

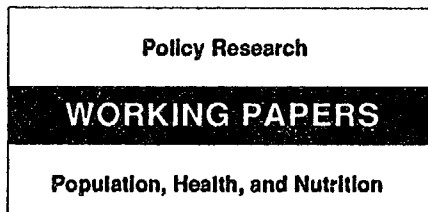
WPS1048

Population and Human Resources
Department
The World Bank
November 1992
WPS 1048

Targets and Indicators in World Bank Population Projects

George Baldwin

The Bank should strengthen its use of international comparisons and trend analysis rather than increasing its use of target setting; the use of demographic and health surveys should be the rule, not the exception, in Bank population and health projects; and more attention should be given to program-level than to project-level performance.



WPS 1048

This paper — a product of the Population Policy and Advisory Service, Population and Human Resources Department — is part of a larger effort in the department to assess the performance of family planning programs in developing countries. Copies of the paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Otilia Nadora, room S6-065, extension 31091 (November 1992, 57 pages).

In reviewing World Bank evaluations of the impact of population projects, Baldwin explains the nature and uses of four families of performance indicators. Two measure inputs:

Project implementation indicators, which are project-specific, are the principal measures used in Bank supervision. They measure success in creating sources capable of conducting certain desired activities.

Process (or activity) indicators measure performance of a project's intended activities but tell nothing about the "yield" or output of those activities.

And two measure output:

Performance (or intermediate-output) indicators measure the yield or output-performance of a project or program. For family planning, the principal indicator in this category is acceptor figures, normally with details about methods used plus the age, parity, and geographical distribution of acceptors. Quality may or may not be good and coverage may or may not be comprehensive. These indicators do not directly measure ultimate demographic impacts — lower fertility and slower population growth.

Demographic outcome (or impact) indicators do measure demographic impacts, usually the contraceptive prevalence rate and age-specific and total fertility rates.

One can use a desired value of any comparator as a target, but a target is only one possible

comparator. Two more widely used comparators for family planning are *trends* (comparing current with past performance) and *international performance* (an external comparison). Baldwin recommends strengthening the Bank's use of world ("successful developing country") standards and of trend analysis rather than increasing its use of target setting.

The Bank's primary interest is normally the performance of the borrower's national program, so more attention should be given to program-level than to project-level performance — except for pilot projects.

Baldwin recommends:

- That the Bank standardize its terminology about these four families of indicators.
- That the Population and Human Resources Department periodically prepare comparator tables and graphs for use in Bank project and sector reports.
- That the Bank discontinue Project Performance Audit Reports on population projects, as they seldom add much to information and judgments contained in Project Completion Reports. The money saved could be applied to more effective evaluation research.
- That operational staff show more concern for a program's contraceptive mix.
- That more attention be paid to a program's service quality.
- That the use of demographic and health surveys be the rule, not the exception, in Bank population and health projects.

The Policy Research Working Paper Series disseminates the findings of work under way in the Bank. An objective of the series is to get these findings out quickly, even if presentations are less than fully polished. The findings, interpretations, and conclusions in these papers do not necessarily represent official Bank policy.

Targets and Indicators in World Bank Population Projects

by
George B. Baldwin

Population and Human Resources Department
The World Bank

Contents

Acknowledgements	v
Executive Summary	1
I. Introduction	4
II. Targets and Indicators	4
The Universe of Projects	4
III. Principal Indicators for Measuring Population Performance	6
IV. Comparators: Projections, Targets, Trends and "Other Places"	9
V. The Use of Targets in Bank Projects: The Record	10
VI. Bank Instruments Used for Evaluating Its Projects	13
Supervision Reports	13
Mid-Term Project Reviews	14
Operational Research Studies	14
Repeater Projects	14
Project Completion Reports (PCRs)	15
Project Performance Audit Reports (PPARs)	15
Non-Bank External Reviews	15
VII. Three Perspectives from Outside Bank Operations	15
The 1985 Report by Simmons and Maru	16
The 1991 OED Report	16
Major Forthcoming Study Sponsored By USAID	18
VIII. Conclusions and Recommendations	18
A Word Of Caution	18
Past Use Of Targets and Indicators	19
A Simple Model Suggests Four Sets of Indicators	21
At Project Level, Emphasize Activity Indicators, Not Outcomes	23
A New And Difficult Dimension: The Quality Of Service	24
Six Recommendations	25
Annex I. PHN Projects with Major Population Objectives (1979-91)	26
Annex II. Synopses of "Targeting" in Past Reports (1970-91)	29
Bangladesh (1975-91)	29
Botswana (1984)	30
Brazil (1986 & 1989)	30
Burkina (1985)	31
Burundi (1987)	31
Colombia (1985)	31
Comoros (1983)	31
Dominican Republic (1976)	32
Egypt (1973 & 1978)	32

Ethiopia (1988)	33
Ghana (1985 & 1990)	33
Guinea-Bissau (1987)	33
Haiti (1989)	33
India (1972-1990)	34
Indonesia (1972-1991)	35
Iran (1973)	37
Ivory Coast (1985)	37
Jamaica (1970-87)	37
Jordan (1985)	38
Kenya (1974-90)	38
Korea (1979)	39
Lesotho (1985 & 1989)	40
Madagascar (1988)	40
Malawi (1983-1991)	41
Malaysia (1972 & 1978)	41
Mali (1983 & 1991)	42
Morocco (1985)	42
Niger (1986)	43
Nigeria (1985, 1988, & 1991)	43
Pakistan (1983)	44
Peru (1982)	44
Philippines (1974 & 1979)	44
Rwanda (1986 & 1991)	45
Senegal (1982 & 1991)	45
Sierra Leone (1986)	46
Sri Lanka (1988)	46
Thailand (1978)	46
Togo (1991)	47
Trinidad and Tobago (1971)	47
Tunisia (1971 & 1991)	48
Yemen Arab Republic (1990)	49
Zimbabwe (1986 & 1991)	49

Annex III. Treatment of Targets and Indicators in PCRs and PPARs	50
Bangladesh (PAR # 6303, June 1986)	51
Dominican Republic (PAR # 5596, April 1985)	52
Egypt (PAR # 4144, October 1982)	52
India (PAR # 3748, January 1982; PAR # 8896, June 1990)	52
Indonesia (PAR # 5675, May 1985; PAR # 6276, June 1986)	53
Jamaica (PAR # 2580, June 1979; PAR # 5589, April 1985)	53
Kenya (PAR # 3536, June 1981)	55
Korea (PAR # 8895, June 1990)	55
Malaysia (PAR # 5641, May, 1985)	55
Philippines (PAR # 5544, March 1985; PAR # 9380, February 1991)	56
Thailand (PAR # 7335, June 1988)	56
Trinidad and Tobago (PAR # 3687, December 1981)	56
Tunisia (PAR 5135, June 1984)	57

Acknowledgements

I was invited to prepare this review of how the Bank has evaluated population projects and programs largely because I had served for seven years in the Population Projects Department, beginning with the Department's establishment in 1970. The review is based, however, not on that experience but on a close reading of the more than 70 project appraisal reports that had been written between 1970 and mid-1991 when the study began. The task of identifying and assembling these reports was given to Mrs. My Thi Vu, Research Assistant in the Population, Health and Nutrition Division. No one could have rounded up these materials more rapidly and thoroughly than she did. I am much in her debt.

For helpful comments on my initial draft I should like to thank the following World Bank staff members: Bruce Carlson, Althea Hill, Ann Levin, Ok Pannenberg and Juliana Weissman. From outside the Bank I received very helpful comments from Ms. Jean van der Tak, retired staff member of the Population Reference Bureau; Drs. Amy Ong Tsui and James E. Veney of The Evaluation Project at the Carolina Population Center; and from Martin Vaessen, Director of the Institute for Resource Development, Inc., in Columbia, Maryland.

George B. Baldwin

Executive Summary

This paper reviews how the Bank has evaluated the impact of its population projects, with specific reference to its use of targets as a performance standard. The review is based on a reading of the appraisal reports for all 75 population projects since the Bank entered the field in 1970, the 17 available Project Performance Audit Reports (PPARs) on completed projects, external reviews of the Bank's approach to performance evaluation, plus one authoritative outside commentary on the subject. This documentary review has been supplemented by talks with Bank and outside population experts.

The review consists of three parts: (1) a main text of 22 pages that discusses the nature of targets and indicators, provides a conceptual framework, and offers a set of six recommendations; (2) a 21-page annex (Annex II) that summarizes the use of targets and impact indicators in each of the SARs; and (3) a seven-page summary of what the more important PPARs have had to say on performance monitoring.

A simple flow-diagram model provides a conceptual framework for identifying and labeling four families of performance indicators. Two measure inputs, two outputs. The suggested categories are:

Inputs

- a. Project implementation indicators
- b. Process, or activity, indicators

Outputs

- c. Performance, or intermediate-output, indicators, and
- d. Demographic outcome, or impact, indicators

Project implementation indicators are project-specific, self-contained within project boundaries, and are the principal measures used in Bank supervision work. Project cost and implementation schedules constitute the main implementation targets, and are often supplemented by other implementation activities (e.g., recruitment and training of manpower). For our purposes, implementation indicators are the least important of the four types: they measure success in creating new resources capable of conducting certain desired activities.

Process indicators measure the performance of a project's intended activities (clinic sessions held, patients attending, contraceptive supplies distributed, training courses conducted, broadcasts aired, film-showings presented, etc.). Process indicators tell us nothing about the "yield" or output of those activities.

Performance indicators measure the yield or output-performance of a FP project or program. The principal indicator in this category is acceptor figures, normally with details of methods used, plus the age, parity, and geographical distribution of acceptors. These figures come from program service statistics. The quality may or may not be good and the coverage may or may not be comprehensive (e.g., is the private sector included?). A key characteristic of intermediate-output indicators is that they provide no direct measurement of ultimate demographic impacts, i.e., lower fertility and slower population growth. For that we need an additional set of outcome

indicators ("impact indicators"); the main ones are the Contraceptive Prevalence Rate¹ and the Age-Specific and Total Fertility Rates. The Natural Rate of Increase and the rate of population growth are bottom-line measures of demographic outcomes, but since they depend on death rates as much as on birth rates they cannot be used as indicators of reproductive behaviour and outcomes.

It is possible to use a desired value of any indicator as a target. But a target is only one possible comparator. Two more widely used FP comparators are trends (comparison with own past performance) and international performance (an external comparison). Indeed, where targets are used they are invariably set with reference to either or both of those two comparators. Thus it is the latter which are the "primary reference points" for setting targets and judging performance in any FP program. Indeed, it is the development of international standards of fertility and contraceptive use, and historical rates of movement towards desired levels of these, that underlie all impact performance judgments. Trend analysis simply measures how successfully, or rapidly, one is approaching international standards. It is for these reasons that the review's recommendations focus on strengthening the use of world (=successful LDC) standards and of trend analysis instead of making greater use of target-setting.

Since the Bank's primary interest will normally be in the performance of the borrower's national program, much more attention should normally be given to program-level performance than to project-level performance. Trying to make the project the primary interest makes sense only if a project can be regarded as an experimental pilot project whose results will need evaluation to see if the project design should be generalized within the national program, or used elsewhere. Few population projects are of this type (mainly the minority of geographically distinct, or area, projects). The activities of most are so commingled with a national program that it is hopeless to try to link project inputs to project outputs; we are normally linking project inputs to program outputs. So while project-based "process indicators" may be feasible, indicators of intermediate and final-outcome indicators will usually be feasible only at program level.

The review finds that the Bank has done a generally good job in using intermediate output targets, often as a technical assistance service to help governments set realistic program targets. Where targets and quantitative indicators have not been used, there is usually a good reason. A caution is stated that performance against project output targets, while sometimes useful, should not be expected to carry the main burden of assessing project success in this sector.

The review finds Bank practice in using targets and indicators sensible and pragmatic, although not immune from some of the confusion and inconsistency that marks the handling of these issues in all population agencies. In addition to suggesting the four families of performance indicators, the review makes six recommendations for improving Bank practice in evaluating population operations. They are:

- a. An effort should be made to standardize Bank terminology, using the flow-diagram model and the four families of indicators derived from it.

¹Strictly speaking, this should be classed with the "performance" or intermediate-output group. However, a robust CPR figure provides such strong evidence of fertility impact that it can be regarded "as if" it were a direct fertility measure. The survey data on which it rests are much closer to those used in establishing fertility estimates than to those used for estimating intermediate outputs (mainly service statistics).

- b. PHN should prepare, periodically, a set of comparator graphs and tables for use in Bank project and sector reports.
- c. Operational staff should show stronger concern for a program's contraceptive mix, both at the stage of project design and in assessing program performance.
- d. More attention should be devoted to measuring a program's service quality, a dimension that has recently begun to receive increased attention by population professionals.
- e. The use of Demographic and Health Surveys (DHS) should be the rule, not the exception, in Bank population and health projects.
- f. Since few PPARs add much to the information and judgments contained in Project Completion Reports (PCRs), the Bank should consider discontinuing them for population projects. The resulting savings could be used for more productive forms of evaluation research.

I. Introduction

This paper reviews past and current Bank practice in "setting targets and evaluating the impact of Bank population projects." The purpose is to provide background "for an eventual experiment with introducing specific impact indicators into one or more Bank projects." The review, it is hoped, may lead to recommendations "for improving target setting and impact evaluation, drawing . . . on Bank staff suggestions, experience in other agencies, and the research literature in the area."

The review combines quantitative and qualitative data in ways which seem most responsive to PHN's main interests in targets and *ex post* project evaluations. With PHN agreement, I have recast the assignment's formal Terms of Reference into the following three tasks:

- ... Tell us what we have done in the past,
- ... Tell us what you think of past practice, and
- ... Tell us what you think we should do in the future.

This approach explains the inclusion of Annexes II and III (summaries of how targets and indicators have been used in Bank reports, and of internal critical reviews by the Operations Evaluation Department) as the main evidence on which the discussion rests. This documentary review has been supplemented by discussions with experienced population professionals in the Bank, plus a few outsiders currently conducting similar reviews.

II. Targets and Indicators

It is useful to maintain a distinction between "targets" and "indicators."² Operational Directive 10.70 has a suggestive footnote which defines an indicator as "a measure . . . which indicates movement towards, or away from, an agreed project target." This need not be the case, however. An indicator can better be regarded as a measure of performance to be used against some comparator—not necessarily a target. An equally common comparator is past performance, i.e. the use of indicators to reveal trends. Another common comparator is the level of selected indicators in other countries, i.e. international performance standards. Without good indicators, no comparisons are possible; with them, many kinds are possible. To focus on targets is to limit attention to a narrower set of problems than is covered by indicators. This paper is concerned with both.

The Universe of Projects

Since entering the population field in 1970 the Bank has made 75 loans or credits, to 42 countries. Annex I lists these operations alphabetically by country, giving the project name as shown on the Staff Appraisal Report (SAR), the date, and the size of the loan/credit. The list does not include all PHN projects since the Bank entered this field: there have been 40-50 Health and Nutrition projects that have not had demographic or family planning as an explicit objective;² such projects have been excluded. The 75 projects which have had Population and family planning as a primary, explicit objective constitute something under two-thirds of all PHN projects.

²Note also the distinction between "targets" and "targeting", the former referring to the setting of numerical performance goals, the latter to specific groups or areas to be served. This paper is concerned with target-setting, not targeting.

There is no pattern or consistency to the names given to Population projects. Some are indeed called "Population Projects"; others include "Health" or "Family Welfare" or some other term in the title. The word "Population" appears in the title of 53 of the 75 ARs but in half the cases the word is accompanied by one or more other descriptors, e.g., "Health," "Family Welfare," etc. These variations have little significance and can be ignored. They reflect the fact that Governments (or Bank staff) in countries where Population is a particularly sensitive topic often prefer a more diplomatic project title, or the fact that Population objectives are normally pursued, in large part, by offering family planning services through the health-delivery system, especially the Maternal and Child Health (MCH) part of the system. Since Population is only one among many objectives of any health system, a project that is not a single-purpose Population project must deal with multiple objectives and will have multiple achievement criteria, indicators, or targets. This does not present any particular difficulty in reviewing performance on the population objective, however.

Table 1 shows the number of projects approved³ during each five-year period since 1970:

**Table 1. Number of Projects
per Quinquennium**

Years	Number
1970-74	10
1975-79	10
1980-84	11
1985-89	30
1990 to 5/91	14
Total	75

There was a major acceleration in Population projects after 1984. Since the typical project lasts about five or six years, and it takes another year or more to write a Project Completion Report (PCR), we can expect to find PCRs only from the set of 31 projects approved in 1984 or earlier. PCRs (which are required of all projects) and Project Performance Audit Reports (PPARs, done only on a small sample of projects) are not the only source of information on project performance, however.⁴ Equally or more informative, usually, are comments on program performance to be found in repeater projects.

The Bank has done one or more repeater projects in 18 countries. The 30 repeaters, plus their 18 initial projects, give us 48 projects that occur in some kind of sequence. This is almost two-thirds of the total. In 11 of the 18 countries the sequence consists of only one initial and one repeater project; but in seven countries there have been three or more projects (India has had seven, Indonesia five, Bangladesh and Kenya four, and Jamaica, Malawi, and Nigeria three each). In 15 of the 18 repeater countries there is a current project under implementation; in only three countries has a repeater sequence been abandoned so that no continuing project exists (Egypt, Malaysia, and

³Strictly speaking, classification is based on the dates appearing on gray-cover SARs.

⁴The latest OED catalogue of OED reports (September, 1991) lists 14 Project Completion Reports (PCRs) and 17 Performance Audit Reports (PARs) for the population sector. The listed PCRs mean that no PARs were done on those projects; the listed PARs normally attach the relevant PCR as an annex.

the Philippines). In two countries with repeater projects (Jamaica and Tunisia) there was a clear interruption in the sequence; but in the other 16 repeater countries the sequences appear to have been unbroken and to be continuing to the present. Nine of the 31 projects signed before 1984 turned out to be "one off" or non-repeated projects, presumably reflecting either Government or Bank reluctance to go forward with a second project. (Most single projects signed in 1985 or later are still disbursing and little can be said about repeater projects among them.)

III. Principal Indicators for Measuring Population Performance

The Bank's main objective in Population has always been to slow its growth so per capita incomes can grow faster, without the drag of a high population growth that eats up much of a country's growth in GNP. Bringing the personal benefits of family planning (FP) to individual couples and their children has not been the primary objective. Thus the main test of project success is its contribution to slowing population growth. That, however, must be a long-term measure, revealed by censuses or sample surveys. In the shorter term, population growth may actually accelerate, even in the face of a good population program. This is because mortality rates may drop faster than birth rates, depending on where a country is in the demographic transition (the historical, world-wide transition from a high birth-rate/high death-rate pattern to a low birth-rate/low death-rate pattern). So while a falling population growth rate is certainly encouraging, a stable or even accelerating rate is not necessarily a mark of failure.

