of the experimental questionnaire is that interviewers were instructed not to leave any months of the first column of the calendar without a code. In fact, all questionnaires in the experimental survey were complete in this regard. By contrast, although the core DHS survey had a relatively high response rate for most questions, responses are missing that are relevant for this analysis. For example, 2 percent of closed intervals with reported use have missing information either on the method used or on the duration of use.

Yet a third advantage of the experimental questionnaire is that interviewers appeared to prefer it to the standard questionnaire. Although interviewers initially were more intimidated by the lack of structure in the experimental questionnaire, after a short period of training they preferred the calendar because it allowed them to reconcile the timing of different events and to probe for information.

Overall, the results presented here suggest that both the experimental and the standard DHS surveys in Peru obtained reasonably accurate reports of contraceptive use. To the extent that the analyst is interested in current status measures of contraceptive use, or even period-based estimates of contraceptive failure, only modest differences exist between the two survey instruments. The major advantages of the calendar for the analysis of information on contraception are twofold: (1) it obtains more complete reports of use for periods prior to the survey, an improvement that has obvious implications for estimates of trends in contraceptive prevalence and estimates of discontinuation; and (2) it obtains information that is more complete and internally consistent with other types of information. Also, the costs of including a calendar appear to be small: on average, interviewers prefer it and it seems to increase the interview time by only a small amount. As a result of the experimental field test in Peru, the DHS project is planning to include a calendar in its second round of surveys, for countries with a moderate prevalence of contraception. Whether use of a calendar would be advantageous or disastrous in societies with a low level of use and with little familiarity with calendar dates is unclear.

Acknowledgments

The authors are grateful to the staff of the *Instituto Nacional de Estadística*, Peru, for facilitating the completion of the Demographic and Health Surveys, and especially for their collaboration in carrying out the experimental survey. The authors would also like to acknowledge Ozer Babakol, who provided invaluable assistance with the computer programming, and Jerry Sullivan, James Trussell, Anne Pebley, and the anonymous reviewers for their comments on an earlier draft of the manuscript. This research was funded by NICHD Grant 1 R01 HD22417 and the Institute for Resource Development, Inc., Westinghouse through its contract with the US Agency for International Development.

Notes

- 1 For example, the majority of World Fertility Surveys that contained a fertility regulation module determined the last method used in the open and the last closed interval, but did not obtain information about the duration of use.
- 2 The World Fertility Surveys in Venezuela and the Dominican Republic also included a sexual activity table.
- 3 A similar experimental field evaluation was undertaken in the Dominican Republic in 1986, but the resulting data have not yet been analyzed.
- 4 Analyses based on other experimental subjects such as migration, fertility, and infant health are presented elsewhere (for example, Goldman et al., 1988a; Goldman et al., 1989; Moreno and White, 1989; Moreno, 1988).
- 5 In both questionnaires, the first section concerns the respondent's background and the second section contains birth history information. The core questionnaire contains a full birth history, whereas the experimental questionnaire records all births since January 1981, and one prior birth.
- 6 There are, however, two additional minor differences with regard to information on ever-use. In the standard questionnaire, the question reads "Have you ever used...?", whereas in the experimental questionnaire the question reads "Have you or your partner ever used...?" In addition, only the experimental questionnaire contains a probe question to determine if women who did not acknowledge having used any of the methods described by the interviewer had really never done so.
- 7 As with the question on ever-use, the experimental survey, but not the core survey, refers to the woman's partner: "Are you or your partner currently doing something or using any method to avoid getting pregnant?" Both surveys have separate questions for sterilized couples; that is, the date of sterilization is obtained separately from information on the duration of use of the current method.
- 8 Both questionnaires collected information on reasons for termination of use—that is, whether the use resulted in a pregnancy, whether the woman stopped using in order to become pregnant, or whether the method was discontinued for another reason. These data, which are essential for the estimation of failure and discontinuation rates, are described elsewhere (Goldman et al., 1988b).
- 9 This agreement occurred in spite of the fact that only the experimental questionnaire contained a probe for those women who did not acknowledge use of any of the specific methods. A total of 79 women responded positively to this question. If we were to classify these women as nonusers, the estimate of ever-use from the experimental survey would drop from 63.4 percent to 59.0

percent, a value that would be significantly lower than that from the core. This comparison suggests that the inclusion of a probe question for ever-use had a substantial impact on the resulting estimate and that the experimental questionnaire would probably have obtained a lower estimate of ever-use than the core without the probe. It could be the case, however, that interviewers, who were trained with both questionnaires, automatically used a probe question in the core questionnaire. On several occasions in the training session, interviewers admitted to using probe questions from the experimental questionnaire in the core questionnaire so as to improve respondents' recall on various subjects. Such behavior would, of course, weaken the true differences in response between the two questionnaires.

