# DHS WORKING PAPERS

DHS Demographic and Health Surveys

The Mass Media and Family Planning in Kenya

Charles F. Westoff
and
Germán Rodríguez

MACRO INTERNATIONAL INC.

DHS Working Papers
Number 4

# The Mass Media and Family Planning in Kenya

Charles F. Westoff
and
Germán Rodríguez

Macro International Inc.
Columbia, Maryland, USA

August 1993

Charles F. Westoff is Professor of Sociology and Demographic Studies, and Faculty Associate of the Office of Population Research, Princeton University. Germán Rodríguez is Senior Research Demographer of the Office of Population Research, Princeton University.

The Demographic and Health Surveys (DHS) Program is designed to collect data on fertility, family planning, and maternal and child health. DHS is funded by the United States Agency for International Development. DHS Working Papers are intended to disseminate early findings of in-depth analysis of DHS data. Comments about this working paper or requests for information about the DHS program should be sent to: DHS, Macro International Inc. 8850 Stanford Boulevard, Suite 4000, Columbia, MD 21045, USA (Telephone 410-290-2800; Fax: 410-290-2999).

## Introduction

There is, understandably, a strong interest within population policy and family planning program circles in the potential impact on contraceptive behavior and reproductive preferences of mass media messages that try to inform and motivate people on the methods and advantages of regulating fertility. These messages can (and have) taken many forms, ranging from soap operas on radio and television designed to persuade women of the personal and social advantages of smaller families to spot advertisements about methods and clinics. The appeal of such approaches to agencies committed to promoting family planning lies in the wide coverage and in the cost effectiveness that can be realized, especially in the developing countries.

What is the evidence that such media efforts actually do influence reproductive attitudes and behavior? There is a library of literature on mass communications accumulated over the past half-century that has reached disparate conclusions on the general question of whether the mass media can influence behavior. Such questions have been posed about a wide range of behaviors, including the effects of television on violence, its influence on children's education, consumer choice of products, voting behavior, and so on. Needless to say, the evidence is mixed on the effectiveness of such communications,[1] but the belief that there is some effect has persisted and is reflected in the emergence of a whole industry devoted to such activities.

The application of mass communications to influence fertility is a natural extension of the basic idea that the media can both inform and motivate people about even such complex subjects as their reproductive means and goals. Communications efforts have become increasingly widespread in the third world as part of international technical assistance and government programs designed to reduce fertility. A very useful 1986 summary of this coverage by radio alone listed nearly 250 citations[2] describing these efforts and some limited attempts at evaluation. For example, in Iran, a media campaign to increase contraceptive awareness and practice showed large increases in the number of both pill and condom users.[3] In Egypt, a large-scale "before and after" experimental design was implemented to evaluate a mass communications effort to promote family planning[4] that also showed positive results, controlling for many of the variables that might jointly influence both awareness and use. In the early 1980s, several studies in Jamaica[5] indicated that there was widespread awareness of family planning messages resulting from soap operas, songs, and other messages, but there was still a significant fraction of women with an unmet need for family planning. The main conclusion there was that awareness is not enough and that other considerations, such as fear of side effects, must be confronted. Recent evaluations of mass media campaigns in three different areas in Nigeria "suggest very strongly that mass media interventions can play a major role in promoting family planning use in certain situations."[6] Large increases in the number of family planning clients at clinics followed different communication campaigns. Similar results have been reported in Indonesia and in Trinidad and Tobago.[7]

In Latin America, there are two studies that illustrate some of the methodological difficulties involved in interpreting the association between awareness of media messages on family planning and the practice of contraception, an inherent problem that complicates the analysis of DHS data on this subject. One small village study in Mexico in 1978[8] examined the relationship between general exposure to radio and print media as well as family planning messages and the practice of family planning and found the expected results. However, they emphasize the difficulty of inferring causality: "... yet whatever evidence we found does not allow us to exclude the possibility that those individuals who already use family planning, or know about it, are those likely ... to be constant media consumers."[9]

The most relevant earlier study, which corresponds closely with the illustrative analysis presented below and which also underscores these problems of causal inference, was based on an analysis of survey data from Guatemala, El Salvador and Panama.[10] The analysis focused on data on exposure to media messages and family planning derived from the Contraceptive Prevalence Surveys from 1979 to 1982 in these three countries. At the outset, the authors note appropriately that only an experimental design with measurements before and after exposure to the media messages will permit relatively unambiguous inferences about the causal process.[11] Their analysis of earlier survey data shares the same limitations as the DHS data: although the media efforts preceded the survey in time, the evidence that women who report having been exposed to such messages also practice contraception in higher proportions or prefer fewer children than those women who do not report such media exposure, is subject to interpretations other than the basic hypothesis that such messages promote the use of family planning. Even when the association persists after life cycle and socioeconomic controls are imposed, there remains the nagging uncertainty that women who practice contraception for reasons unconnected with media messages may simply be more likely to hear such messages and report having heard them when asked. We shall return to this problem.