A more revealing demographic indicator of program success is what is happening to fertility (i.e., forget about the death rate). If fertility is declining, the program must be succeeding (this assumes that program-offered FP and motivation activities are a major source of falling fertility, as they normally are). There are four fertility measures commonly used and frequently encountered in Bank SARs:

The Total Fertility Rate (TFR)
 The Crude Birth Rate (CBR)
 Age-Specific Fertility Rates (ASFR), and
 The Net Reproduction Rate (NRR)

These terms are explained in the summary definitions on the inside cover of this review. Census analyses can provide numbers for all four of these indicators; but censuses are infrequent, analyses are often delayed, and census quality not always reliable. So better numbers are often provided by special sample surveys. The key indicator is the TFR, which tells us how many births a woman will have over her lifetime if she follows the reproductive experience which women in defined age-groups are currently experiencing. Thus if a survey shows that women 15-19, 20-24, etc. are having, on average, x, y, etc. number of children, then it is easy to aggregate the age-specific births to see how many an average woman will have throughout her reproductive years. Advocates of "zero population growth" want to see the TFR fall to about 2.1: this implies (with some allowance for non-marriage) that each woman will have one daughter, i.e., will just reproduce her own reproductive capacity, over her lifetime. That represents a Net Reproduction Rate of 1.0, a steady-state demographic condition (population does not stop growing when a NRR of 1.0 is reached, since increasing numbers of women will be entering their reproductive years. Only when this historical momentum has been worked off will population growth gradually fall to zero).

TFRs can be derived only from ASFR data, actual or assumed. Since only actual trends are of interest for judging program performance, such indicators become available only as often as surveys can be conducted. These are rarely done more frequently than every few years, at best (they are rarely conducted on any regular basis—usually when some donor presses for them and offers to fund them). At best, a few years must pass between successive calculations of the TFR; at worst, we get one figure and never a second or a third or a fourth.

So demographic indicators are too spread out in time to provide current information on how people are responding to population programs. For that we must turn to data on program "acceptors." If we have reasonable detail on FP users, it is possible to convert such figures into estimates of the demographic measures reviewed above. Thus FP acceptor data provide much more current and useful data on program performance than are available from demographic surveys and censuses. They are also the main source of data needed to "fine tune" any program.

The principal "acceptor" indicators are these:

- a. The total number of current acceptors (=users of FP) for a government program, outside the government program, by contraceptive method
- b. New acceptors (e.g., per month, per year) government plus non-government, by method
- c. Continuation rates (or their complement, dropout rates: 100 acceptors minus dropouts = continuing acceptor rate).
- d. The contraceptive prevalence rate (CPR)⁵

Acceptor data come from the service statistics of FP clinics and hospitals. They should be reflected in individual patient records, so that special analyses could generate figures by age group, FP method used, and other characteristics. Most service-statistics reports are based on programs alone, however, with no regular reporting on acceptors in the private sector (the significance of that omission will vary greatly from country to country, among regions within countries, and within countries as the balance shifts between government and the main private service-providers—pharmacies, physicians, and non-government organizations. Although private, community-based distributors are normally tied in to the government reporting system). Gross acceptor figures mean relatively little; acceptor figures by method mean considerably more; but only acceptors by method, combined with continuation rates, can provide good estimates of "couple years of protection" (CYP). One then needs to compare such figures, by age groups, with the fertility of unprotected couples in the same age groups, to arrive at good estimates of how strongly FP is affecting fertility.

For projecting demographic trends, the Contraceptive Prevalence Rate (CPR) is a considerably more significant figure than acceptor figures. Taken with information about the proportions using different methods (which have different birth-prevention effectiveness), the CPR permits estimates of how many births are being averted during a particular period (but only fairly

⁵A fifth measure is "Couple Years of Protection." This measure applies weights (based on the estimated effectiveness of each contraceptive method) to the total number of users, based on the estimated numbers using each method. USAID makes considerable use of this indicator, but Bank staff, after some use in early projects, rarely if ever use it today.

crude estimates, unless continuation rates are known). This allows estimates of the CBR and, if the ages of acceptors are known, of TFRs. CPRs are not generated by service statistics, however; they require special surveys (occasionally they can be "piggy-backed" onto other surveys, but most often they must be conducted on their own).

There is a rule-of-thumb, often cited by the late Dr. Barney Berelson, the former head of The Population Council, that a 30 percent acceptance rate is typically matched by a CBR of 30—the "30/30 rule." This loose relationship is linked to another CPR/CBR rule-of-thumb, namely, that a one percentage-point rise in the CPR generates an approximately equal fall in the CBR (this relationship holds only for the mid-ranges of the two variables). So if a weak, early-stage FP program has a CPR of under five percent, and the CBR is above 50, then a 25-point movement in each rate will find them crossing at about 30.

The above review correctly suggests that demographic and acceptor indicators can be used to infer one from the other, in a loose way. Thus demographers who make projections about a country's future fertility and the size of its population can translate these projections into CPR rates implicit in those projections. And they can make statements about the annual numbers of new acceptors, and their continuation rates, that will be compatible with a given CPR. A great many assumptions will be necessary, with possible alternative values capable of producing the same result (e.g., a slow build-up of acceptors but with high continuation rates vs. high numbers of acceptors with lower continuation rates). Similarly, demographers can begin with acceptor data and, if the data are good enough (e.g., with totals, ASFR by method, and dropout rates) infer future trends in fertility and population growth. Ideally what is wanted are direct, frequently-collected demographic data that provide hard evidence on what is happening. But few countries collect demographic data on an ideal pattern, so governments and researchers must draw on whatever demographic and acceptor data are available, see if they appear compatible, and draw reasonable inferences about demographic trends.

One further comment about the linkage between demographic and acceptor indicators. If FP were the only variable influencing fertility, then there would be a 100 percent linkage; changes in acceptors and FP methods would explain 100 percent of any changes in fertility, and vice versa. No such tight linkage exists, of course. However, where the age of marriage, the prevalence of marriage, and the mores of extra-marital intercourse are all relatively stable, then there is a high linkage and fertility rates can be explained largely by changes in the numbers of acceptors, their continuation rates, and the methods they use. The literature almost invariably cites FP as the principal explanation of falling fertility; it is not uncommon to see FP cited as explaining, e.g., "two-thirds of the drop in fertility." The more difficult questions concern why FP is or is not increasing, i.e. demand-side questions. Almost all national programs, and Bank projects, contain components designed to increase the demand for FP (these are normally the Information-Education-Communication, or IEC, components. Longer-term approaches to increasing demand for FP include wider education for girls, increasing women's employment opportunities, and the introduction of incentives or privileges for smaller families).

In summary: demographic and acceptor data are the two sets of outcome indicators that allow judgments on how well population programs are working. Demographic statistics are the ultimate indicators but are generated much less regularly, and are much more costly, than acceptor data generated by service statistics. Both have important roles to play in providing insight into program performance. Both can serve as complementary program targets. But largely for pragmatic reasons of cost and availability, acceptor data normally offer a more current, more operationally

useful source of outcome targets than the longer-term, ultimate authority of demographic data. There is a difficulty, however: the demographic output data that are the ultimate indicators of progress are rarely if ever available on a project basis (either geographically or time-wise). If one wants output data geared to a project, the most one can hope for are acceptor data (even those are sometimes cast only in program, not project, terms).

The question arises whether program outcome targets are more useful for judging project performance than another kind of data, namely, project input targets (e.g., facilities constructed and put into service, staff recruited and trained, new FP methods introduced, workshops held, IEC materials produced and distributed, intended studies completed, the service statistics system improved, policies reviewed and announced, agreed operating budget levels reached, etc.). The provision of these hardware and software inputs constitutes the "system-building" which is often a primary objective of Bank projects. The creation of such motivational and service-delivery systems is especially important during the early years of any program.⁶ Later, making the system perform effectively ("organizational development," in one sense) is likely to become of greater concern. A major conclusion of this review is that input indicators deserve to be rescued from the *infra dig* limbo into which they have fallen and should be used, for project performance targeting, in tandem with program output indicators.

Up to now I have discussed targets solely in terms of demographic outcomes or acceptor statistics which suggest demographic outcomes. This perspective is too narrow, however. Something like one half of all "population projects" are in fact health projects where FP services, and demographic objectives, play a distinctly subordinate role. Such project design may reflect Government sensitivity in addressing population questions more directly, or a greater interest in health than in population objectives on the part of particular Bank staff, or a sensible judgment that improved health indicators may be a precondition of changed attitudes towards family size. In any event, we must recognize that in many health projects non-population objectives will be dominant and the primary targets will be health not population indicators. There may be no population targets at all; indeed, there may have been staff judgments that to set any kind of population or acceptor targets would have stirred resistance and been counter-productive.

IV. Comparators: Projections, Targets, Trends and "Other Places"

Appraisal reports routinely present demographic projections in addition to whatever targets they may contain. These are entirely independent exercises and should not be confused. The PHN Demographic Unit, using a standard projection model, prepares projections, every year, for about 190 countries. These projections use as their starting values recent estimates of fertility, mortality, and migration. Not infrequently, Bank operating divisions request PHN to construct alternate-scenario projections to illustrate how population growth would behave assuming faster, or slower, declines in fertility and mortality than are assumed in the single-valued annual projections.

⁶Note the following (footnote) from OED's 1991 survey of the Bank's population assistance to eight countries: "In countries just beginning to build a population program, a necessary first step is to create a delivery system and the institutional capacity to operate it. This in itself, is likely to be sufficiently difficult and time consuming that several projects could usefully be devoted exclusively to that. The Bank has effectively helped with the physical aspects of the delivery system . . . but not so effectively with the institutional aspects." (Population and the World Bank: A Review of Activities and Impacts from Eight Case Studies, OED Report No. 10021, October 22, 1991, p. 66.)

The resulting "high, low, medium" projections are used fairly frequently in PHN and other Bank reports.

Projections are arithmetic exercises that are neutral with respect to what "ought to happen." Targets are not neutral; they are normative and announce goals that managers and staff will try to achieve. They exercise a hortatory psychological influence intended to elicit a more successful outcome than would be attained without them. They may even be viewed as "quotas," with corrupting effects on service quality and on service statistics. Demographic targets are usually set at a national level; at that level they have little operational impact and typically need to be translated (through the linkage relationships discussed above) into acceptor targets of increasing specificity as one moves down in program levels from the national to that of individual service-points. Target-setting is of course a tricky business that must find the high ground of encouragement, credibility, and challenge and avoid the discouragement and cynicism that come with targets that few people believe can be achieved. While everyone feels good when targets are met or exceeded, and bad when performance falls short, evaluators of performance must worry about the reasonableness of the targets as well as about performance in meeting them.

In addition to comparing program performance with projections (a long-run comparison) and with targets (a short-term comparison) two other types of comparison are possible—(i) trend analysis, or comparing current performance with the past, and (ii) comparing country or area performance with that of other countries or areas. Trend analysis is perhaps the most important standard of all, since it compares "facts" with "facts" (current and past performance) within the same socio-cultural setting (impossible when the geographical areas are widely different).

Projections, targets, trends, and geographical comparisons are not alternatives. They are complementary and additive, in the sense that they all add something to our understanding of a country's demographic situation. But international comparisons and trend analysis are particularly useful standards for judging project or program performance.

V. The Use of Targets in Bank Projects: The Record

Table I is based on Annex II. It shows the frequency with which 11 target-related topics appear in Bank projects in each of the 42 countries covered by our universe of 75 SARs. Before turning to the table, a few words of explanation and caution are in order. **First:** the list of characteristics is somewhat fuzzy, imprecise, and arbitrary. Several of the characteristics combine a number of closely-related indicators which, if separated out for individual recording, would greatly complicate the table, make it more difficult to understand, and would add little to our objective of reporting the kind of targeting used in Bank projects. **Second:** the footnotes should be read in conjunction with the labels put on the 11 "characteristics." **Third:** the "scoring" has been done by countries, not by projects, i.e., if a country has had more than one project, all of them have been treated as if they were in fact a single project for that country. This loses any sense of evolution within successive country projects, but that can be found in Annex II if anyone needs that degree of detail. **Fourth:** Table I is a summary of a working table which showed how each country scored on these 11 characteristics. No table showing how individual countries scored on each of these characteristics is shown for the simple reason that the country summaries in Annex II are felt much more useful, and less misleading, than a country-by-country table.

Table 2. Types of Targeting/Evaluation Provisions Found in Population Lending to 42 Countries

	Characteristics	Countries
a.	No demography or family planning targets	15
b.	Government to form/announce population policy	5
c.	Main objective is to create a functioning health-delivery system (targets premature)	9
d.	Bank evaluation of government demographic targets and acceptor requirements	20
e.	Target CPR specified	21
f.	Government obligated to do key surveys (CPR, KAP, ASFR)	14
g.	MIS to be improved (especially service statistics)	6
h.	Mid-term or other project evaluation exercise	8
i.	Demographic data-base to be improved	4
j.	Upgrade institutional capacity to do operational research	9
k.	Targets for acceptors, by method	8

Notes on each characteristic

- a. Self-explanatory. But in all such countries FP services are offered in the government MCH system. Reflects either sensitivity or uselessness of setting targets in a system just getting organized.
- b. Found where government is offering FP services and has shown willingness to adopt a formal population policy.
- c. Self-explanatory (see a above).
- d. Here the Bank analyzes the consistency between a government's national demographic and program acceptor targets (or translates the former into the latter). The result is almost invariably a set of demographic and acceptor targets agreed as reasonable between the government and the Bank.
- e. A target contraceptive prevalence rate (usually set by government and accepted or modified by the Bank) is the single most common target used.
- f. A common complaint in Bank reports is that the CPR is not known. So a common project component is a requirement that such surveys be undertaken at one or more points during a project's life.
- g. Improving the Management Information System (MIS), especially the system of FP service statistics, is a fairly common component, to permit better program monitoring.
- h. Several projects (probably more than indicated) use a mid-term evaluation procedure to monitor performance. The criteria to be used by such teams are almost never spelled out (i.e., they are left to the teams themselves).
- i. Where basic demographic data are especially weak, projects may fund activities to collect and analyse such data.
- j. Several projects push for doing operational research studies, but often need to strengthen organizations that can do them.
- k. It is almost as important to know the distribution of acceptors by method as to know their overall number. A fair number of projects have set acceptor targets in terms of contraceptive methods.

Strictly speaking, only items a, d, e, and k involve target-setting (item b implicitly involves targets as well). Item c involves the establishment of pre-conditions before targets become possible or meaningful. But all the other five characteristics (items e to i) are concerned with creating the data, or the institutional capacity, that will permit performance monitoring (either against targets,

past performance, or against other, less objective criteria). Targets are meaningless unless current performance data also exist. So projects have often been concerned not only with targets but also with generating performance data for making target comparisons.

Comments on the table

The following observations are suggested:

In over one-third of the countries, the SARs make no mention of either demographic or FP targets. Most of these countries are in Africa or Latin America, where population-control and FP have only recently been accepted as government objectives, and where the subject is often still sensitive. So FP goes forward, invariably as one of many MCH services, but without any targets. In four or five countries with more than one project one can see an evolution from no targets to some kind of targets (the 15 "no target" countries do not include those countries).

The single most common target, found in exactly half the countries, is the contraceptive prevalence rate (CPR).

Used almost as frequently as the CPR (in 20 of the 42 countries) are demographic and acceptor targets, almost invariably linked. The Bank has done a lot of "consistency testing" of demographic targets (always set by the government, not the Bank) to see how many acceptors would be required to meet the demographic targets. Such analyses were particularly common in the earlier Bank projects, where extended annexes on this subject were the rule. Such analyses are less common today, probably because the early borrowers were countries with pressing population problems and where the subject was not politically or culturally sensitive; today, there are many more projects in countries which, while willing to offer FP/child spacing services, are not yet willing to talk in terms of population control and are therefore reluctant to set demographic targets.

In over a fifth of all countries the Bank is involved in helping a government get its PHC health system up and running. The earlier the stage of system development, the more primitive the institutional base, the greater the manpower shortages and experience—the less relevant seem the setting and monitoring of FP targets (see fn. 4, p. 7). That seems a more plausible explanation for the absence of targets than assuming Bank staff just forgot to include targets in the project.

In nearly one-fifth of the countries projects have set acceptor targets in terms of contraceptive method mix. Sometimes these can be quite specific, and are obviously related to decisions that must be taken on contraceptive procurement. But sometimes the "method" targets are quite general—e.g. reducing an emphasis on sterilizations and giving greater emphasis to temporary methods (which appeal more to younger couples who have not yet completed their families). Often, too, a significant number of declared "acceptors" or "users" (as revealed by CPR surveys) will respond that they use the rhythm or withdrawal methods. The effectiveness of these methods is much lower than the so-called "modern" methods; so there is often a determination, reflected in target-setting, to persuade a higher proportion of users to accept one of the four or five more effective methods—and to do this FP staff need to be given the education and pressure which method-specific targets can help provide.

Finally, and most importantly: this review of all Bank population projects makes clear that the Bank has been deeply involved in target-setting ever since it got into this sector. There is a lot of experience to draw on. It is difficult to see any need for new "experiments" to test the

usefulness of targets. Some understandable confusion among inexperienced staff assigned to population projects, and the innocence of managers who may never have worked in the sector, may be the main sources of doubt about how energetically the Bank has used target-setting for evaluating project performance. In my view, however, what the Bank needs more than greater use of targeting more thinking and discussion on how target-setting fits into a larger framework of project and program evaluation, i.e., of performance measurement. The final section offers such a framework.

My own view is that outcome, or impact, targets can often be useful as political and staff motivators within countries and can provide one important yardstick for measuring performance. They should therefore be used wherever the local political and cultural climates make them acceptable. But I put much more emphasis on monitoring trends than on setting and meeting targets. The "bottom line" is really a simple one: is the proportion of couples using effective contraception rising? If this is happening, it means that the net result of acceptor numbers, the methods they are using, and continuation rates is moving in the right direction (not all components of this trio may be rising, and we should know what is happening to each of them . . .). A rising and robust CPR (one that reflects "modern" contraception) inevitably means falling fertility. One might argue, therefore, that the CPR is "all we really need to know." While demographers may be able to work out demographic trends from the CPR (by age groups), we always want direct, independent data to cross-check our impressions. So there will always be a need for separate CPR and ASFR surveys (the separate sets of data may in fact be gathered in a single survey—see the Recommendations section for the need to increase staff knowledge, both in the Bank and in host governments, about conducting surveys. Note esp. fn. 11). I neglect the use of births recorded in vital statistics systems because of an impression, perhaps wrong, that these are almost invariably weak and unreliable.

VI. Bank Instruments Used for Evaluating Its Projects

The Bank has long given considerable attention to monitoring and evaluating its lending operations. As noted, current guidance to staff on these questions is contained in Operational Directive 10.70; this distinguishes between "monitoring" and "evaluation," defines Bank policy and the roles of PCRs and PPARs, and discusses the need for including in projects a management information system that will generate "a minimum core of quantitative and qualitative indicators for monitoring progress towards [quantified project objectives]." But the guidelines wisely note that, "There can be no standard list of indicators, even within an individual sector . . ." Such humility throws the ball into PHN's court to define its indicators as it thinks best. Most of these are "second nature" to experienced staff working in population (see above, para. 12). But, as noted earlier, evaluation of project performance concerns more than using a set of standard sector indicators.

It is useful to remind ourselves of the large number of instruments the Bank uses to tell us how well a project is doing or has done. Some clearly have different primary objectives. Not all are available for all projects. Most become available at different times in a project's history. But presumably all do have, or should have, something to tell us about a project's performance. Here's the list:

Supervision Reports

Internal bank progress-reports, focused on how well project implementation and expenditures are keeping to schedule. Provides a check on the periodic reports, on essentially the

same subjects, which the borrower is required to submit to the Bank. Supervision reports rarely report on program performance,⁷ being preoccupied with project implementation.