- 10 This agreement is not surprising since the only difference between the questions on current use is a reference to the partner in the experimental questionnaire. In general, estimates of current use of contraception seem to be robust to the specific wording of the question (see, for example, Anderson and Cleland, 1984).
- 11 This information was not used in the calculation of current use in Table 3.
- 12 In the core survey, the question reads: "For how long have you been using (CURRENT METHOD) continuously?" In the experimental survey, "long" is replaced by "many months."
- 13 The fact that the extent of heaping is less for respondents with more education suggests that the less heaped responses are more accurate.
- 14 The basic logic of the imputation scheme was to attribute all months within the interval to contraceptive use if the reported duration of use was equal to or greater than the length of the interval. If the segment of use was shorter than the interval, we randomly imputed the starting date of use according to the notion that all possible dates (conditional on the reported duration of use) were equally likely.
- 15 Only in rare circumstances did we impute missing durations of use. For example, out of a total of 1,207 last segments of use within closed intervals, 15 segments had no reported duration. Of these, 5 were reclassified as nonuse and the remaining 10 were given an imputed duration; see Goldman et al. (1989) for details.
- 16 The percent of women using a contraceptive method was reconstructed for January and July of each calendar year between 1981 and 1986. The values for interview date (plotted as October 1986) differ from those presented for current use in Table 3 for three reasons: first, the numbers in Table 3 are for currently married rather than ever-married women; second, estimates in Figure 1 (which are derived from the simulated calendar) are based on responses to the question on current method and to the question on other methods used regularly during the same month (see discussion in the previous section); and third, responses to questions on whether the respondent is currently using a method in the experimental questionnaire are not entirely consistent with the codes entered in the interview month in the calendar. Note that some of the latter inconsistencies (all of which are due to reported use in one case and nonuse in the other) may be real: that is, women may not be currently using but may have used a method at some earlier time in the month. These discrepancies highlight the ambiguity of the concept of "current use."
- 17 Our simulations indicate that this is the case with regard to closed intervals. An additional question on duration of use of the penultimate method in the open interval (for women not currently using a method) might well have a substantial effect on the resulting estimate of prevalence.

18 For example, among intervals that began subsequent to January 1981, 52.9 percent and 51.8 percent of open intervals, and 30.7 percent and 29.9 percent of closed intervals, were reported with use in the experimental and core surveys, respectively. The mean duration of use of the last method used in closed intervals was 13.8 months and 12.2 months in the experimental and core surveys, respectively.

19 These estimates are based on ever-married women.

20 These estimates include only live births since nonlive birth information was not collected from the standard survey. Estimates for all methods combined and rhythm are significantly different between surveys at a 1 percent level.

References

- Anderson, John and John G. Cleland. 1984. "The World Fertility Survey and Contraceptive Prevalence Surveys: A comparison of substantive results." *Studies in Family Planning* 15, 1: 1-13.
- Gaslonde, S. and E. Carrasco. 1982. "The impact of some intermediate variables on fertility: Evidence from the Venezuela National Fertility Survey, 1977." *World Fertility Survey Occasional Papers No. 23*. Voorburg: International Statistical Institute.
- Goldman, N., C.F. Westoff, and L. Moreno. 1988a. "Full vs. five-year maternity history data for fertility and child mortality measures." Paper presented at the International Union for the Scientific Study of Population (IUSSP) Seminar on Collection and Processing of Demographic Data in Latin America, 23-27 May, Santiago, Chile.
- Goldman, N., L. Moreno, C.F. Westoff, and B. Vaughan. 1988b. "Estimates of contraceptive failure and discontinuation based on two methods of contraceptive data collection in Peru." Paper prepared for the United Nations Expert Group Meeting on Methodologies for Measuring Contraceptive Use Dynamics, New York.
- Goldman, N., L. Moreno, and C.F. Westoff. 1989. *Peru Experimental Study: An Evaluation of Fertility and Child Health Information, Final Report*. Columbia: Demographic and Health Surveys, Institute for Resource Development, Inc., Westinghouse.
- Goldman, N., S. Rutstein, and S. Singh. 1985. "Assessment of the quality of data in 41 WFS surveys. A comparative approach." *World Fertility Survey Comparative Studies No. 44*. Voorburg: International Statistical Institute.
- Jejeebhoy, S. 1988. "Measuring contraceptive use failure and continuation: An overview of new approaches." Background paper for the United Nations Expert Group Meeting on Methodologies for Measuring Contraceptive Use Dynamics, New York, 5-7 December 1988.
- Laing, John E. 1984. "Measurement of contraceptive protection for fertility analysis." Paper prepared for the Seminar on Integrating Proximate Determinants into the Analysis of Fertility Levels and Trends, International Union for the Scientific Study of Population and World Fertility Survey, London.
- Moreno, L. 1988. "A comparison of the effects of demographic and health covariates on child mortality from two surveys in Peru." Paper prepared for the Annual Meeting of the Population Association of America, 29-31 March 1989, Baltimore, Maryland.
- Moreno, L. and M. White. 1989. "The use of a calendar to collect migration data." Paper prepared for the International Union for the Scientific Study of Population (IUSSP) General Conference, 20-27 September, New Delhi.
- Pebly, A.R., N. Goldman, and M.K. Choe. 1986. "Evaluation of