Mid-Term Project Reviews

A fairly high proportion of projects provide for these, and have done so since the first one in 1970. Where a project is large in relation to a country's total program, the reviews will examine the program's performance, not just that of the project alone. The early reviews were usually done by outside experts. More recent reviews (and there are several) are often internal, done by the borrower's own staff. Both types are intended to result in written reports copies of which would come to the Bank. However, I have not seen more than one such review and have no judgment on how useful they are as a monitoring device that leads to program/project adjustment. Would it be useful to have a YP or Research Assistant review such Reviews to see if they are anything more than a bureaucratic nuisance that makes Bank project designers feel virtuous when they mandate them? A different kind of external review is the Bangladesh model: the program's several donors meet annually with program administrators to review recent progress and agree on a work plan for the next year. The government is provided a list of indicators to be used in briefing those who attend the meeting. The incentive influence of this annual, in-person reporting requirement may be as useful to the system as the existence of acceptor targets. The purpose (said to be well realized) is to provide a forum for dialogue that can help identify needed mid-term adjustments. The approach is, of course, additive to the simple existence of targets and is intended to help reach (or revise) them.

Operational Research Studies

A fair number of projects have components that aim at creating increased borrower capacity for conducting such studies. Needed individual studies are almost never listed, however. I doubt anyone knows which country programs may have received useful information from these components—it is probably enough just to provide generic support for capacity-building and hope for the best. A different but related route to program evaluation and "fine tuning" is to mandate that in-country staff themselves develop objective indicators for their activities and meet periodically to assess how well they are doing (cf. the two Zimbabwe projects, esp. the second).

Repeater Projects

These are one of the richest sources of program and earlier-project performance data. Each successive SAR invariably contains information about program progress. They give high visibility to such progress to Bank staff and to the Board. If I were looking for progress reports on program performance in a country with more than one Bank project, I would turn first to successive SARs.

⁷Operational Directive 13.05 instructs staff to attach to Supervision reports an annex (normally not longer than one page) that displays "in tabular form...quantitative measures of project performance in critical areas, such as construction progress, yields, adoption rates, enrollments, target groups reached, traffic, operating efficiency, and finance. The indicators should be selected on the basis of the SAR and agreed with the SOD/COD chief concerned." I have not reviewed any supervision reports to see how often this is done.

Project Completion Reports (PCRs)

These report primarily on project implementation, not its subsequent performance. They are concerned with essentially the same data that supervision reports deal with during implementation. Written by the Department responsible for project supervision, PCRs can be written only after projects are complete and typically appear a year or so after the final disbursement has been made. As yet only 31 of the 75 projects covered by this review have been completed long enough to have their PCRs written. They are not an important source of information on project performance after completion.

Project Performance Audit Reports (PPARs)

While every completed project gets a PCR, PPARs are done on only a sample of projects. Seventeen Population PPARs had been written as of October, 1991. Their purpose differs entirely from the PCR: the latter addresses implementation, the PPARs address what happens after implementation is complete. They are meant to be evaluative and judgmental and to throw up lessons that will lead to better projects in the future. They are written not by those who designed or supervised the project but by neutral, uninvolved outsiders, responsible to OED. A PPAR is begun about when a PCR has been completed, i.e., roughly two years after implementation. So the operational period brought under review is relatively short. It can, however, answer the question, "Is the project off to a good start or a bad one, and why?" To the extent that project components began functioning before final completion, the PPAR may be able to offer useful comments on early project performance. A review of the 17 PPARs shows that where targets were included in the SARs, the PPARs reported on performance against them but were usually silent on targeting when no targets had been set. The PPAR for the first Malaysia project, however, specifically criticized the project for having failed to specify indicators.

Non-Bank External Reviews

PPARs are not the only type of independent performance reviews employed by OED, nor has OED been the only sponsor of external reviews. Within the past few years, OED has commissioned eight such reviews of Bank lending to eight different countries (Bangladesh, Brazil, Colombia, India, Indonesia, Kenya, Mexico, and Senegal) and has prepared a useful summary of the findings (see below, para. 29 ff.). In 1985 the old PHN Department (i.e., the operating department before the Bank reorganization of 1987) commissioned a review of the Bank's population lending and sector work by a two-man team from the Univ. of Michigan (the late George Simmons and Rushikesh Maru)⁸. Relevant parts of both those reports are summarized in the following section.

VII. Three Perspectives from Outside Bank Operations

It will be helpful to review what others have said about the role of targets in judging FP project and program performance. There is of course a large literature on evaluation. I have made no attempt to review that literature. For our interest (i.e., judging the Bank's approach to performance evaluation, with special reference to the use of targets) it is sufficient to summarize (i)

⁸Published as a blue-cover Working Paper (WPS 94) under the title, The World Bank's Populations Lending and Sector Review (September, 1988.97 pp.).

the two reports cited in para. 25 (vii) that deal specifically with Bank practice and (ii) a forthcoming major research study, funded by USAID, that will try to bring more order into a somewhat confused and contentious field.

The 1985 Report by Simmons and Maru

Simmons and Maru are not very helpful on how they think performance-monitoring should be handled. On the one hand, they complain that supervision missions pay too much attention to humdrum questions of project implementation and tend to neglect "measures of program activity such as the quantity and quality of transactions or of intermediate output measures such as prevalence" (p.40). If they really mean to load such reporting on supervision missions, I can only disagree. They themselves note (p.60) that "Supervision missions provide neither the time nor the incentives for the mission leader to give adequate attention to process evaluation." But if time and incentives were provided, how useful is it to worry too much about "process" events while the project is still being built? Nor do I understand their proposals for measuring the "quality of transactions" in a project or program—a concept they seem to favor but nowhere define.⁹ They note that "The problem of evaluation is compounded due to weak project designs, which do not detail evaluation procedures and measures"; this reflects, they say, "lack of evaluation design and targets at the project formulation stage" (p.61). They apparently felt that the Bank could do more with the use of targets. At the end of the day, however, they say that "it may never be able to evaluate [project] effects adequately . . . because of the complexity of projects and the environment in which they are implemented" (p.74).

Their final advice to the Bank is that it "adopt a systems approach to the evaluation of projects and that evaluation be undertaken on a more regular basis to avoid the delays that have been experienced in the past" (p.74, para. 7). They do not say what they mean by "a systems approach." If they mean that repeater-project preparation and appraisal missions, or OED internal or external evaluators, should collect all the quantitative and qualitative performance information they can find and come up with a holistic judgment on how well a project (or program) has performed, I would agree. Performance against targets would provide one useful input to such an overall judgment. But one does not come away from Simmons and Maru with very clear ideas on how evaluations ought to be made and what role targets should play in making them.

The 1991 OED Report

As noted, OED has recently completed a review of Bank population work in eight countries. The case studies were done by outside experts and all were then summarized, and lessons drawn, by OED (*Population and the World Bank: A Review of Activities and Impacts from Eight Case Studies*, October 1991, Report No. 10021, 92 pp.). Among the major conclusions of the review are the following:

⁹In its proposal to USAID for the EFPPI project (see below), the winning contractor has some suggestive comments about how the "quality of care," or service, might be assessed. It suggests six elements: choice of methods, client counseling, staff competence, interpersonal relations, follow-up, and service acceptability. It goes on to suggest how these quality-elements might be evaluated. See also p. 27 ff. below.

Bank staff and Bank projects have been too oriented to the supply side of population work and have paid insufficient attention to the processes of social and economic change underlying the demand for family planning. If staff had searched more diligently for "selective interventions into the development process that [have] the possibility of changing implicit benefits and costs of large families—far more might have been accomplished" (p. xi). It is not enough, the report feels, to rely on the Bank's general program of development lending to produce the kinds of changes that generate a desire for smaller families. The assertion is made that had the Bank searched for non-population projects and components that lead to smaller families, it could have identified them and results would have been stronger (para. 8.05). On the lending side, therefore, the main conclusion is that the scope of population activities should have been broader.

As for "monitoring and evaluation," the report notes that some such component "is required in all projects"¹⁰ but few have one, and the subject is "generally neglected during supervision." A frustrating aspect of the eight-country review has been "our inability to determine more clearly whether population projects are having the desired impact or not. It is precisely because of the neglect of monitoring, evaluation and research components in these projects that this is the case" (8.25). The report then suggests that "without baseline and control group data . . . proper assessment is impossible."

Part of the problem in getting useful evaluation work done is that local capacity for doing such work is absent or weak. While there has been "a large volume of so-called operations research, its quality is often dubious." The conclusion is drawn that if the Bank and other donors had paid more attention to building such capacity, the quality of research output would have been better.

Within the Bank, demand for strong monitoring, evaluation, and research activity is said to be weak. The Board should require "evaluations of past efforts . . . to explain and justify projects proposed to it. It would also help if senior management were to make it clear that a proposed project has not been properly designed—and therefore is not ready for Board presentation—until it includes a convincing plan for establishing a monitoring and evaluation capacity and then for utilizing that capacity to answer questions about the project" (8.29).

At the end, however, the report's evaluation focus shifts from the level of project performance to progress in achieving sector goals. "In evaluating population projects, more weight should be given to proximate output measures, focusing on those believed to be the most important for achieving sector goals" (8.55). There is no indication of what "proximate output measures" should be used or how such sector-level indicators are to be used alongside the project-level monitoring and evaluation indicators referred to earlier. Indeed, the report complains, near its end, that too often evaluation focuses "on projects rather than non-project activities, to ask whether the project did what the SAR . . . said it would do, and within projects to focus on the more expensive and quantifiable inputs. While this is important for accountability purposes, it leads to an excess focus on projects, on inputs rather than outputs, and more important, on the specific inputs included in the project rather than on whatever is needed to advance the goals of the sector" (8.54).

While some of the OED comments are not as explicit, transparent, and actionable as one might wish, I agree with the final comment that monitoring program (=sector) output performance

¹⁰See para. 4 of Operational Directive 10.70: "Plans for monitoring and evaluation are to be included in all Bank-funded projects..."

should hold a higher priority than project monitoring. The latter is not irrelevant, just less relevant, in many cases.

Major Forthcoming Study Sponsored By USAID

In the fall of 1991, USAID awarded a multi-million dollar, five-year evaluation contract to three well-known U.S. population groups, led by the Carolina Population Center at the University of North Carolina. The effort, known as the "Evaluating Family Planning Program Impact," or EFPPI, project, aims to develop a methodology for assessing the impact of USAID'S generation-old, world-wide population programs. The project is the most ambitious such effort ever undertaken and will be of great interest to all agencies active in the field. The main outputs of the project are expected to be available (in draft, at least) by the end of 1993.

It is a commentary on the unsettled state of evaluation methodology that USAID has felt the need for such help. The winning bidder's 64 single-space pp. proposal (copy on file with Mr. Bulatao in PHN) itself provides a useful sketch of the present state of the art and lists six documents which it expects to have ready, in draft, within two years (a state-of-the art review of evaluation methodology; a conceptual framework paper; a topology of family planning programs; an evaluation manual for practitioners; a handbook of consistent indicators, with standardized definitions; and interactive software that can be used for computerizing the results of Demographic and Health Surveys (DHS). In addition, the project will include a series of country and methodological studies that will presumably apply and test the recommended methodologies.

While the EFPPI project will remove the need for any major, formal work on evaluation by others for the next few years, it should not stop the Bank from trying to improve its own performance indicators without waiting for results from the EFPPI project. The recommendations of this review are made in that spirit.

VIII. Conclusions and Recommendations

A. Word Of Caution

This review was commissioned because there is apparently some uneasiness among Bank managers that not enough attention has been paid to target-setting, and to the evaluation of project impacts, in the Bank's population lending. My reading of the evidence does not lead me to this view. One of my principal conclusions would be that the better one understands the nature of targets and indicators, and their applicability to different levels and stages of population work, the higher the marks one gives the Bank on how it has used them. There is always room for improvement, however; the six recommendations at the end may help achieve it.

There are two core evaluation questions Bank managers and staff are typically interested in when looking at population operations:

- ▶ How well is a country's program doing? (present tense)
- ▶ How well did a particular Bank project do? (past tense)

One of the difficulties in dealing with targets and indicators is that people are not always clear which question they are asking, or they assume that measurements applicable at one level must be equally

useable at the other. Unfortunately, this is not so. As we will see when we look at a simple model later on, targets and indicators are very useful when looking at the input end of project activity; but when attention turns to "impacts" and "outcomes," the most useful indicators are program indicators.

Perhaps the most controversial point I want to make, especially for non-population, non-project staff, concerns how the Bank can best learn from its project experience so it can design better projects. Some people apparently feel that more and better targets and indicators, built into project design at the start, are needed to support such evaluations, especially the evaluation of major project components or activities. Population projects are complex enough without overloading their managements and information systems with requirements for data which, even if successfully collected, are unlikely to make much difference to the design of future projects, either in the same or other countries. "Strategic thinking" inspired by imagination and based on a knowledge of international population programs, and on a country's own situation and on the activities of other donors, is decisive for project design. What is to be encouraged for better evaluation work are well-focused operational research studies on specific topics, and general program reviews, not the multiplication of targets and indicators which may or may not someday be used.

Past Use Of Targets and Indicators

Projects should be evaluated in terms of their objectives. Not all population projects, however, make it a stated objective to increase FP acceptors and the CPR, and reduce fertility. Those may be long-term Bank goals for a country, and increasingly shared by most governments. But a long-term goal for a country is not the same thing as a five-year goal for a project. As noted, a project, particularly a first project, may simply aim at basic system- and institution-building: getting a Primary Health Care network built and staffed, organizing an MCH program, introducing FP services for the first time, or helping establish a population policy and program. Beginning in the early eighties it became a requirement that all SARs should include an explicit statement of the project's objectives. These SAR paragraphs are usually informative, free of "boilerplate," and are one of the likely sources of information about targets if targets are included. It would be a mistake to mandate that quantitative output or outcome targets (e.g., acceptor or CPR levels, fertility reductions) be included among the "statement of objectives" in all projects. Sometimes targets are appropriate, sometimes they are not (never when governments don't want them).

The Annex II SAR summaries show great interest in targets, beginning with the very first project and continuing down to the present. The fact that about one-third of all country operations have not used targets does not reflect neglect and weakness of project design but realism, pragmatism, and a respect for Government wishes. So, by and large, I think the Bank has done a good, professional job when it has used quantitative targets, and has not been sloppy or negligent when it has not used them.

More often than not, the Bank has analyzed and commented on government-set program targets, sometimes urging higher ones, sometimes lower and more realistic ones. In two or three cases the Bank has recommended that programs without targets establish them. Generally speaking, the Bank is certainly "pro-target." Staff know targets spice up the game and provide a way of keeping score. They may also realize that targets provide an arena for dialogue, both in their setting and during subsequent monitoring. However, playing the critic to government program targets serves a different function than setting targets at the level of Bank projects: the latter may or may not be sufficiently independent of national program activities to make project targets feasible. Even where feasible they may not be desirable, since they can distort programs and record-keeping.

The kind of targets just discussed are normally "intermediate output targets"—i.e., acceptor and CPR targets, although they may sometimes include "final outcome" demographic targets; any others are much less important (for types of indicators and targets, see below, para. 46).

When we move from the level of national program to that of the Bank project, we move into a different ball-game. It is the level at which Bank accountability for the use of Bank funds is important (providing a major reason for the Bank's traditional emphasis on project supervision). It is also the level at which Bank Management seeks to draw lessons for the effectiveness of Bank operations, an interest institutionalized in the work of the Operations Evaluation Department (OED) and the preparation of Project Completion Reports (PCRs) and Project Performance Audit Reports (PPARs).

Annex III summarizes the findings of the 17 PPARs that have been completed to date. A recurrent theme in the OED audits is that Bank population projects have not developed a sufficiently clear hierarchy of objectives, indicators, and targets that can be used for program, project, impact, and process measurements. OED has not attempted to provide such a hierarchy, however; it has simply urged PHN to do so. Although not undertaken for that purpose, this paper counts as a step in that direction.

My generally supportive evaluation of how Bank operational staff have used targets and indicators does not leave Bank practice with a perfect score, however. The main problems I detect are the following:

- ▶ Bank managers may entertain more ambitious roles for targets, esp. at the project level, than the Bank's population professionals, few of whom believe outcome targets are useful at either project or program level.
- ▶ There may not yet be enough consensus on the differences between targets and indicators, and not enough understanding that targets are not the only comparators available for judging performance. All targets require indicators but not all indicators require targets.
- ▶ Indicator terminology is not standardized, and while confusion is not great, it would be helpful, and not difficult, to achieve greater consistency in how different types of indicators are labeled.
- ▶ I suspect that only a handful of the Bank's non-population staff have an adequate understanding of the three key indicators and their construction (i.e., "acceptors," the Contraceptive Prevalence Rate, and the Total Fertility Rate). In particular, I see a need to give greater attention to the problem of "contraceptive mix," a key aspect of a program's ultimate effectiveness, i.e., its demographic impact.
- ▶ Finally, Bank staff need to recognize the new concern for the quality of FP services and should help develop better indicators for measuring this dimension of projects and programs (see paras. 50-54).

Further comments and suggestions on these points will be found in the paragraphs that follow. The whole discussion will become clearer, however, after we walk through a simple model.

A Simple Model Suggests Four Sets of Indicators

I find it helpful to use a simple model of what a FP program is and what we want it to accomplish. The model is shown in Figure 1. On the left we have the FP program, or delivery system, showing some of its main components. The program is split into two main parts (elements concerned with the delivery of FP services, and elements concerned with stimulating demand). The system (=the FP program and its physical, human, and financial resources) represents a static potential that has no significance independent of its functioning. But to function, it must first be built—not usually all at once, but over many years, the result of many projects. "Systems" and "programs" are never finished and complete. New activities and projects are always coming along to enlarge or modify existing systems.

The design of the system, or increments to it, and bringing them into existence belong to project design and implementation. Project supervision is the monitoring of implementation, a process that involves an intense period of physical and financial reporting to measure progress towards project completion. As project components reach completion they are ready to function, to begin the repetitive activities that contribute to the purposes for which the system was built.

The "activation" or functioning of the system produces a stream of input activities whose sum constitutes a process intended to recruit FP acceptors, reduce unwanted births, and lower fertility. It is desirable to know the activity levels (and, if feasible, the quality levels) of key activities during system operation (an ongoing, long-term process). To the extent we can invent indicators of such activity, we have a set of "process indicators." But looking at the process indicators tells us nothing about what effect they are having. We want to know how the system is performing in terms of recruiting and retaining acceptors. So we develop a set of "performance indicators" that measure output performance, or "yield." Performance indicators suggest what is probably happening to the ultimate outcomes or effects we are hoping for, but they cannot give us conclusive information. The results they measure are intermediate between system input activities and the desired final outcomes (i.e., lower fertility and slower growth in population). To see if we are achieving those final outcomes we need yet another set of indicators which will quantify, and either confirm or revise, the inferences we draw when we look at the intermediate output indicators.

All this is simple and straightforward and should be clear from Fig. 1. We now see that we have four sets of indicators, two of which measure system inputs, two outputs or results:

Inputs

- ▶ Project implementation indicators;
- ▶ Process, or activity, indicators, which measure recurrent, post-implementation activities.

Outputs

Performance which measure the intermediate outputs, or non-demographic results, of system functioning. Occasionally, in Bank documents, these indicators are referred to as "proximate variables" (unhappy, as it conflicts with demographers' use of that phrase);

Figure 1. SCHEMATIC MODEL OF A FAMILY PLANNING PROGRAM AND INDICATORS OF ITS PERFORMANCE ¹
(Sequential flow: read from left to right)

I. Inputs: The FP System and its activities			II. Intermediate Outputs: system performance or results					III. Final Outputs: Demographic outcomes or impacts				IV. Economic Benefits (beyond FP system)
Policies	Service Delivery	Demand Activs.	New Acceptors	Continus./Dropout Rates	Current Users	Contrace. Prev. Rate	Cple Yrs. of Prot'n	Current Fertility	Births Averted	Net. Rate of Incr	Pop. Growth Rate	
Pop. & FP	govt. serv. prnts. staff & their trg. priv. physicians pharmacies CBD distributors NGOs methods offered med. policy on method delivery availability of supplies staff incentives records & stats. supervision syst. trg. facils. / progs. budget support	IEC: software & hardware user FP incentives small-family social incentives polit. & relig. support	by method, age, parity service-point, etc. (usually govt. prog. only)	ests.fr. serv. stats. (client re-visits & supplies distr' d) hard data fr. tracer-interviews and/or surveys	cumula. "new acceptors" minus est'd dropouts priv. sect.?	$CPR = \frac{\text{current users}}{MWRA^2}$ by method (rough est. w. many assumptions, or requires survey)	CYP	ASFR; TFR; CBR ²	(est'd)	NRI	ests. or censuses	Higher p.c. Incs. Better job mkt. Less pressure for basic social expend. Less pressure on land & environment
			<p>————— Routine Service Statistics —————</p> <p>(from within the FP system, generated at high frequency— weekly, monthly...annually)</p> <p>————— Performance, or Intermediate Output, Indicators —————</p>					<p>————— Demographic Outcome Indicators —————</p> <p>(low frequency, usually multi-year intervals)</p>				
			<p>(Indicators of initial performance results achieved by the FP system These suggest what is probably happening at the final-outcome stage.)</p>					<p>These are all different measures of progress in achieving the ultimate program goal of lower fertility and slower population growth. Calculated by demographers from data independent of the FP system. By quantifying final outcomes independently, such measures either confirm, or revise, the outcome inferences drawn from earlier readings of the intermediate indicators.</p>				
<p><u>The Four Families of Indicators</u></p> <p>Inputs</p> <p>1. Project implementation indicators</p> <p>2. Process (or program activity) indicators</p> <p>Outputs</p> <p>3. Performance, or intermediate output indicators</p> <p>4. Demographic, or impact, indicators</p>												
<p>1. A. Created over time, though a series of projects each of which has a start, finish, and implementation period in between. Project implementation involves its own supervision indicators to monitor and report on project progress. When finished, these indicators are no longer needed.</p> <p>B. Upon completion, projects/programs start performing the repetitive activities intended to achieve system objectives. The activity or "process" events of holding clinic sessions, making home visits, showing films, airing broadcasts, conducting studies, retraining staff, issuing and analyzing statistics, etc. can be measured through process indicators. An emerging frontier is evaluation of the quality of key processes; but few programs yet have quality indicators, except dropout rates (when available) Service quality is of course reflected indirectly in performance indicators.</p> <p>2. Married Women of Reproductive Age, Age-Specific Fertility Rate; Total Fertility Rate; Crude Birth Rate.</p>												

Demographic outcome indicators, which measure the degree of success in realizing ultimate objectives. These are sometimes called "impact indicators" since they usually refer to ultimate demographic results beyond which no questions are asked.

The position of these indicators should be clear from a careful reading of Fig. 1. They cover a sequence of activities in which the output of one stage becomes an input of the following stage—or in which good performance shown by a lower-stage (left-hand) indicator is a precondition for good performance at a higher-stage (right-hand) indicator. Notice also that different individual indicators, and classes of indicators, have different frequencies, and involve quite different costs for constructing and maintaining them. The final-outcome indicators have the lowest frequency—it is costly, and a considerable administrative effort, to conduct surveys and censuses. Within the performance indicators, CPR figures, (or continuation and dropout rates, both of which require surveys), are constructed much less frequently than acceptor data, which are produced weekly or monthly by a system's routine service statistics. The routine, lower-cost, frequently-issued indicators are often taken as proxies for what is happening at the "ultimate outcome" stage; however, their often low quality, and limited detail (e.g. acceptor figures may not show distribution among methods) may mislead inexperienced readers. Thus it is very important to construct outcome indicators periodically to see what is really happening. Today, DHS surveys are the standard way of doing this.¹¹

The above terminology, and the stages to which they are tied, are not used consistently in Bank documents. One encounters use of "performance," "process," "output," "outcome," and "impact" indicators but they are almost never defined, are not always clear, not always logical, and not always consistent. If the Bank could standardize terminology along the lines of para. 46, it would be consistent with usage in most of the leading population institutions.

At Project Level, Emphasize Activity Indicators, Not Outcomes

The issue of improving indicators is far more complex and important than wondering if the Bank makes sufficient use of targets. As noted, indicators without some kind of comparators are meaningless. Only three types of comparators are possible: targets, past performance, and the performance of other systems (usually other countries). While targets can play a useful role (particularly at national or program level), most population specialists strongly prefer trends and

¹¹The DHS (Demographic and Health Surveys) program of sample surveys has been developed in recent years by IRD/Macro International of Columbia, Maryland. The phrase "DHS surveys" has, in fact, become almost a proprietary term for the work in which IRD/Macro Int'l is now pre-eminent. In addition to its technical quality, the DHS program is given high marks for the speed with which it compiles and releases survey results.

Today's DHS surveys are the direct descendants of two similar programs conducted during the 1970s and early 1980's, the World Fertility Surveys (WFS) and the Contraceptive Prevalence Surveys (CPS), both sponsored by the International Statistical Institute in The Hague. The WFS were funded jointly by USAID and UNFPA, the CPS by USAID alone. Today's DHS surveys have so far been funded entirely by USAID (an initial Bank-funded DHS, in Namibia, is expected to be authorized soon). DHS surveys have been conducted in some 35 countries, with another 20 or so scheduled for coverage within the next year or two. Repeater surveys (ideally every three to five years) will be done in 10-12 countries, bringing the total number of DHS to over 50. The typical individual survey involves local costs of U.S. \$250-400,000 (the additional cost for IRD's technical assistance has so far been covered under a multi-year USAID contract). Before the Bank can make greater use of the DHS resource, (i) more staff must become aware of its existence, and (ii) borrowers must be willing to include the costs of such surveys in Bank projects if grant funds are not available.

international comparisons for measuring performance, not success in meeting targets.¹² PHN could perform a useful service by assembling tables and (especially) graphs that provide Bank staff with convenient comparators—the levels attained by representative countries on key FP and demographic indicators and the rate at which progress has been achieved. One summer intern should be enough to get such a data-bank established.

The reservations expressed about most project-level output indicators do not apply at the input and activity levels, however. Simple service statistics should of course be produced (they typically contain a mixture of activity and performance indicators). The entire process of project implementation, for example, is one of performance against targets—getting facilities built, equipped, staffed, and put into service; procuring vehicles; procuring supplies; training staff; etc. Every "project description" in an SAR is a list of targets, full of quantitative indicators; project supervision is a detailed monitoring of progress in meeting those targets. The Bank is widely credited with doing an excellent job at the "supervision" stage, assuring that funds are spent for their intended purposes. The game becomes more difficult when we move to the next stage—establishing targets or other indicators for the activities made possible by successful project implementation (see para. B at bottom left of Fig. 1). More attention deserves to be given to incorporating indicators of program (incl. project) activities into the program's normal MIS. Types of activity or process indicators that should be considered for collection are illustrated below:

- ▶ no. of active service-points, by type
- ▶ vacancies, by category, at service-points ("fill ratio")
- ▶ no. of daily/weekly visitors by type of service-point
- ▶ no. of home visits made by home visitors
- ▶ training and retraining courses held and numbers attending
- ▶ contraceptive supplies distributed, by type and by service-points
- ▶ IEC activities (radio/TV spots; films produced and no. of showings; schools using pop. ed. materials).

Most of these indicators should be generated by a program's normal service-statistics or MIS system. Again, they should normally be reported at program, not project level (it should not be difficult to do special analyses at the project level if there is reason for such; but routine reporting of such figures for project facilities alone only generates too much unused paper).

A New And Difficult Dimension: The Quality Of Service

There is one subtle but important recent development in FP programs that deserves to go on the agenda for "improved indicators." I refer to the increased concern for measuring the quality of services being offered. Its measurement involves a mixture of input and outcome indicators.

¹²Some experienced project staff warn against the temptation for Bank population experts to push their "good performance" projections on government officials as "targets." It is inherent in the work of senior government officials that they often get caught up in political considerations likely to affect program design and implementation — including the setting of targets. Bank staff must be sensitive to such considerations and remain flexible in deciding when it may be useful, and when dangerous, to establish program or project output targets. Trends and international comparisons provide comparators that avoid political sensitivity.

Concern for service quality reflects the broadening of FP objectives beyond the demographic goal of slowing population growth which has dominated programs for nearly three decades. That "top down," supply-driven goal has not become invalid and obsolete. It is, however, being supplemented by increased concern for giving clients (mostly women) more information about their options, for widening the number of options being offered, for educating rather than "selling" them, for worrying about their reproductive health and sexuality as well as about their fertility, and for treating clients so that they become satisfied, healthy, continuing contraceptive users. As one advocate of the shift has expressed it, programs are moving from demographic objectives with individual benefits to the provision of individual benefits with demographic impacts. Such an emphasis is demand- or client-driven; it is in fact the orientation that private FP providers have had for many years.

The top down (demographic) and bottom-up (client-service) goals need not be opposed to each other, although it takes a special effort to serve both effectively. The keys are (a) the tone of program leadership and (b) the quality of staff training. With this summary, what can we say about indicators that might measure the quality of project or program services?

The PHN Board paper (Spring, 1992) devoted nearly eight pages to service quality as an element in program effectiveness. The paper identifies the following factors as key elements in service quality: access to services, the variety of methods offered, continuation of use, sensitive counseling and informed choice, and staff technical competence. Some of these components can be measured, as a first-approximation, with check-lists: e.g., what methods are offered at which types of facilities? Which categories of staff are permitted to provide which kinds of services? Do training/retraining courses include modules on counseling techniques? Other aspects might require sample surveys—e.g. of client waiting times and travel distances; of client satisfaction with services received.

An important indirect indicator of service quality, of course, is the dropout rate (or its obverse, the continuation rate), since low rates reflect higher satisfaction than high rates. Dropout-rate studies are much rarer than CPR sample surveys, because they are more difficult to do (data need to be collected over extended time periods, for different methods, since most clients eventually discontinue use, or change methods, for one reason or another, even in high-quality programs). Bank projects probably ought to push harder for more dropout studies, although their feasibility will depend in part on professional and institutional resources available in a country or on program use of technical assistance (the DHS program, for example, includes some dropout-rate questions as part of its core DHS questionnaire). The "bottom line" with regard to quality-of-service indicators seems to be: it is a relatively new field, with progress likely to consist of a combination of one-time check-lists, sample surveys of client satisfaction, and increased emphasis on continuation-rate studies.

Six Recommendations

We close this review by listing six specific recommendations for improving how the Bank uses targets and indicators in its population work. They are:

- ▶ Staff should standardize terminology by using the terms illustrated in Fig. 1 and listed in para. 46.
- ▶ PHN should prepare, periodically, a set of comparator graphs and tables for use in Bank project and sector reports.

- ▶ Operational staff should show stronger concern for a program's contraceptive mix, both at the stage of project design and in measuring program performance. iv. Increased attention should be devoted to measuring a program's service quality. This requires sensitivity to an important recent development in the field and a need to agree on and use a set of appropriate indicators (see above, paras. 53-54).
- ▶ The use of Demographic and Health Surveys (DHS) should be the rule, not the exception in Bank population and health projects. While a U.S. firm is now pre-eminent in this field, the Bank should work with it to help create additional survey capabilities, both national and international.
- ▶ Serious thought should be given to discontinuing PPARs for population projects. They rarely add much to PCRs, and the resources devoted to them could be better used in funding internal or external operational research studies, and to external reviews of national population programs.

Most of these recommendations concern clarification of Bank thinking and terminology, first among professionals engaged in population work, then among managers whose job it is to worry about the effectiveness of Bank assistance. We can achieve these modest objectives by holding a few staff seminars and circulating a paper or two. The hoped-for consensus may point to a need for wider use of certain indicators, involving some additional project record-keeping, costs, and staff. But instead of pressing for a maximum set of refined but probably never-used indicators, the pressure should all go in the opposite direction—seeing how little one can settle for. Periodic evaluation reviews by joint Bank/borrower teams or by external review missions, or *ad hoc* operational research studies, can alone provide the rounded judgments on which *ex post* project-design assessments must rest. Such evaluations will of course need to assemble whatever quantitative measures they can find; but serviceable proxies can almost always be found without having to burden a project or program with "indicatoritis," a disease for which there is no good vaccination except restraint.

Annex I. PHN Projects with Major Population Objectives (1979-91)

Country	Project Name	SAR Date	Ln/Cr Amount (million US \$)	
Bangladesh	1. Population Project	2/3/75	15	Cr
	2. Second Population & Family Health Project	5/16/79	32	Cr
	3. Third Population & Family Health Project	12/24/85	78	Cr
	4. Fourth Population & Health Project	5/20/91	130	Cr
Botswana	Family Health Project	11/11/84	11	Cr
Brazil	1. NE Basic Health Services Project	4/30/86	59.5	Ln
	2. Second NE Basic Health Services Project	10/25/89	267	Ln
Burkina Faso	Health Services Development Project	5/20/85	26.6	Cr
Burundi	Population & Health Project	11/16/87	14	Cr
Colombia	Health Services Integration Project	6/18/85	36.5	Ln

Comoros	Health & Population Project	6/22/83	2.85	Cr
Dominican Rep.	Population & Family Health Project	8/31/76	5	Ln
Egypt	1. Population Project	9/28/73	5	Cr
	2. Second Population Project	8/14/78	25	Cr
Ethiopia	Family Health Project	5/12/88	33	Cr
Ghana	1. Health & Education Rehabilitation Project	12/24/85	15	Cr
	2. Second Health & Population Project	11/16/90	27	Cr
Guinea-Bissau	Population, Health & Nutrition Project	3/27/87	4.2	Cr
Haiti	First Health Project	11/28/89	28.2	Cr
India	1. Population Project	5/15/72	21.2	Cr
	2. Second Population Project	1/15/80	46	Cr
	3. Third Population Project	11/9/83	70	Cr
	4. West Bengal—Fourth Population Project	6/17/85	51	Cr
	5. Fifth (Bombay & Madras) Population Project	5/25/88	57	Cr
	6. (National Family Welfare & Systems Development) Population Project	6/8/89	11.3	Ln
			<u>113.3</u>	Cr
		124.6		
	7. Seventh Population Project	4/24/90	10	Ln
		<u>86.7</u>	Cr	
		96.7		
Indonesia	1. Joint IDA-UNFPA Population Project	2/29/72	13.2	Cr
			(+13.2 UNFPA)	
	2. Second Population Project	5/31/77	24.5	Ln
	3. Third Population Project	5/19/80	35	Ln
	4. Fourth Population Project	4/10/85	46	Ln
	5. Fifth Population Project	2/8/91	104	Ln
Iran	Population Project	4/23/73	16.5	Ln
Ivory Coast	Health & Demographic Project	4/29/85	22.2	Ln
Jamaica	1. Jamaica Population Project	6/2/70	2	Ln
	2. Second Population Project	5/26/76	6.8	Ln
	3. Population & Health Project	5/1/87	10	Ln
Jordan	Primary Health Care Project	4/11/85	13.5	Ln
Kenya	1. Population Project	2/20/74	12	Cr
	2. Integrated Rural Health & Family Planning Project	4/14/82	23	Cr
	3. Third Population Project	4/20/88	12	Cr
	4. Fourth Population Project	3/9/90	35	Cr
Korea	Population Project	11/16/79	30	Ln

Lesotho	1. Health & Population Project	4/1/85	3.5	Cr
	2. Second Population, Health, & Nutrition Project	6/10/89	12.1	Cr
Madagascar	Economic Management & Social Action Project	10/27/88	22	Cr
Malawi	1. Health Project	3/24/83	6.8	Cr
	2. Second Family Health Project	2/26/87	11	Cr
	3. Population, Health, & Nutrition Sector Credit	n.d. (2/91?)	55.5	Cr
Malaysia	1. Population Project	12/21/72	5	Ln
	2. Second Population & Family Health Project	6/16/78	17	Ln
Mali	1. Health Development Project	11/15/83	16.7	Ln
	2. Second Health, Population, & Rural Water Supply Project	2/22/91	26.6	Cr
Morocco	Health Development Project	5/15/85	28.4	Ln
Niger	Health Project	2/26/86	27.8	Cr
Nigeria	1. Sokoto Health Project	2/21/85	34	Ln
	2. Imo Health & Population Project	9/1/88	27.6	Ln
	3. National Population Project	4/21/91	78.5	Cr
Pakistan	Proposed Population Project	4/1/83	18	Cr
Peru	Primary Health Project	10/29/82	33.5	Cr
Philippines	1. Population Project	5/22/74	25	Ln
	2. Second Population Project	5/23/79	40	Cr
Rwanda	1. Family Health Project	3/6/86	10.8	Cr
	2. First Population Project	5/15/91	19.6	Cr
Senegal	1. Rural Health Project	11/29/82	15	Cr
	2. Human Relations Development Project	3/12/91	35	Cr
Sierra Leone	Health & Population Sector Support Project	4/22/86	5.3	Cr
Sri Lanka	Health & Family Planning Project	3/25/88	17.5	Cr
Thailand	Population Project	1/18/78	33.1	Ln
Togo	Population & Health Sector Adjustment Program	1/23/91	14.2	Cr
Trinidad & Tobago	Population Project	4/28/71	3	Ln
Tunisia	1. Population Project	3/3/71	4.8	Cr
	2. Population & Family Health Project	2/19/91	26	Ln

Yemen Arab Rep.	Health Sector Development Project	5/7/90	15	Cr
Zimbabwe	1. Family Health Project	6/20/86	10	Ln
	2. Second Family Health Project	4/30/91	25	Ln

Annex II. Synopses of "Targeting" in Past Reports (1970-91)

Bangladesh (1975-91)

Four successive five-year projects in a country with an alarming P problem, a government with a long-standing commitment to ZPG, a high consensus on P goals, and a strong sense of urgency among government and donors.

Pop I ('75): By far the most explicit and comprehensive targeting/evaluation coverage of the 75 SARs. Four detailed annexes, totalling 47 pages, describe base-line and progress indicators and their generation. Indicators, however, are to be used to measure how much progress is being made towards long-term program goals and not as operational targets for the project. Strictly speaking, the project does not make use of output targets (demographic levels or acceptor yields). There are to be separate evaluations of the national program and of project activities in four pilot areas. Base-line and follow-up KAP surveys are to be conducted. Fertility changes are to be measured by changes in ASFRs and TF which are to be determined by special surveys and cross-checked with vital registrations (births and deaths) and estimates based on acceptor statistics. One annex describes and evaluates the service statistics system and needed improvements. Another annex (#26) analyzes the feasibility—with the government's present program - of recruiting enough acceptors, and averting enough births, to reach the government target of a NRR of 1.0 by 2000 and reducing the '75 CBR of 47 to 43 by '78. The analysis shows the government's demographic targets to be unrealistic even with maximum performance of the project. The only hope of reaching the targets would be to extend the intensified service and motivational systems throughout the country, to integrate FP with MCH care, and to increase system efficiency—tasks for future projects. Internal and external task forces will evaluate the contribution of each major project component to the program.