contraceptive history data in the Republic of Korea." *Studies in Family Planning* 17, 1: 22-35.

Population Information Program. 1985. "Fertility and family planning surveys: An update." *Population Reports, Series M, No. 8:*

Special Topics (September/October).

United Nations. 1987. "A comparative evaluation of data quality in thirty-eight World Fertility Surveys." ST/ESA/SER.R/50. New York: United Nations.

Appendix

Figure A1 Calendar used for the experimental questionnaire in the Peru 1986 Demographic and Health Survey

INSTRUCTIONS: BEGIN COLLECTING INFORMATION FOR MONTH OF INTERVIEW. ONLY ONE CODE SHOULD APPEAR IN ANY BOX. FOR COLUMNS 1 AND 6 ALL MONTHS SHOULD BE FILLED IN.

INFORMATION TO BE CODED IN EACH COLUMN

Col. 1: Births, "Other" Pregnancies, Contraceptive Use

0 NO METHOD
1 PILL
2 IUD
3 INJECTIONS
4 VAGINAL METHODS
5 CONDOM
6 FEMALE STERILIZATION
7 MALE STERILIZATION
8 RHYTHM: CALENDAR
9 RHYTHM: BODY TEMPERATURE
10 RHYTHM: CERVICAL MUCUS
11 RHYTHM: TEMPERATURE AND MUCUS
12 WITHDRAWAL
13 RHYTHM AND CONDOM
14 RHYTHM AND WITHDRAWAL
15 CONDOM AND WITHDRAWAL
16 OTHER

Col. 1A: Discontinuation of Contraceptive Use
1 BECAME PREGNANT WHILE USING
2 WANTED TO BECOME PREGNANT
3 OTHER REASON

Col. 2: Breastfeeding
1 BREASTFEEDING

Col. 3: Post-partum Amenorrhea
0 PERIOD DID NOT RETURN

Col. 4: Post-partum Abstinence
0 NO SEXUAL RELATIONS

Col. 5: Marriage/Union
1 IN UNION (MARRIAGE OR LIVING TOGETHER)

Col. 6: Moves and Places of Residence
0 CHANGE OF RESIDENCE
1 COUNTRYSIDE
2 TOWN
3 CITY

Col. 7: Type of Employment
1 SELF-EMPLOYED
2 WORK FOR FAMILY MEMBER
3 WORK FOR OTHERS

	1	1A	2	3	4	5	6	7
DEC								
NOV								
OCT								
SEPT								
1 AUG								
9 JUL								
8 JUN								
6 MAY								
APR								
MAR								
FEB								
JAN								
DEC								
NOV								
OCT								
SEPT								
1 AUG								
9 JUL								
8 JUN								
5 MAY								
APR								
MAR								
FEB								
JAN								
DEC								
NOV								
OCT								
SEPT								
1 AUG								
9 JUL								
8 JUN								
4 MAY								
APR								
MAR								
FEB								
JAN								
DEC								
NOV								
OCT								
SEPT								
1 AUG								
9 JUL								
8 JUN								
3 MAY								
APR								
MAR								
FEB								
JAN								
DEC								
NOV								
OCT								
SEPT								
1 AUG								
9 JUL								
8 JUN								
1 MAY								
APR								
MAR								
FEB								
JAN								

Wyclwood 57 976

School

Rye St Anthony

62802

9.45