Pop II ('79): Useful quantitative review of program achievements to date, with international comparisons (p. 12 ff.). By 1978 the program had succeeded in recruiting almost 50 percent more acceptors than targeted (note that no targets were in the Pop-I SAR; they were set subsequently by the government). The absence of continuation-rate data prevent any estimate of demographic impact, however. In Pop-II, evaluation is handled by appointing an inter-agency government committee which is to "prepare in consultation with the Association a comprehensive plan of action for evaluating the population and health programs and project including its various components . . . This plan will: (a) define the measurable objectives of each component; (b) identify corresponding quantitative performance indicators; (c) ascertain the availability of baseline values of these indicators; (d) determine the data to be collected; and (e) identify agencies for carrying out specific studies. Illustrations of performance indicators and timetable for evaluation studies are given in Annexes 4 and 5" but, alas, are illegible because of excessive photo reduction. As in Pop-I, no setting of output targets in the SAR (although government continued to make heavy use of them).

Pop III ('85): Compares program performance with targets during Second FOP (80-85) and concludes that the major shortfalls partly reflect unrealistically high government targets (demographic, FP, and health). Mid-term Bank and donor advice led to adoption of more realistic targets. Long-term FP and demographic projections are made, and "estimates" are given (Annex 4) of the number of acceptors, by method, by year, for the period '85-90. Those (Bank) estimates constitute *de facto* program targets. The project commits the government to conduct CPR surveys in the first, third, and fifth years of the project (weak knowledge of continuation rates was referred to in Pop-II). Because the project finances "most of the major activities of the national family planning and MCH program . . . its benefits cannot be separated from those of the program as a whole" (6.01).

Pop IV ('91): Again, the Bank feels that Fourth Plan demographic and FP targets are too high—. . . based on past experience and the maximum achieved by some of the best performing FP programs in the world, more realistic targets would be a CPR of 45-50 percent, a CBR of 30-35 per 1,000 and a TFR of around 4." The government's quantitative MCH and health targets are stated (1.39) and presumably accepted although the only quantitative health target used in the statement of project objectives is to reduce maternal mortality from "the current 6-8 to 4-5 per 1,000 live births" (2.1). Annex 17 is a useful one-page summary of "satisfactory understandings" reached between the Government and its donors on "demographic targets, the contraceptive mix, the FP and MCH service balance, and incentives." Annex 21 is a table showing the government's six-year projections, in percentage terms only, of the distribution of acceptors by contraceptive method; a fn. notes that these projections are revised semi-annually and "Major revisions tend to occur" By now, the Bangladesh Fertility Surveys (permitting estimates of trends in ASFR and TF rates) and CPR surveys appear to have become institutionalized. Also, there has now been established (starting with Pop-III) an elaborate series of annual technical and program reviews with the donors. Indicators to be used for these reviews are specified and include the CPR, TFR, IMR, and EPI (immunization) coverage figures. It seems fair to say that the project puts greater emphasis on monitoring the trends of key indicators than on the achievement of specific targets.

Botswana (1984)

The government has no quantitative fertility objectives but does want to reduce the population growth rate and therefore wants to make FP advice and materials available to all potential parents. This health project will be evaluated (including its FP impact) "by existing information systems which the project will strengthen." USAID will sponsor a CPR survey at the time of project start-up and is expected to fund a second one during the project's final year. Without establishing any quantitative targets, the project's improved MCH/FP services "should result in: (a) an increase in the numbers of acceptors of family planning; (b) more women breastfeeding; (c) longer birth intervals resulting from encouragement of breastfeeding and use of contraception; and (d) decrease in infant morbidity and mortality and healthier mothers, due to birth spacing and improved child health" (6.02). No quantitative targets but trend monitoring of FP through the use of standard acceptor data from service statistics and CPR surveys.

Brazil (1986 & 1989)

The two very large basic health services projects in the NE are directed to expanding and strengthening the health system, with emphasis on PHC as the primary means of improving maternal and child health. The '86 SAR noted that population and FP are sensitive subjects and while the government has acknowledged FP as a basic human right it "refuses to set up targets on

fertility reduction, in order to avoid political conflict." The project will therefore support the implementation of family planning services through the government's MCH services, "aiming at meeting the spontaneous demand, and promoting child spacing for health reasons, and at creating the infrastructure and organization for the eventual expansion of these services" (2.21). Annex 5 lists seven "indicators" (=targets) to be reached "after five years of service"; only one of these relates to FP and says only that in two of the project's four states the system will "have met the request of contraceptives to 100 percent of women, except for those with a medical or legal contraindication." The '89 project simply extends the basic health system farther and improves its functioning, with, again, no reference to demographic or FP targets, nor to service statistics.

Burkina (1985)

The country is in the initial stages of developing a population program, demographic analysis, and FP services. The project will help establish all of these. As for monitoring, the project will establish a number of baseline health indicators, including "the number of family planning acceptors."

Burundi (1987)

An initial Bank project in a country with a severe population problem. The promotion and development of birth spacing and FP, and the improvement of the demographic data base for population policy development, are two of the project's five explicit objectives. Each project component "has a full set of quantitative objectives, including outcome and process targets." These include: "(a) reducing maternal mortality by 20 percent; (b) reducing infant mortality from 1215 per 12,000 to 100 per 1,000; and (c) increasing contraceptive prevalence from 1.6 percent to 14 percent" (by 1992 - in five years). These government targets (called "projections" but viewed as targets by both government and Bank) are shown in a table (T 12, Annex 1) listing FP users, the CPR, the TFR, and births averted for 1988-1992. T. 13 shows the projected contraceptive method mix. The basis for the acceptor target is explained in 3.17. P. 18 of Annex 1 lists 17 "Key Project Indicators," grouped into three classes—"Process Indicators," "Outcome Indicators," and "Impact Indicators." [This nomenclature is not standard in Bank practice. Its use here is consistent with Fig. 1 in the main report].

Colombia (1985)

This initial project is in a country with quite good data on fertility and contraceptive prevalence (there were two National Fertility Surveys, and one Contraceptive Prevalence Survey, between 1969-1980). The project will fund a new National Contraceptive Prevalence Survey to update the '80 figures, and a fertility component in a 1985 National Health Survey. Targets for total and new acceptors were set for each of the project's nine health regions and sub-regions. These targets "are based on the Government's quantified objective of reducing total fertility rate (TFR) from 4.3 to 4.0. The figures imply an increase in the contraceptive prevalence rate (CPR) from about 37 percent to about 50 percent." Quantitative health targets are also specified.

Comoros (1983)

Government just beginning to show concern for population matters, which are sensitive and risk political and religious backlash if pressed too rapidly or forcefully. But FP services have been permitted as part of MCH services for three years, although a FP Division, which the project

will support, has just been created. There are no reliable data on contraceptive prevalence, estimated at about 3 percent. Likewise, no data exist on dropout rates. As an initial system-building project in a weak environment, its principal objective is expressed in these words: "By the end of the project the MPHP should be in a position to provide good quality basic health—especially MCH—and family planning services in all primary level facilities . . . It would have the appropriate facilities, drugs, equipment and procedures in place, and would be carrying out a plan for the optimal deployment and training of staff." The Bank's own population projections show that if the CPR can be raised from the present 3 percent to 22 percent within 10 years (implying the recruitment of 9,000 new acceptors annually by 1992), the rate of population increase could be slowed from 3.2 percent in 2005 to 2.3 percent—"a realistic objective for the Comoros" (2.06). With a stronger but "still attainable" effort, the rate could be brought down to 1.8 percent. These are targets the Bank is suggesting to Government

Dominican Republic (1976)

The review of the government's early-70's CBR and CPR targets shows them to have been, even after downward revision, much too optimistic. A new system of service statistics is a considerable improvement and will be relied on, with new health statistics which the project will introduce, to evaluate program and project performance. Evaluation through service statistics will be supplemented by surveys to establish numbers of acceptors, continuing users, active clients, and couple-years of protection, and to allow estimates of fertility declines based on ASFRs. "Program efforts," or inputs, will also be evaluated (e.g. personnel-hours, cost, contraceptives distributed). The project concentrates on improving motivational work and service delivery in two Regions. Estimates of FP users and CPRs in the two Regions show nearly three times as many FP users with the project as without it (within five years); no attempt is made to make quantitative estimates of changes in Regional fertility, however.

Egypt (1973 & 1978)

The Bank's two projects (in '73 and '78) encountered a difficult administrative environment and did not succeed in establishing a continuing Bank/borrower relationship. The first found a "middling" CPR of 13-14 percent, a weak service-statistics system, and a disappointing growth-rate of new acceptors. The government had no recruitment targets; instead it operated a staff incentive scheme whose weaknesses made targets look a more attractive system. The project did not establish any targets, however. Instead it provided for three operational research studies intended to suggest ways of making the program more effective.

The '78 project paid much more attention to targets, monitoring, and evaluation. Starting from a government demographic target to reduce the CBR by one point per year over a 10-year period (i.e., from 34 in 1972 to 24 by 1982) the SAR works out targets of required acceptors for the nation and the project Governorates. The result was an implied CPR of 30.5 percent (up from an estimated 18.2 percent in 1977) and a need to recruit about 3.5 million new acceptors ('78=82). The figures were also disaggregated by project Governorates and districts, taking account of their considerable differences. The derived results were considered "ambitious" and doubtful of attainment; a more realistic target would be to allow two more years for doubling the baseline numbers (3.13). To monitor progress towards these targets, the project established a new Research and Development Unit in the FP Dept. of the MOH. Among the Unit's functions would be (a) helping the MOH's new Statistical Unit (specially established to overcome the serious inadequacies of the existing FP service statistics) design a management information system for FP, (b) helping the

program director set FP targets in terms of continuing users and new acceptors, and (c) evaluating data on fertility and MCH in the project Governorates. The R&D Unit was also given funds to mount an external research program, to be conducted by Egyptian institutions, focused on "primary health care with emphasis on MCH/FP." Nine specific priority studies were listed.

Ethiopia (1988)

This is a pilot project to improve the health status primarily of mothers and children in two regions. Specific health targets are set, as are simple FP targets (a) to provide FP services at all health facilities as well as on a community-based distribution basis, and (b) to increase the CPR to 17 percent from the current 2 percent (over the project's eight-year disbursement period, presumably - not specified). A target distribution of the contraceptive mix is also specified (3.06 and 3.09). While "primary emphasis in monitoring and evaluating the project would be placed on output and process indicators . . . management strengthening and the reactions of project beneficiaries would also be evaluated systematically" (through an independent management review and sociological assessments, the latter to test beneficiaries' KAP on FP, health, and nutrition (5.07, 5.08).

Ghana (1985 & 1990)

This initial Bank project in either of the two main social sectors is intended—as part of the Bank's support for Ghana's economic recovery program—to provide "emergency assistance" to sectors that have been slipping backwards for several years. In health, the principal aim is therefore to revive a moribund, ineffectual delivery system. Within health, "one of the major obstacles to an effective family planning program would be addressed, i.e., convincing health workers that family planning is an important element of maternal and child health services." The project projects that the CPR can be increased from around 5 percent in 1985 to 16 percent by 1988 (with the commercial sector contributing seven percentage points, NGOs three, and the MOH six—para. 35). The 1990 project (which followed a 1988 PHN sector review) identified, as among the major FP program issues, the need for "broad target-setting and active monitoring of progress" and "adoption of an explicit objective of increasing the number of family planning acceptors/users as fast as possible, and of specific targets for the program" (2.06). Annex 2-9 presents the Bank's own longer-term (1995 and 2000) targets for both population and health. The key population target is a CPR of 21 percent by '95 and of 31 percent by 2000; no demographic targets are specified.

Guinea-Bissau (1987)

This is an initial PHN project in a weak, early-stage system. Hence major emphasis is on basic institutional development and the strengthening of a PHC delivery system. No quantitative demographic or FP goals are mentioned. The only explicit FP goals are to conduct (a) a FP study aimed at developing guidelines for identifying reproductive-risk women (3.14) and (b) an unelaborated "special policy study" on family planning (one of four such studies - 3.02). Support for in-service training will include material on FP.

Haiti (1989)

The Bank's initial assistance to a chaotic, poorly-financed, administratively top-heavy, urban-biased health system aims primarily to introduce organizational and policy rationalization as a basis for more effective, more equitable future service delivery. Thus the project's contribution to strengthening FP during the project period will be extremely modest; the main hope is that the

measures taken to reform and strengthen the MCH/FP system will "lay the basis for intensifying FP" in the future. A government population policy paper was due for completion by March, 1990. No demographic or FP targets are specified. Annex 3 lists six "process outputs" to be monitored and three "outcome outputs" for which targets are set: the only FP activity to be monitored is the number of sterilizations performed on (a) men and (b) women.

India (1972-1990)

The Bank has been assisting the Indian population program continuously since 1973 when the first of its seven projects was approved. These SARs will be reviewed together, using Roman numerals I-VII and para. numbers to refer to individual reports and projects. Each of the seven projects addresses the problem in a different, well-defined area where, through the project, intensive efforts are made to improve program performance. All seven SARs begin by reviewing the national program—GOI policy, demographic and acceptor trends, broad program strategy and organization, major problems that exist, arrangements for monitoring and evaluation and the underlying data-base, and similar relevant background topics. Each project then develops a strategy for extending and improving the program within its chosen area. The strategy may be the government's or a mixture of Bank-suggested and government strategy for intensive application in the project area. Throughout this 18-year period, Bank reports have shown a strong interest in monitoring and evaluation, including the making of projections, target-setting, and operational research (research studies to see what works well, what less well). A constant Bank concern has been to deepen and improve the quality of the data-base on which performance judgments rest.

India-I was so impressed with the inadequacies of the Indian program's data-base that it made the establishment of two independent (i.e., not university-attached, like some 16 others) Population Centers a major part of the project. These were charged with designing an improved management information and evaluation system (MIES), of which service statistics would be a centerpiece, and to engage in program-relevant demographic analysis and operational research. This was an attempt at major institution-building in support of monitoring and evaluation. In addition, SAR-I contains an impressive 18-pp. annex on demographic and acceptor targets for each of the two project areas. It is a model of cautious, explicit projections that show three possible outcomes depending on differing assumptions. Indeed, as is so often the case with low-medium-high projections, only the medium one deserves to be called a target (low projections normally assume little or no performance improvement, only a slow expansion at existing efficiencies, while high projections are usually based on assumptions more optimistic than it is reasonable to expect). The GOI has long made strong use of targets, in the sense of "this is what your administrators expect of you." Bank SARs never show a comparison between its area targets and those set (mainly by administrative fiat) by government administrators. Presumably (because Bank projections are more realistic and less hortatory) their targets would almost always be lower than the Government's.

The two major Population Centers did not develop as rapidly, or with as much operational relevance and support, as originally hoped. SAR-II reported that after six years they had become well established but SAR-III complained that they seemed somewhat isolated and academic in their work; SAR-IV credits them with having made some valuable operational studies and with having improved the MIES in the two states they were meant to serve. This is the last we hear of this major innovation—although all projects talk about local data inadequacies and support early and late surveys to provide area data against which performance judgments can be made. A complaint made in several of the projects is that almost no data exist on continuation rates, making it impossible

to translate acceptance data into demographic impacts except by use of norms imported from other countries.

By 1983 (Pop-III) two major Bank and GOI concerns with the program were (a) the need to give greater emphasis to MCH activities as a basis for increasing the demand for FP and (b) that the national program had over-emphasized sterilization and had not given enough emphasis to the temporary methods that might attract more younger couples. Simple MCH targets therefore appear, alongside the former acceptor and demographic targets/projections. In addition the broad area targets now reflect the desired shift in method-mix. Pop-IV continues with method-mix targets and, like some of its predecessors, sponsors some baseline area surveys for later performance measurement—and which "will provide more precise values for the project indicators [a disappointingly summary one-page annex is repeated from earlier projects . . .] on the basis of which agreement will be reached with the Bank on this project's quantitative goals." It is not clear how these post-project goals relate to those already expressed in the SAR.

Pop-V is an urban project but the targeting/monitoring/evaluation problems are familiar ones. A text table presents input and output targets in exceptionally clear format. Input targets are limited to the expansion of physical facilities (nothing on staff numbers, training courses, IEC work); outputs are expressed as "service delivery achievements" and refer to a CPR target, a desired shift in method mix, and coverage with simple MCH procedures. Once again, the project will fund baseline and end-of-project surveys to generate performance data.

By the end of the '80s the GOI and the Bank were both convinced that improving the quality of MCH services could play a key role in increasing the demand for, and effectiveness of, FP services. In Pop-VI and -VII staff training is viewed as the key to quality improvements. In Pop-VI three scenarios are worked out to show the effect on the TFR of differing assumptions about how improved training may affect acceptance rates, the method mix, and continuation rates. The results have to be taken on faith as the tables are not backed up by explanations (indeed, the annex numbers do not seem to agree with those used in the text).

SARs VI and VII each reflect a feeling that both the Bank and the GOI give too much emphasis to quantitative acceptor and demographic data and not enough to qualitative judgments and operational research (1.13). SAR-VII notes that "the performance indicators which are most heavily relied on include estimates of target achievement. These tend to over-emphasize the achievement of method-specific (sterilization and IUCD) quotas and under-emphasize other dimensions of the revised program strategy, particularly the delivery of priority maternal and child health services. The project would assist the State in developing improved measurements of program outcome, and train workers and managers in the use of these measures and the interpretation of the data generated." Areas within the five states that develop alternative indicators are to be released from "present target-setting procedures" (2.36). The SAR nevertheless contains three-scenario projections, although these end with a cautionary note that "projections of increases in contraceptive prevalence are difficult because . . . current continuation rates for temporary methods are unknown, and also because changes in the number of acceptors from year to year for all methods have been quite erratic." The most one can say, then, is that if the project is able to achieve its process objectives, prevalence and fertility outcomes should come correspondingly closer to scenario B than to scenario A" (p.136).

Indonesia (1972-1991)

The Bank's five projects reflect a continuous association with this highly successful program since it moved into high gear about 1970. The Bank's initial help was precisely to help the government set targets for the program: in late 1969 the Bank joined the UN and WHO in a joint mission to suggest the scale and type of program the GOI should organize. The mission's report suggested the government should aim at recruiting 6 million new acceptors between 1971-76; this was double the target of the then-current five-year plan. The government accepted the figure (even increasing it slightly) and developed a program that followed closely the three-agency report. Thereafter the government has set annual new-acceptor targets, steadily raising them, with performance consistently beating them.

The first ('72) Bank project devoted considerable attention to starting up evaluation and research studies, and supported needed continuation-rate and KAP surveys (the best summary of P-I's E&R work is in para. 4.07 of SAR-II). A six-page annex in SAR-I (no. 38) provides one of the clearest explanations anywhere of how long-run acceptor targets should be set. After this initial project, however, the next three SARs show far less concern for evaluation, research, and targeting than one finds, for example, in Bank projects in such other major countries as Bangladesh or India. Concern for these matters returns with a thunderclap in SAR-V, however. A principal explanation of the relative silence of projects II, III, and IV on these matters is that other donors were much concerned with them and Bank "neglect" of them reflected a division of labor.

The Bank was not completely silent on targets after 1972. In the second (1977) project, the Bank stated its agreement with the GOI's CBR target of 34 by '84 and with the very large acceptor targets this implied. The Bank justified its acceptance of these high targets by referring to the program's very strong recent performance. The project also supported an interesting experiment with a "community incentive scheme" in 14 kecamatans to see if the offering of communal prizes (more money for community development projects) would be effective in raising local CPRs to target levels (a 16-page annex describes the experiment). Later ARs make no reference to this experiment, and the subsequent development of a community-based approach to recruitment and supply (perhaps the distinctive hallmark of the Indonesian program) appears not to have been modeled on it.¹³

Pop-III ('80) contains an excellent short review of acceptor targets and program performance during the '70s. The project notes and accepts the GOI's demographic target of cutting in half the country's 1971 fertility rate by 1990; it also accepts the 50 percent increase in acceptor targets for '80-84 as compared with '75-79. But it is only in SAR-IV ('85) that we see interest in looking at the effectiveness of high acceptance rates—the project commits the government to conducting biannual surveys to establish CPR and continuation rates (Prof. Freedman notes that the continuation-rate figures accepted by government and donors in the mid-80s appear to have been serious over-estimates).¹⁴ The project also reflects some disappointment with the performance of

¹³See Prof. Ronald Freedman's discussion of the communal structure of the program at pp. 15-16 of his report for OED, *The World Bank and Indonesia's Population Program* (February 13, 1991, 48 pp., Bank rept. no. 9370).

¹⁴"The information system, which produced credible estimates of prevalence until at least 1976, subsequently greatly exaggerated the rise of prevalence rates, especially between 1982-83 and 1985-86. Exaggeration of prevalence rates by 15-20 absolute percentage points resulted in gross overestimates of resources needed for maintenance and gross underestimates of new users required to reach targeted levels. The illusion that abnormally large reported prevalence increases were real created a temporary euphoria that was misleading about what was possible in short time periods." (*op cit.*, p. 19). A Bank review in 1986 analysed this problem in some depth but Freedman says that changes made since then have not prevented the system from generating prevalence rates that are still misleadingly high.

Indonesia's demographic institutions and provides funds for the training of more demographers. It also provides major funding for three new research centers (for FP Studies, FP Policy Development, and Biomedical and Human Reproduction) to be established by the National Family Planning Coordinating Board.

As noted, Pop-V ('91) shows a greatly heightened interest in evaluation and research matters. The interest is not primarily with the nature or validity of GOI demographic and acceptor targets. It is directed instead to the evaluation of project components and to the development of indicators for routine "process" and "performance" monitoring for each project component. ". . . Recent evaluations," the SAR states, "draw attention to lack of adequate indicators as to what targets project components should meet; in the absence of such targets, performance judgments become rather subjective" (2.09). Perhaps drawing on the Bank's work in Bangladesh, P-V introduces the use of Annual Implementation Plans, which are to be submitted for Bank review before a new Annual Review Mission every October. The Implementation Plans (to be prepared by each of BKKBN's project-implementing units) are directed to include "performance and process indicators." To help Indonesian staff develop appropriate indicators, SAR-V contains four annexes (nos 6-9, totaling 27 pp.) which suggest how the "process indicators" problem might be handled for each major project component. This seems a daunting task, and one may fairly ask if this kind of detailed tracking and reporting (which goes considerably beyond normal project implementation reporting) is worth doing. Such monitoring asks if specific tasks are getting done; it does not evaluate the design, impact, or cost-effectiveness of tasks and whether they deserve to be continued, modified, or dropped in favor of possible other means of accomplishing an objective. In short, component monitoring should not be mistaken for operational research: the Bank needs to give more attention to this distinction and to what types(s) of evaluations it and its clients most need (or what the balance should be between them). Pop-V also provides for external technical assistance (involving a two-year residency—TOR in annex 10) to review the program's management information system (MIS), essentially the service-statistics system. With the program now undergoing some major changes in direction (including a new emphasis on privatizing many activities, and stronger attempts to serve hard-to-reach groups) it is felt that the present service-statistics system may no longer provide the routine data the program should be getting.

Iran (1973)

This project (later canceled) does not contain any output targets of any kind. There is a review of the service statistics system, and comments on the weakness of evaluation and research activities, but nothing more.

Ivory Coast (1985)

This initial project intends to help convert a recently pro-natalist, anti-FP country with almost no demographic data into one with basic demographic data and a remodeled MCH system (but without FP services, although nurse training is to cover that topic). The project will fund the country's first census. The main population "targets" are to produce basic data, encourage a shift in government population policy, and bring the MCH system to a stage of development where it can begin to offer child-spacing services.

Jamaica (1970-87)

This series of three projects shows modest, explicit, and increasing interest in output targets. P-I was the Bank's first population operation. It takes note of the GOJ's hope that the CBR, which was 35 when the program started in 1968, can be reduced to 25 by 1975. In a summary and somewhat simplistic paragraph, the SAR makes use of acceptor data from neighboring Barbados and notes that if the government's acceptor targets are met it could result in a high enough CPR to meet the 1975 demographic target. The project provided for an annual external review during each of its four years "to assist the government in establishing and maintaining high standards of cost-effectiveness, i.e., so that the Board, the government, and the Bank will have an independent judgment on whether or not the program is being conducted at acceptable levels of efficiency and to identify specific problems requiring special attention." An annex listed 24 specific questions these reviews might consider; these included looking at the establishment of local acceptor targets and at efforts to establish continuation rates.

Pop-II contained funding for the calculation (from FP service statistics) of ASFRs annually, 1976-80. Funds were also provided for studies of service statistics (and follow-up surveys, if needed) of method-specific dropout rates, which were not yet known.

Pop-III avoids taking a position on the GOJ's very ambitious target of raising the CPR from the 1983 figure of 43 percent to 70 percent (a very high figure, by world standards) in 1991. The SAR says only that "The project will support MOH efforts to improve the family planning service delivery programs in support of the Government effort to increase contraceptive prevalence from 40 percent in 1983 to around 70 percent by 1991." A mid-term project evaluation by the MOH "will include assessment of the contraceptive prevalence rate," which can be compared with results of the 1983 CPR USAID-funded survey. A one-page annex explains the relationship between CPRs and fertility projections, which are presented to 2005.

Jordan (1985)

This initial PHC project provides reasonable data on the amount and sources of FP services available in an environment where the government has no population policy and FP is still a sensitive subject and is normally referred to in terms of child-spacing as a private health benefit. The MOH does offer contraceptive services as one of its MCH services, although the private sector is the major source of supply. Most of the project's targets are health targets; among these are a specific target of raising the MOH's share in the CPR from 1 percent to 6 percent. If this is reached, the SAR estimates that the overall CPR will rise from 26 percent to 31 percent. But that remains a Bank projection, not a project or even a government target.

Kenya (1974-90)

The four Kenya projects between 1974 and 1990 provide an ironic lesson in target-setting: all the acceptor targets in the first three projects were wildly missed. Those of the first two projects were not nearly attained, and generated a widespread mood of disappointment and frustration among government and donors that the ambitious program just was not working. Then suddenly, unexpectedly demand began to explode around the end of 1987. By the end of 1988 the carefully set targets for that year had been greatly exceeded—and fears developed that the country might run out of supplies. P-IV in 1990 is essentially an emergency contraceptive supply project to make sure supplies would be adequate to accommodate the unexpected explosion in demand. Almost a decade and a half of unexpected disappointment, followed by three years of unexpected success. Targets provided psychological baselines that fed these mood-swings (service statistics, providing routine

monitoring measurements, obviously push in the same directions but without the fillip of target-comparisons).

Pop-I was based on Bank acceptance of GOK health, demographic, and acceptor targets (4.02); indeed the Bank had helped prepare them. Annex 23 of the SAR then worked out the favorable demographic outcomes which attainment of the acceptor targets would mean, with their benefits to the economy. P-II noted the program's disappointing performance and offered some reasons to explain it (3.17-3.19). Pop-II was designed to try to overcome the perceived constraints. Acknowledging the difficulty of setting targets for a "stalled" program, the SAR nevertheless said it would "not be unreasonable to expect an increase in the contraceptive prevalence rate from about 5 percent now to between 10 percent and 17 percent by the end of 1987 This would be reflected in a decline in the CBR from about 53 per thousand now to between 51 and 48 by 1987, approximately" (4.04). No disaggregated regional or district targets appear in the SAR.

SAR-II contains an excellent (and the only) definition of the three types of targets which a project may contain—targets for inputs, processes and outputs (7.12).

Pop-III (1988) noted the program's disappointing performance over the previous decade, blamed it on a gap between government and private perceptions of the benefits of FP and on lingering reluctance to discuss population matters openly, and then went on to establish, for the new project, a more detailed and explicit set of targets than either of its predecessors. The P-III targets (Annex 3) use a model developed by Bongaarts and Stover at the Population Council to work out district-by-district targets, by contraceptive method, for each of the four project years (1988-1991). This is perhaps the most explicit and most detailed set of acceptor targets to be found in any of the 75 Bank SARs. In addition to setting acceptor targets, the project gave considerable attention to strengthening the institutional capacities for population research, and to the development of mechanisms to conduct "comprehensive evaluations of GOK and NGO population program activities." Independent evaluations of program sub-components are to be conducted, by local consultants and organizations, in the second and fourth years of the project. This, at minimum, was a token bow to the need for operational research (2.09-2.10)

SAR-IV has the pleasant task of reporting the dramatic program turn-around which had resulted in making the P-III targets obsolete even before that project was out of the starting gate. SAR-IV gives no acceptor targets but presents a table of "contraceptive usage, 1990-1994" the underlying assumptions for which are in Working Paper I in the project file (not in an SAR annex). The project provides for biannual surveys to monitor the CPR. The project also provides for annual work plans "based on a component-by-component review of progress in the preceding year." A mid-term project review will also be conducted—an operational review, if not operational research

Korea (1979)

This project, in a country with an outstandingly successful FP program, contains explicit output (demographic and acceptor) targets. The aim is to reduce the CBR from 24.3 in '76 to 23.9 by '81 and to 22.9 by '86. "To do this it is necessary to increase the number of current users of contraceptives from 2.0 million in 1976 to 3.1 million in 1981 and 3.9 million in 1986." The report notes the existence of effective domestic demographic, FP, and health-research institutes capable of generating and analyzing the data needed for setting targets and monitoring performance (2.35). But since the project will establish some new service-delivery systems for which a new reporting system, and end-of-project evaluations will be needed, funds are provided to support these activities (3.18).

Quasi-targets are set for the new MCH facilities which the project will establish: "By the end of the project period, the MCH centers should be able to handle about 150,000 births a year or about 30 percent of the number of births in rural areas. In addition, the multipurpose workers would have the capacity to help in about 200,000 deliveries at home. This compares with direct attendance for only about 27,000 births at present. The services provided by these workers . . . will reach at least 50 percent of the rural population. It is estimated that about ten million people would be provided primary health care services under the project" (5.01).

Lesotho (1985 & 1989)

Although this southern African country does not yet have a formal population policy, the government clearly recognizes the need to control population growth, has made several strong statements in support of FP, and will use the project to strengthen the provision of FP services as part of its MCH services. No demographic targets have yet been set by national planners; the SAR does not go beyond the inclusion of normal Bank three-scenario population projections. The SAR estimates that this MCH/FP/Health project should result in an increase in the CPR from 5.5 percent at present to 8 percent by project completion. Beyond this, the report notes only that the project's combined components "would all lay the groundwork for a more substantial impact on fertility in a subsequent project" (6.02).

The subsequent project came along four years later (1989). It again uses standard Bank projections to estimate future population size (in 2000 and 2050) "If Lesotho is successful in promoting an accelerated decline in fertility While it will not be easy for Lesotho to pursue this latter scenario, the experience of other developing regions suggests that it is achievable" (2.10). The Fourth Plan ('87-91) sets a target of raising the CPR from under 10 percent to 25 percent. The SAR, citing recent rapid increases in contraceptive use in neighboring countries, comments that "this target is probably achievable" (2.12). Later the report estimates that the project will deliver FP information "and contraceptive methods" to about 60,000 women by the end of the project period (6.03). Project impact is to be evaluated by three studies (covering FP, food and nutrition, and disease control) to be done late in the project by independent local and foreign consultants (TOR are left for later). Baseline data are not entirely absent but are weak Input targets for FP training are explicit, and a date is specified by which the Government is to finalize its population policy (3.13). The SAR contains no discussion of service statistics but the project will fund a "Health Information Systems Advisor" to review and make recommendations on "MOH's basic data needs."

Madagascar (1988)

This is not a PHN project but the operation does address population and FP as part of a structural adjustment operation. The Government's Letter of Development Policy notes that the Government "is resolved to promote greater consistency between the economic growth rate and the population growth rate" and that a new Population and Development Unit in the planning agency will submit to Government, by July '89, a proposed population policy. The only FP services now being provided come from a Non-governmental Organization (NGO) to which the government "proposes to extend assistance . . . to enable it to improve and extend its operations" (p.51). With the help of this assistance "the recruitment of 15,000 new clients" by the FP NGO would be possible. That one output target is matched by specific input targets, among which are the training of 250 motivators and 750 health professionals to be trained in FP clinical skills (3.10). Such is this initial operational nudge to a country in the very early stages of dealing with an admitted but still sensitive problem.

Malawi (1983-1991)

Malawi is one of several African countries which, partly as a result of dialogue with the Bank, have begun to recognize their population problems and have moved to introduce FP as a child-spacing program within their MCH services. Pop-I accepts the Government's Child Spacing Plan and supports the introduction of FP services, but without any targets. ". . . The project will not directly lead to any amelioration of health problems or reduce health indicators to acceptable limits; without the project interventions, however, the Government would not be able to do this effectively in the future" (6.01). Input targets for system expansion (numbers of facilities at which FP services are to be offered, staff numbers to be trained) are specified (3.30,3.31).

Four years later (1987) Pop-II noted the evolution of government and public attitudes, the "rapid and unforeseen growth in demand" for FP services, and the over-fulfillment of the training targets in Pop-I. The second project was therefore able to deal explicitly with future output targets. With an estimated present CPR of 2 percent and a TFR of 7.7, the government target of raising the CPR to 10 percent by 1991 would bring the TFR down to about 7, the number of FP users would increase about 45 percent to 55,000, and population in 2015 would increase only 70 percent, not the 200 percent if fertility were to remain unchanged. A table presents the estimated numbers of users, by method, for 1987, 1988, and 1989. "A paper detailing how these targets were calculated is available in the project files" (by R. Bulatao). Performance towards goals will be monitored by a number of "process indicators" [specified] as well as "several surveys and studies" [specified]; they are to include "annual impact studies of various programs"—unspecified.

Pop-III (Feb. '91? - date omitted from SAR) notes that although the government has no explicit population policy, it has begun to more actively promote its child-spacing program, now quite widely available. However, the Bank had come to feel that the Pop-II CPR target of 10 percent by 1991 (a Government target which the Bank had apparently accepted) was now unrealistic and should be postponed to 1995/96. Six percent should be taken as an "interim" 1992 target. To monitor progress towards its contraceptive target, new family formation surveys will be undertaken (one in April 1992, another three years later). No review of service statistics is to be found in this or in either of its two predecessor SARs.

Malaysia (1972 & 1978)

Two projects were all the Bank FP assistance the government sought before sensitive ethnic rivalries ended earlier interest in slowing national population growth. Although Pop-I had very explicit demographic targets, set by the National FP Board, and global acceptor targets, there was not yet any FP strategy that related program activities to the Board's targets. An extended annex argued that the Board's acceptor targets were too low; the project would permit raising them by 40 percent. An important part of the project was the selection of 14 districts as "Intensive Input Demonstration Areas" to test the ability of special efforts to recruit higher levels of acceptors. No targets for those areas appear in the SAR, however, and although there is clear assurance, there is no description of the "detailed evaluation of this demonstration" that the project will conduct, beyond a reference to "special management and research studies and surveys by the research advisers and a management consulting firm" (annex 5, p. 2). Since all research on Malaysian population problems had been done by expatriates, local evaluation/research capacity was extremely weak; the project provided support for its strengthening (4.01-v).

Pop-II (1978) accepts the government's demographic and acceptor targets but notes that their validity will depend on a shift from less to more-effective contraceptive methods. The CBR is to be brought down to 26 by 1985, which will require a CPR of 44 percent (vs. c. 40 percent in '74-75). The improvement of the method mix is assumed to require a shift to a higher proportion of acceptors in the government program and a decreased proportion from non-program sources (2.21). The project funds a built-in evaluation system "on the basis of quantitative indicators specified by [the government]" ["a listing of these indicators is given in Annex 3": these list eight project components, the "objective" for each, the "quantitative indicators" to be used, and the "status and sources of measures." Nearly two years of technical assistance is provided for project evaluation work. Additional TA, plus other funding, is provided to build up the Research, Evaluation and Management Information System—relying heavily on recommendations of the External Review Team under Pop-I (4.61-4.66).

Mali (1983 & 1991)

This was the first Francophone country of W. Africa to accept FP as part of its MCH services. In 1983 the CPR was estimated at 0.2 percent. UNFPA had recently helped establish a population unit in the Planning Ministry to work on development of a population policy. Pop-I is overwhelmingly a primary health care (PHC) project; as MCH services are improved, and the high infant mortality rate declines, increases in FP will be slow. No demographic or acceptor targets are mentioned. Most of the project's "process indicators" (Annex I, T. 7) are health indicators; among them will be "the number of FP acceptors."

Pop-II (1991) shows a much stronger Bank and government interest in population. Although primarily a health project, it aims to raise the national CPR from 1.2 percent in 1991 to 8.5 percent by 1997. A draft population policy from the planning agency (1990) proposed demographic targets that were unrealistically high; a more modest operational plan from the MOH is judged actionable and is accepted (2.33-2.34). The project, which will reach nearly one-third of the population, aims to raise the CPR in the project area from 1 percent in 1989 to 10 percent in 1997 (3.53). A 29-pp. Annex on "Population and Family Planning" includes a table, based on the "World Bank standard fertility decline scenario," on the projected number of users, by method, 1990-2000—the last page of an annex overwhelmingly concerned with the program details of the project's FP component.

Morocco (1985)

While the government recognizes the drag of too-rapid population growth, it has shied away from explicit demographic targets. It has, however, provided strong support for FP as part of its MCH services, which have clear FP objectives. Within the project area (three provinces), the target is to raise the average CPR from 22 percent in 1985 to 33 percent by 1989 through a combination of reaching out to rural women, improving continuation rates, and offering a wider range of methods (2.06-b). The average rate is disaggregated/differentiated for each of the three provinces; interim as well as end-of-project (1989) targets are given. In addition, the project aims to increase women's knowledge of modern contraceptive methods from the present 63 percent to 90 percent by 1989 (Annex C, p.2). The conceptual statement of the project's monitoring and evaluation activities is unusually clear, i.e., the division of indicators among inputs, outputs, and outcomes (4.12). The project will fund a number of technical evaluation studies, some of which concern FP (annex F, pp.2-3).

Niger (1986)

This is a basic health-system project in a country that is only just beginning to show concern for population problems, where FP is still a sensitive problem, but where the government is willing to offer FP services as part of its MCH facilities. UNFPA and USAID have been playing a stronger role in encouraging the government to focus on its population problem and the project will support, mainly through studies, their efforts to get population and FP policies adopted. There is no attempt to set demographic or outcome targets of any kind.

Nigeria (1985, 1988, & 1991)

The Bank has been following the evolution of Nigeria's population problems and policies since the '70s. From an unconcerned, pronatalist, anti-FP political and cultural posture the country (or at least the Government) has moved to a highly concerned, strongly pro-FP position. When, in the mid-'eighties, the government became interested in Bank assistance in the health sector it was agreed to develop a series of health projects on a state-by-state basis. The Sokoto and Imo projects were the initial steps down this road. As national concern heightened, however, the Government and Bank both realized that a state-by-state approach would be a much slower response to health and population problems than a direct, nation-wide approach. Hence the development of the highly innovative Pop-III, an explicitly National Population Project (April, 1991).

The first two (state) projects incorporated population and FP objectives within wider health projects that emphasized strengthening of PHC. The introduction of FP as one of the basic MCH services did of course require special training of female health workers; the training targets (4.16) had to be extremely modest. The output target is perhaps the most cautious to be found in any Bank project: "The number of family planning acceptors at the health clinics is expected to rise with the training of CH workers and TBAs in family planning" (7.02). [T.12, Annex I, provides a set of "indicators" for monitoring the project; the only FP indicator, however, is the "no. of FP acceptors"—for which annual targets are to be fixed, annually.

Pop-II in Imo (1988) contains a more direct attack on population problems, in line with changed national policy. It is concerned with creating a State Population Council, a Population Studies Center at the State university, and strengthening FP in the state MOH. Among the future work of the Population Studies Center would be the conduct of baseline surveys, operational research and impact studies—but no dates or details are set. No output targets are mentioned. Staff training to be supported by the project gives heavy emphasis to FP, with targeted numbers of courses and trainees (Table, Annex 3-11).

The central government finally adopted a national population policy, and launched a National Population Programme, in early 1989. The policy contains 11 quantitative targets, the achievement of which "will be extremely difficult and very expensive." Annex 2-5 of the SAR presents these targets, "most of which appear . . . unrealistic within the time frame set." Notwithstanding, "a commendable step has been taken by Nigeria in setting such a broad array of concrete targets for itself." The SAR comments that "it would be highly unfortunate if Nigeria's inability to come even close to achieving most of its population targets were to result in a loss of interest on the part of the Government and external agencies and a collapse of morale of program staff and managers." (2.11). The Bank does not suggest revised targets; it does however include its standard population projection tables, using a three-scenario approach of fast, medium, and slow fertility declines.

The SAR is primarily concerned with launching a highly innovative and admittedly risky operational program, and with the development of mechanisms for monitoring its progress. The project consists essentially of establishing a national line-of-credit (the Population Activities Fund) on which any agency, public or private, can draw to finance population activities it can get approved by a controlling authority. This will generate many scattered, dissimilar sub-projects which it would be difficult or impossible to monitor by "most conventional measurement tools such as contraceptive prevalence surveys." But "simple indicators" will be worked out, an existing experimental survey system (the Sentinel Survey System, developed at the Univ. of North Carolina, and described in Annex 3-9) will be spread over a much wider area (4.14). Annex 2-8 notes that "The overall approach to monitoring, evaluation, and, to a lesser extent, research activities still needs to be developed for the National Population Programme." The three-page annex suggests "a possible framework."

Pakistan (1983)

This was an attempt by the Bank and co-financiers to help a revised national program make a fresh start. Some Bank assistance was national but assistance for service delivery was concentrated on 13 districts. In those districts the outcome targets were to raise the CPR from 7 percent in 1983 to 14.5 percent in 1986. That target "implies" a reduction in the CBR of about 3 points in the project districts. Those targets were close to the government's national targets (2.06 and 1.19). Annex 5, p.2. lists seven indicators (expressed as targets) which, if met, would demonstrate "good progress" in the implementation of the Family Welfare Centers project component. Two national CPR surveys (to be funded by USAID) would provide baseline (1983) and progress (1985) data. A fairly full package of evaluation surveys and studies was also provided (4.04-4.05).

Peru (1982)

This project will strengthen PHC in four of the country's health regions. PHC includes FP among its MCH services. "Specific operational targets have been developed against which the effectiveness and impact of project supported activities and services could be measured . . ." (6.01). A one-page annex lists the targets ("after 5 years of service delivery"), expressed in percentage moves up or down, as appropriate. The only FP target used is a percentage increase in FP users; the other 31 targets are all health targets.

Philippines (1974 & 1979)

Pop-I (1974) contains an 18-pp. demographic annex that works out the numbers of acceptors and continuation rates (and thus the increase in the CPR) needed to reach certain demographic targets. The latter were defined by calculating the Philippine population in 2000 (a) assuming no significant change in the present high fertility and (b) assuming a "moderately rapid" (= "reasonable") decline in fertility. The difference of about 25 million was taken as a target number of births to be averted over the next quarter century. The SAR judged that "there is a fair chance that the target numbers for the program . . . may be achieved with reasonably moderate strengthening" of the program ("the program" referred to the government program—some 40 percent of contraceptors were receiving supplies outside the program in 1974, a figure expected to rise to 50 percent by 1991. Thus the government program was scaled to accomplish 60 percent of the target, or 15 mn. averted births, the other 40 percent falling to the private sector). A series of tables in the annex worked out the continuing and new acceptors, and the number of FP outlets needed,

year-by-year, 1974-1985. It also worked out the ASFR and TFR for 1970-75-80=2000; under the target program (and allowing for the substantial private-sector contribution), the TFR would fall from 6.4 to 3.8.

In the 1979 SAR the Bank expressed some reservations about the program's ability to meet its increased targets of a 50 percent CPR by 1987 and a slowing of the population growth rate from 2.5 percent in 1978 to 2.1 percent in '87 and 1.6 percent by 2000. Major increases in the scale, intensity, and efficiency of MOH services would be necessary to reach these goals. Pop-II would increase from 33 percent to 50 percent the proportion of population reached by MOH services. "Contraceptive effectiveness" would be increased from 86 percent in 1978 to 90 percent in 1982 "by emphasizing more effective and lasting methods"

Rwanda (1986 & 1991)

These two projects (both primarily population projects) attempt to help build a population control program in a country with one of the most pressing problems in the world (in 1989 the TFR was 8.5, population was rising at 3.6 percent p.a, and land was becoming exceedingly scarce). By the mid-80's the government had become convinced of the need to slow population and now supported FP. A government target of increasing the CPR (0.9 percent in 1986) to 10 percent by 1991 was judged realistic, although no significant reduction in fertility could be expected until about 2000. Table 7 in Annex I shows acceptors, users, and births averted, by year, 1987-91; these provide specific global targets. Table 8 shows the 1984 and projected/targeted 1991 method mix. Table 23 of the same annex is a one-page display of "Key Project Indicators" (13 "process" indicators and four "outcome" indicators, including the above-mentioned CPR) that also serve as quantitative targets. The project also funds a study, by the national population agency, of acceptors and their reasons for continuing or dropping out, by method.

SAR-II continues to highlight the urgency of the country's demographic situation. Using simple projections, it demonstrates the time required for a moderately-rising CPR to affect overall fertility. Thus under the six-year 1991 project "only the conditions will be set in place for a sustained reduction in fertility and related improvements in health conditions." The report summarizes the government's demographic targets (3.02) which it characterizes as "extremely difficult to achieve with the time frame allotted." But it goes on to say that the targets "should provide an impetus to the program, motivating providers and users." To achieve a reduction in the TFR from 8.5 in 1985 to 6.4 by 2000 would require a CPR of about 30 percent by 2005 (this result was produced by simulations, using the Bongaarts model - Annex I shows the calculations). The implications of that target CPR for the number of FP users by 1995 and 2000; they imply more than a 200 percent increase in the number of users between 1991 and 1995.

Senegal (1982 & 1991)

Pop-I was essentially a PHC project in which an initial attempt is made to help the government extend FP services (recently introduced as part of a USAID-funded project) more widely within the MCH delivery system. Although the government had begun to show some interest in its population problem, the new National Population Commission had not yet developed a population policy. No demographic or acceptor targets are mentioned.

Pop-II (nine years later) reflects a considerable maturing of the population environment. Projections of the levels to which the CPR should rise (from 3.4 percent in '91 to around 14 percent

in '95 and around 22 percent by 2000) are offered in effect as targets. There is no reference to demographic targets. A considerable program of population research and studies will be supported, including two iterations of the 1986 Demographic and Health Survey (to be done in '91 and '96). The DHS studies are expected to provide monitoring data on fertility, contraceptive prevalence, knowledge of and attitudes toward FP, and infant and child mortality. The common device of a comprehensive mid-project evaluation is also provided. An eight-page annex provides summary performance indicators (few of them quantitative) for all project components.

Sierra Leone (1986)

A typical mid-'80s Africa project where a government has just begun to recognize its population problem and the MOH is just beginning to offer FP services, which are entirely urban and led by NGOs. The project's assistance in extending fledgling FP services into MCH facilities (especially into an expanded network of rural PHC clinics) will not do more than "help lay the foundation for the development of a national population and family planning program" (6.91). No mention is made of any demographic or acceptor targets—beyond reporting that the local project preparation task force recommended the adoption of a national population policy with explicit demographic targets (2.29). A major project aim is to strengthen the capabilities of the young National Population Commission.

Sri Lanka (1988)

A project with clear, explicit demographic and global FP targets in a country with a quarter-century-old strong and successful population program. Since 1960 the TFR has fallen from 5.4 to 2.8-3.0 and the CBR has declined from 38 in the '40s to 24. The age of marriage is about 25 for women, 28 for men. A 1987 demographic and health survey (DHS) estimated the CPR at 60-62 percent, with recent increases half accounted for by sterilizations (mostly female). Temporary methods are estimated to rely two-thirds on traditional methods (rhythm, withdrawal, etc.); this accounts for one of the project's targets, namely, to raise the CPR for modern methods from 40 percent to 50 percent. An explicit set of service and IEC targets (3.33) is intended to help the country reduce the TFR by 0.4 by 1993 (if this is attained, the country would likely reach replacement fertility five years earlier than the latest projection showed).

The fact that NRR does not yet equal 1.0 means that increasing numbers of women are continuing to enter their reproductive years (the annual increase is estimated at 100,000 fertile women through 1993). It will therefore be difficult just to maintain the present CPR, especially since the already-high sterilization rate (30 percent) is threatened by recent substantial falls in sterilization acceptors. So there will have to be increasing emphasis on temporary methods. No method-mix targets are given, however. Six annual five-day "population strategy workshops," involving public and private actors, will develop annual service-delivery plans (no mention if those annual plans will include targets, by method . . .)(3.55 and 3.56).

Thailand (1978)

This SAR has perhaps the most explicit set of demographic and FP targets of any Bank report. This was made possible by the clarity of Government population policy, its many years of a strong and successful FP program, and a good data base. A detailed 13-page annex reviews and endorses government acceptor targets, which are spelled out by method and province for each year of the five-year project. The demographic target is to reduce the rate of population growth from 2.5

percent to 2.1 percent by 1981; this represents a decline in the CBR from 34.5 to 29 and in the CDR from 8.3 to 8.0. This target will require the prevention of 2.6 million births over the five-year period. These demographic targets will require the recruitment of some three million new acceptors (600,000 p.a. compared with an average of 490,000 p.a. during the preceding five years), adding 1.6 million additional continuing users by the end of 1981 (3.11 and 4.01).

Togo (1991)

Another typical Africa late-'80s project where a Government has come to acknowledge its population program, has adopted a formal population policy and program, and is introducing FP services into its MCH delivery system. The 24-page Annex reproduces the new National Population Policy under which the government will launch activities "to sensitize the general population, increase availability and use of family planning services and improve PHC services to combat high mortality. These measures are critical for setting and achieving, in the future, specific targets to reduce fertility" (para. 71). Later on the SAR reports a government goal to raise the CPR from the present 3 percent [5 percent?] to 15 percent by 1995 (para. 83). Annex 1 (b) lists 22 quantitative goals (to be reached by 1995) of the MCH component of an expanded PHC program; six of the 22 goals are FP targets, i.e.,

- i. promote a birth-spacing interval of at least two years among 70 percent of new mothers, by 1995;
- ii. reduce the number of births among women under age 20 and over age 35 from 30 percent in 1990 to 20 percent by 1995;
- iii. reduce the percentage of high order pregnancies (more than six) from 60 percent to 50 percent, by 1995;
- iv. reduce maternal mortality from unwanted pregnancies from 16 percent to 8 percent, by 1995;
- v. increase the availability of services to women and married couples from 10 percent to 60 percent, by 1995;
- vi. increase the use of modern contraceptive methods from 5 percent to 15 percent, by 1995;

Trinidad and Tobago (1971)

Like several of the early SARs, this one includes a long (11 pp.) Annex that links estimates of acceptors to births averted and other demographic outcomes. The Annex's Outline shows the calculations made:

- a. Number of New Acceptors
 - (i) Acceptors in the Postpartum Program
 - (ii) Acceptors in Project Health Centers
 - (iii) Acceptors in Existing Facilities
 - (iv) Total New Acceptors
- b. Births Averted
 - (i) Contraceptive Methods
 - (ii) Continuation Rates
 - (iii) Women-years of Protection
 - (iv) Births averted
- c. Effect on the Birth Rate

- d. Effect on the Population Size and Growth Rate
- e. Effect on the Labor Force

While not labeled "targets," the estimates are considered "indicative of possible achievements," which is what a well-set target is!

The results of the calculations are sprinkled through the SAR and are summarized in para. viii of the "Summary and Conclusions:"

The project is estimated to increase new acceptors from 10,000 in 1970 to 22,000 per year from 1980 onwards. As a result of this, the birth rate is estimated to decrease from 27/1000 in 1969 to 15.5/1000 in 1980. The rate of population growth will be reduced from a projected level of 1.5 percent per year in 1980 without the project to a level of 0.9 percent with the project. By the year 2000 the population would be 6.2 percent smaller than otherwise and the labor force 5.2 percent smaller.

Tunisia (1971 & 1991)

Tunisia has much the strongest and most successful FP program in the Arab world. The first Bank assistance (1971) came after the national program had been in existence nearly five years. The SAR contains a 4-pp. demographic annex very similar to that just reviewed for Trinidad and Tobago. It established the following acceptor and demographic targets:

On the demographic side, the project should add 55,000 new acceptors to the 1970 level of 21,300 resulting in a total of 76,300 new acceptors a year by 1976. This would result in about 21,300 averted births per year in 1979, of which 15,300 are attributable to the proposed project alone The total number of averted births will reach 551,000 by the year 2000. The effect of this family planning program would be to make the population about 4.6 percent smaller than otherwise in the year 2000, while that of the project alone would be to make it 3.2 percent smaller than otherwise (6.02 and 6.04).

Based on its analysis, the SAR concludes that "the targets of program achievement by the mid-1970's can be set much higher than recent performance." (para. v, Summary and Conclusions). The SAR's projections are treated as targets, with a clear distinction between project and program targets.

Pop-II followed Pop-I by 20 years. The project's objectives are expressed in explicit target terms:

The objectives of the Population and Family Health project are to increase the Contraceptive Prevalence Rate (CPR) from 40 percent to 45 percent and to improve the health status of the population, especially of women and children in rural areas. These objectives would be pursued through two strategies:

- a. strengthening of family planning and maternal and child health services (FP/MCH) within an improved Basic Health Care (BBC) package;
- b. targeting BBC for population groups with the highest fertility and mortality levels and inadequate access to BBC services (2.1).

While other projects have also used group targeting in project design, Pop-II is more explicit in saying so. The SAR states that "The target number of women to be served will be set by the [ONFP] . . . to meet the national target to increase the CPR (for modern methods) from 40 to 45 percent by 1996. The specific targets per governorate will be agreed upon by the Regional Health Directors and approved by the Project Steering Committee in the context of the Yearly Resource Allocation and Implementation Plans" (2.5). The report makes no reference to targets by contraceptive method.

Pop-II makes use of one other kind of target not seen in any of the other 75 SARs: the number of consultations for FP per Married Woman of Fertile Age. The project would raise this figure from a current level of 0.23 to 0.6, which "is adequate for IUD and oral contraceptive users." As there are large regional variations in this norm, "the project's focus on underserved areas would significantly narrow these disparities." With an expected doubling of pre-natal and post-natal consultations, it will help identify high-risk pregnancies and reduce maternal and neo-natal mortality (5.1).

Yemen Arab Republic (1990)

This is a second health-sector development project, the first (1983) having experienced great difficulties in implementation. Although the SAR reports on the country's demographic and FP situation (very bad demographic numbers, very little FP), the project has no population or FP objectives or targets, explicit or implicit. "Direct policies with quantitative demographic targets for population growth rates are deemed unacceptable." (1.07). The SAR notes that "The Government supports maternal and child health programs, family planning training and services, and public information." However, there is no reference to FP in the description of the health services, or training activities, whose expansion the project will support.

Zimbabwe (1986 & 1991)

These two projects, and their respective SARs, are very similar. They are set in a country with a relatively high per capita income (for Africa), where fertility has already started to decline (the only such country in Southern Africa), and whose government is strongly supportive of slowing population growth even though it does not yet have a document formally stating its policies. Both projects choose a set of under-served districts for extension of MCH/FP services. Both have population and FP as primary project objectives. Both have explicit demographic and FP targets.

Pop-I accepts the government's FP target for the project area: "The target contraceptive prevalence rate for the eight districts by the end of the project is about 50 percent—a target considerably above the present estimated rate of about 20 percent in these districts, but not unreasonably high in light of Zimbabwe's achievements since Independence and the inputs through the project If achieved, there would be approximately 136,000 FP acceptors in the eight districts in 1991, compared with approximately 55,000 if the present contraceptive prevalence rate were to remain constant. Similar calculations can be made for the other MCH interventions" (6.04). The project will fund a baseline survey, plus a second survey in its third year, "and the targets would again be reviewed and revised if appropriate" (3.04).

In addition to monitoring progress towards meeting its outcome targets, Pop-I makes considerable use of "process," or input, indicators: the output of trained nurses and village health

workers year by year; the use of mass media by the IEC unit; the production of manuals and booklets; and the holding of planned training workshops.

Pop-II covers 16 under-served districts, and involves six donors. Its design is explicitly linked to an analysis of lessons learned in implementing Pop-I (which had gone well). Pop-II carries quantitative targeting somewhat beyond the level of Pop-I. It endorses, for example, the series of nationwide "process indicators" developed by the government to monitor H/FP/N status and service delivery. These are:

- (a) lowering the infant mortality rate from 53 in 1990 to 43 in 1996;
- (b) increasing the percentage of children receiving the complete vaccination course, at the correct age/interval, from 70 percent in 1990 to 85 percent in 1996;
- (c) reducing the share of children aged 6-36 months who are malnourished (less than 80 percent of the reference weight for age) from 12 percent in 1990 to 5 percent in 1996;
- (d) increasing the percentage of pregnant women receiving antenatal care from 90 percent in 1990 to 95 percent by 1996 and increasing the percentage who deliver in a health facility from 70 percent to 85 percent;
- (e) lowering the total fertility rate from 5.5 in 1986 to 4.5 in 1996, and raising the percentage of married women of reproductive age using modern methods of contraception from 36 percent to 48 percent;
- (f) increasing the percentage of married women using permanent and semi-permanent methods from 4 percent in 1988 to 12 percent in 1996;
- (g) training 3,500 nurses and 120 doctors in family planning;
- (h) increasing the proportion of nurses in district and rural service trained in midwifery to 60 percent by 1996;
- (i) increasing the level of cost recovery by MOH from 3 percent of recurrent budget expenditures in 1990 to at least 10 percent by July 1, 1995. Government's annual target increases will be monitored during implementation.

The SAR goes on to note that, "For each of the project objectives and related components/subcomponents a more detailed set of progress indicators has been elaborated. In many cases, the provincial and district health planning teams have developed their own targets" (3.2).

Annex 2 contains a fuller statement of the government's national FP strategy. This includes two "goals" and seven subsidiary "objectives." Included is a table showing 1995 national targets for numbers of FP users, by method and by source of service. No comparable table is given for the 16 project districts, however.

Annex III. Treatment of Targets and Indicators in PCRs and PPARs

The Bank requires that a Project Completion Report (PCR) be written by the implementing division within nine months of final disbursement. These reports are usually written by one or more of the individuals who have been involved in supervising the project, although the authors may not have been around when the project was under design. PCRs are then reviewed by the (independent) Operations Evaluation Department (OED) which, upon approving the document, sends it to the Bank's Board. OED must then also decide whether or not it will conduct a separate, independent audit (i.e., a comprehensive review and evaluation, not a financial audit). Audits are

done on only a sample of all PCRs—in the past, about one-third of all PCRs have been selected for audit. Project Performance Audit Reviews (PPARs or PARs) are conducted by OED staff members or outside consultants; no one from the division responsible for the project participates in the OED review. The intent is to conduct an independent review, free of the biases and temptations to "cover up" or "excuse" which might rob a PCR of objectivity.

PARs use the PCRs as their starting-point, then construct whatever kind of review OED thinks appropriate. The writing of both PCRs and PARs involve field visits; both are circulated in draft to the borrower (and often to co-financiers) for comment. The PARs normally pay somewhat less attention to project implementation matters than PCRs do, and more to project design and impacts. PARs result in a Project Audit Review Memorandum which is circulated to the Board and placed in a bound file in the OED library; PARs are not published, however. Memorandums (typically 10-20 pp. in length) invariably include the full text of the PCR (of comparable length, plus frequent annexes). Unlike Appraisal Reports, the authors of PCRs and PARs are not identified. For credibility, one relies on the quality of the review and the majesty of impersonality.

Of the 75 population projects approved by the Board to the end of 1991, 28 had been completed long enough to have had their PCRs written. Eighteen PCRs had been selected for PARs, and the Reviews completed, by Feb. 1, 1992. The 28 completed PCR cover population/health projects in 16 countries, reflecting coverage of two or more projects in nine of the countries. There is a clustering of PARs in '81-82 (five) and, especially, in '85-86 (seven).

The summaries below extract comments on targets and indicators from PCRs and PARs for 14 of the 16 countries for which PARs exist to date (reports on the two omitted countries covered failed projects that had nothing to contribute—not even the study of their failures). The documents provide a good historical record of project implementation (usually of implementation difficulties) and of initial results, or impacts, with the PARs offering judgments on project design and explanations of delays and impacts, or the lack thereof. They do not have much to tell us, however, about targets and indicators that we have not already learned from our review of SARs and other sources.

Bangladesh (PAR # 6303, June 1986)

The PAR on the first of the Bank's four Bangladesh projects arrived at a general conclusion that the project had been too complex, was insufficiently focused, was too big, and required an "enormous" supervision effort (which it praised). It noted that ". . . although it is impossible to estimate causal links between project inputs and program results, it is important to assess the results of the overall population program . . ." Both the PCR and the PAR agreed that "the major contribution of the project was in developing a population program in Bangladesh." The PAR noted that OED had used this project as a case study in a special review it had made of project supervision. That study noted the difficulty the project had experienced in establishing a monitoring system—five years after the project's start, "a system for monitoring the impact of the project had still not been satisfactorily established."

The PCR for the second project reviewed the successive reductions in the Government's demographic and contraceptive targets. Tables are included showing trends in the CPR, by method, 1975-85. It noted that the Government and donors were now in agreement on the key demographic and CPR targets, including the need for stronger emphasis on temporary methods of birth-prevention (which recruit younger couples before they opt for sterilization) if the demographic goals were to be

met. The PCR was reporting trends in key (program, not project) indicators and on agreed broad program targets.

Dominican Republic (PAR # 5596, April 1985)

The Review judged the project successful in terms of "fostering a sound population policy through the strengthening of primary health services." The project's precise demographic impact "cannot be measured," however, despite original intentions to do so. National fertility has been falling, and "the project's likely impact probably comes reasonably close" to the original intent that it should account for one-quarter of the national decline. There is no discussion of the methodological difficulties which prevented a closer linkage of the project's contribution to national trends.

Egypt (PAR # 4144, October 1982)

The Bank/borrower relationship was a difficult one from the start. The project set no quantitative objectives: its main objective was "The development of a mutually satisfactory long-term relationship" (the SAR had actually recommended against the project "because it would have too little demographic impact!" The PCR and Review abundantly confirm that predicted outcome).

Despite the difficult relationship, a second project was signed before the first had finished. The second project did set acceptor and CPR targets. Subsequent surveys reported conflicting trends, making it impossible to know what was happening. The second project fared little better than the first, becoming a "problem project" almost from its start, and being canceled in year 7 because of poor progress. The PCR concluded that the Government was not much interested in population and the Ministry of Health not much interested in family planning. So, the Review noted, the Bank must pursue dialogue through other channels.

India (PAR # 3748, January 1982; PAR # 8896, June 1990)

The first project (1973-80) showed little or no concern for targets but the Review noted that acceptor results in the two project states were somewhat better than in non-project areas. Perhaps more important, the project provided key ideas for the Government's Model Plan, and "invented," and established, the two Population Centers (PCs), in Bangalore and Lucknow, to conduct operational research. The project left it up to the new PCs to develop an appropriate set of indicators for this and other projects.

The PCR and PAR for the second project used standard FP indicators to demonstrate that, contrary to expectations, performance in the project districts was no better than in non-project districts. Both reviews concluded that the Bank's typical area-by-area approach in India was not likely to show above-average results unless the Government could change some national program policies. The PAR faulted the Bank for not undertaking sector work that might have led to such policy changes.

As for Bank attention to indicators, the PAR noted the attention given them in the SAR but also noted that the surveys on which they heavily depended (to be carried out by the two PCs mentioned above) were not very satisfactory. "The vital rates reported were questionable, and in a number of instances differences between baseline and mid-term estimates seemed quite unlikely."

Indonesia (PAR # 5675, May 1985; PAR # 6276, June 1986)

These two Reviews again make the point that project accomplishments cannot be objectively judged unless project indicators are specified in advance. The first Review makes the point summarily—and promises a more elaborate discussion in Review no. 2, which is in fact presented. Review no. 1 cannot accept the PCR's claim that "the project has undoubtedly succeeded in meeting its ambitious objectives: contraceptive acceptance targets have been exceeded, family planning is no longer considered an especially sensitive issue on religious or political grounds and the total fertility rate has been considerably reduced."

Neither in the PCR nor in the appraisal report is there a specific framework or set of indicators on the basis of which causal relationships can be established between the project's stated objectives on the one hand and project activities on the other. As a result, there is no hard "causal" evidence as to the extent that "micro" project activities contributed to "macro" project objectives. This point has been elaborated upon in earlier audits and will be pursued in the audit of the Second Population Project.

The Review of the second project attempts to spell out what OED means when it refers (as it repeatedly has done) to the need for "a hierarchy of objectives and targets to make evaluation of population programs and projects more meaningful" Indeed, the PAR suggests "that PHN develop such an evaluation methodology soon." The two-page effort to lead PHN down this path is unsatisfying, however. The advice given is more hortatory than specific and practical. This disappointment does not invalidate the Review's basic advice that PHN needs to work up a better set of project indicators (an effort to which the present paper may contribute).

The PCR for the Third Population Project offers the following comments on the "project indicators" problem:

As stated . . . direct relationships between project inputs and fertility decline cannot be demonstrated. The appraisal stated that Population III would assist the GOI to achieve its interim goal of reducing the 1979 TFR from 4.8 to 3.5 in 1984. There is an ongoing debate whether more monitorable and direct indicators should be used to assess project impact rather than using outcome indicators such as reductions in fertility and mortality The project design would have benefitted from a greater emphasis on process indicators and a systematic listing of when certain indicators should be monitored during project supervision. Additional impact indicators to be included in the project design would be contraceptive prevalence rates and number of current users during each year of project implementation. A small survey towards the end of the project, focusing on knowledge, attitudes and practices, and designed to produce results comparable to surveys previously undertaken, would have provided an evaluation of some aspect of immediate project impact.

Jamaica (PAR # 2580, June 1979; PAR # 5589, April 1985)

1. PAR 2580 was OED's first review of a population project (the Bank's first in this sector). The Review said nothing about targets and indicators. The second Review had much to say, however. It began by expressing the view that project performance would have been better if the project had set "realistic objectives" and had specified "performance indicators to assess the

effectiveness of project strategy." Instead, OED noted, the SAR used program objectives to justify future project performance. But these are

unrealistic yardsticks . . . because program goals are achieved by multiple Government and non-Government interventions and efforts by other donor agencies, and are influenced by socio-economic development in general. The Second Population Project is only one component of the Government's . . . efforts. Even if program objectives are being achieved, it would be difficult to establish the causal link between results and program achievements.

During project design, quantifiable goals to measure the outcome of the various project strategies . . . were not stated. In the absence of such quantifiable targets and related indicators, real project achievement and impact could not be assessed. To what extent project strategies have been effectively implemented becomes a matter of subjective judgment, always open to debate. More objective measurement of strategy performance would be preferable.

The Review then goes on to make these suggestions:

Even if program objectives need to be specified as part of project justification, the audit suggests that during project preparation a clear hierarchy of project objectives be established: program objectives, strategy performance indicators and physical implementation and utilization targets. Indicators at the delivery and the institutional levels will both assist in linking project impact to the overall program goals as well as provide guidance as to what type of data need to be collected, processed and analyzed for purposes of management information and project impact evaluation.

With reference to institutional development, the objectives should be stated keeping in mind the time frame necessary for bringing about organizational changes

A few comments on OED's suggestions may be in order. Establishing "program objectives" normally presents no problems—higher CPR (using modern methods) and a falling TFR are standard, valid goals. But a question arises whether or not these same "output" goals should be applied at the project level, using the same sample-survey techniques used to measure national (= program) performance). The answer would seem to depend on how sharply a project's catchment area can be separated out from the rest of the program. I have not identified any Bank project where special surveys have been conducted to get this kind of outcome data at the project level (a few projects, e.g. in India, have compared project with national performance using service-statistic data, which measure earlier-stage performance, not demographic results).

We come next to [project] "strategy performance indicators" and "physical implementation and utilization targets." I am not sure what these two sets mean. I would prefer to separate out "physical implementation" targets, which are almost always adequately covered by supervision reports on the project implementation stage. Facility utilization seems a particularly important, and widely neglected, indicator. Establishing utilization standards is the key problem—some combination of "capacity," trends, and comparison with comparable facilities is usually feasible. "Utilization" of clinics is of course a neutral measure of "activity," revealing nothing about results, e.g. acceptance or termination of contraception (intermediate outputs). Service statistics (which are essentially the same for any clinic, whether it be offering pure health, pure

population, or mixed services) can provide all the needed activity indicators and at least some useful contraceptive output data. What is needed, in my view, is agreement on the main classes of performance indicators and a listing of the data-requirements for each class of indicator. I have tried to provide such a framework in the Conclusions and Recommendations section of the main text.

Kenya (PAR # 3536, June 1981)

This early PAR was done in the gloomy days when all government/donor population efforts were being met with perverse results—rising fertility and acceptor rates far short of targets. The PAR noted these disappointing results and that the project was not designed to generate data that might explain them. The PCR did note, however, that program acceptor targets were based on Asian norms, which proved "unrealistically high."

We know today that by the end of the '80s the demand for FP services has shot up higher and faster than anyone had expected, for reasons not yet well understood. One wishes one could say that the success of the late-'80s has reflected changes in the national program made in response to analyses (by the Bank and others) of program weaknesses earlier in the decade. Any such claim would require a detailed analysis of multiple reports and subsequent adjustments in program content.

Korea (PAR # 8895, June 1990)

This PAR is useful because of its graphical presentation of 15-25 year trends in the overall CPR (urban, rural, total—a seven-fold rise from about 10 percent in 1964 to over 70 percent in 1985), the CPR by method (it shows the over five-fold explosion in sterilizations after 1975, with continuing high use of birth-spacing methods by younger couples), and the steady decline in the TFR (from 6.0 in 1960 to 2.1 by 1985).

The PAR is skeptical that the project (1980-87) made any major contribution to the strong national trends. Socio-economic developments generated by Korea's remarkable economic growth doubtless played a much stronger role, leading to widespread reliance on non-government sources of FP advice and supplies. Utilization of project facilities, while showing some increase during the project period, was judged lower than expected and well below capacities. If the project had come a decade earlier, before people had the incomes to visit private providers, it might reasonably have claimed a larger contribution to the national trends.

Malaysia (PAR # 5641, May, 1985)

This PAR appears to have been written by the same author who wrote the one for Jamaica (see above). He makes precisely the same criticism, namely, that project impact cannot be assessed because the SAR established indicators only for the national program, not for the project itself. The author then pleads for exactly the same set of project yardsticks also found lacking in the Jamaica project:

A hierarchy of project objectives should be established at the time of appraisal of population projects indicating (i) national family planning program goals; (ii) project strategy performance indicators; and (iii) physical implementation and utilization targets.

There should be a distinction between long-term, medium-term and short-term goals of the project. Similarly, for institutional strengthening (of major importance in this

project), outcomes to be expected once the institutions are developed and the time frame needed for institutional development should be specified.

These are right instincts but not as operationally transparent as one might hope. Does a short-term project goal mean getting the buildings built, equipped, and staffed? A medium-term goal to achieve rising utilization and acceptor rates? A long-term goal (over 5-10 years after project completion?) to maintain the facilities and reduce the fertility of those who use them? Does this mean sample surveys, two or three years after project completion, to establish CPRs and TFRs for project-only users? And is progress towards institutional goals susceptible of measurement through regularly-collected periodic indicators—or perhaps better through occasional *ad hoc* reviews (including reviews conducted in the course of preparing repeater projects)?

OED reviews and judgments deserve a high place on the agenda of professional discussion about how to assess the worthwhileness of Bank population projects—but no more. They do not deserve any extra-ordinary deference simply because they represent the Bank's "judgment of last resort" in a sequence of formal procedures. Whether or not PARs add significantly to the data and insights of PCR's is a separate, important issue. My own view, based solely on my reading of the 18 population PARs done to date, is that the cost of conducting PARs far exceeds their benefits. There is simply too much duplication between the two sets of reports to justify continuing to do both. The Bank would learn far more about how to improve the design of population projects if it stopped doing PARs and devoted those resources to *ad hoc* operational research and to external reviews of national population programs. OED is said to be moving cautiously in this direction, at least in the population sector.

Philippines (PAR # 5544, March 1985; PAR # 9380, February 1991)

Government interest in FP has never been strong, and steadily weakened during the course of two Bank projects. This dominant political fact presented so many problems for useful dialogue and for project implementation that one wonders why OED considered either of the projects worth auditing. For our purposes the only relevant part of either Review is the lament, in the first Review, that the Bank decided not to undertake a survey on the use of project facilities which had been part of the original project. That, said the Review, was a mistake. With utilization reviews as rare as they are, this might have provided a useful model.

Thailand (PAR # 7335, June 1988)

This Review concludes with an observation close to those noted in the Jamaica and Malaysia Reviews:

Evaluation of population projects is inherently difficult but would be facilitated if objectives were clearly formulated and linked with program objectives. A hierarchy of program, project, impact and process objectives is needed, together with clear targets for project components.

Trinidad and Tobago (PAR # 3687, December 1981)

This early Bank project, begun just before the oil boom of the early '70s, encountered many implementation difficulties, and the national program badly missed the acceptor targets set for it in the project.

. . . It is difficult to escape the conclusion that the national program should have done very much better. Government commitment has weakened steadily. Management has been weak at the central and county levels. Although . . . the more efficient private Family Planning Association has had to curtail its operations, in 1979 the FPA recruited at least half as many new acceptors with only two clinics as all of the 86 government centers combined.

The PAR does not conclude that the project should not have gone forward. It says only that it should have been a somewhat different kind of project—and perhaps should not have been approved until the government had demonstrated stronger commitment by remedying the institutional weaknesses which later hobbled project implementation. The PAR did not address, however, the following critical point, drawn not from the PAR but from the PCR:

It must be remembered that in the years following 1968, the Bank, in the light of the position taken by Mr. McNamara at the 1968 Annual Meeting and on subsequent occasions, regarded it as its duty to draw the attention of member countries to the threat to development posed by rapid growth of population and to encourage them to take active steps to check this growth. It achieved positive responses, such as that of Trinidad and Tobago, to Mr. McNamara's initiative, and felt that it should go out of its way to foster the interest thus displayed. The Bank's willingness to make the loan at an early stage of the project cycle reflected this attitude. This, like many first projects, was intended to initiate and sustain a sectoral dialogue between the Government and the Bank.

In short, the Bank's speculative investment in "dialogue" often has to be accepted, in many of the Bank's early projects, as the basic project justification. No amount of disappointment thrown up by weak performance indicators could justify an *ex post* judgment that the project should not have gone forward (as noted, the PAR does not contain any such judgment). One can legitimately wonder, however, how well Bank staff make the judgment as to whether or not a proposed project, in a given situation, will produce a dialogue worth having. In the present case, there seems general agreement that no one could have foreseen the rapid worsening of the project environment that set in shortly after the project was launched. So there was very little useful dialogue, and lots of bad numbers.

Tunisia (PAR 5135, June 1984)

The PAR on the first of two projects notes that project design included good specification of project indicators. At time of writing, however, these had proved of little use because the management information system (MIS) had developed too slowly to generate indicator data. The PCR for the second project ('81-89) noted the system's continuing inability to develop a meaningful MIS system. The project did include explicit FP targets, based on the major project objective of integrating the vertical FP delivery system into the Ministry of Health's Basic Health System (BHS). Since the integration failed to take place, it was clear that the targets (based on the unrealized BHS system) were not met.

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS1033	Latin America and the Caribbean Region (and Northern America) Population Projections, 1992-93 Edition	My T. Vu Eduard Bos Ann Levin	November 1992	O. Nadora 31091
WPS1034	Revising Financial Sector Policy in Transitional Socialist Economies: Will Universal Banks Prove Viable?	David H. Scott	November 1992	K. Waelti 37664
WPS1035	How Import Protection Affects the Philippines' Motor Vehicle Industry	Wendy E. Takacs	November 1992	D. Ballantyne 37947
WPS1036	Output Decline in Hungary and Poland in 1990-91: Structural Change and Aggregate Shocks	Simon Commander Fabrizio Coricelli	November 1992	O. del Cid 35195
WPS1037	Vocational Secondary Schooling, Occupational Choice, and Earnings in Brazil	Ana-Maria Arriagada Adrian Ziderman	November 1992	C. Cristobal 33640
WPS1038	Determinants of Expatriate Workers' Remittances in North Africa and Europe	Ibrahim A. Elbadawi Robert de Rezende Rocha	November 1992	A. Marañon 31450
WPS1039	Education, Externalities, Fertility, and Economic Growth	Martin Weale	November 1992	PHREE 33680
WPS1040	Lessons of Trade Liberalization in Latin America for Economies in Transition	Jaime de Melo Sumana Dhar	November 1992	D. Ballantyne 37947
WPS1041	Family Planning Success Stories in Bangladesh and India	Moni Nag	November 1992	O. Nadora 31091
WPS1042	Family Planning Success in Two Cities in Zaire	Jane T. Bertrand Judith E. Brown	November 1992	O. Nadora 31091
WPS1043	Deriving Developing Country Repayment Capacity from the Market Prices of Sovereign Debt	Stijn Claessens George Pennacchi	November 1992	R. Vo 33722
WPS1044	Hospital Cost Functions for Developing Countries	Adam Wagstaff Howard Barnum	November 1992	O. Nadora 31091
WPS1045	Social Gains from Female Education: A Cross-National Study	Kalanidhi Subbarao Laura Raney	November 1992	M. Abundo 36820
WPS1046	World Bank Project-Financed Research on Population, Health, and Nutrition	J. Price Gittinger Carol Bradford	November 1992	O. Nadora 31091

Policy Research Working Paper Series

Title	Author	Date	Contact for paper
WPS1047 Côte d'Ivoire: Private Sector Dynamics and Constraints	Enrique Rueda-Sabater Andrew Stone	November 1992	P. Infante 37642
WPS1048 Targets and Indicators in World Bank Population Projects	George Baldwin	November 1992	O. Nadora 31091