## DHS Data and Kenya

Two-thirds of the DHS-I surveys included one or more questions on exposure to family planning messages in the media, i.e., whether the respondent had heard, read, or seen messages on family planning on the radio, in newspapers, magazines or on posters, or on television.[12] The most common question asked about the radio.

For several reasons, we have selected Kenya in order to illustrate the procedures for analyzing such data. Although it clearly has begun the transition toward lower fertility, the level of fertility is still very high (a Total Fertility Rate of 6.7 in the late 1980s) and is therefore of great program and policy interest. Moreover, there have been major efforts to use the media to promote family planning prior to the time of the Kenya DHS survey in 1988-89. The Family Planning Association of Kenya has played an important role in such activities, producing many booklets, posters, films and videos as well as radio programs about family planning. A recent report [13] noted that there were three regular radio programs on family planning and population issues being broadcast. Although radio is the most important communication vehicle,[14] there was also a popular television soap opera that emphasized the problems of early pregnancy.[15] In addition to radio and television, USAID/Kenya has supported the production and dissemination of family planning calendars, posters, and leaflets. Most of the messages are aimed at creating general awareness of small family norms.

One of the radio programs — a soap opera aired twice weekly and listened to by 39 percent of the country (up to 60 percent by one account that described it as having "the highest rating of any radio program in the history of broadcasting in Kenya"[16]) — was subsequently evaluated with interviews on samples of rural and lower class urban communities in an effort to gauge audience reactions.[17] The messages communicated on the program emphasized monogamy, delaying marriage, limiting family size, equal treatment of male and female children, and the problems engendered by rapid population growth. Although the study design had serious limitations, the conclusion was reached that nearly half of the respondents claimed to have acquired their initial knowledge of family planning through the media.[18] The authors observe, however, as in other studies cited above, that such people may have been more favorably inclined toward family planning initially. The anecdotal evidence seems highly convincing, however. A survey of rural clinics is reported to have revealed the universal testimony of clinic personnel that "everyone says they have come because of the radio."[19] A sample survey in 1988-89 of 4800 persons in the provinces served by the Family Planning Association of Kenya determined that radio is the primary source of both women's and men's first information on family planning.[20]

## Media Exposure

Many more Kenyans listen to the radio than read the newspaper or watch television. A total of 68.4 percent of married Kenyan women report that they listen to the radio at least once a week. Radios are owned by 60.7 percent, while only 4.5 percent report owning a television set.

The focus of this report is to illustrate the analysis of the effects of mass media on reproductive behavior, in particular the impact of media messages on family planning. The Kenya DHS survey included a question inquiring about whether the woman had heard or read about family planning in the last six months. The question was repeated for the radio, television, newspaper or magazines, posters, and friends or relatives. The proportions responding "yes" are shown in Table 1.

The importance of the radio is clear from these data. The category "friends and relatives" is included along with the mass media because it was part of the question and also probably serves frequently as a reinforcement and as a link between media messages and women who may not have been exposed to these media directly.[21] Since "friends and relatives" are not in the category of the media, no further reference is made to their role in the communication link.

There is a strong cumulative association between having heard such messages on different media. If women report having seen television messages, they are very likely to report having seen messages in the print media or on posters, and to have heard family planning messages on the radio. If they have not seen television messages but have seen such content in print, they are also very likely to have heard them on the radio. In view of this result, we created a cumulative scale of media exposure scoring one point for having heard a message on the radio, one for either papers or posters and one for television. This Guttman scale[22] has a coefficient of reproducibility of 0.955, which means that only 4.5 percent of women do not conform to the cumulative logic where a score of 1 means radio, 2 means radio and print and 3 means radio, print and television. The distribution for this media scale is shown in Table 2.

Table 1. Percent of Married Women who Heard or Read about Family Planning in the Last Six Months, Kenya 1989

| Source | Percent |
|---|---|
| Radio | 65.9 |
| Television | 7.8 |
| Newspapers/Magazine | 20.9 |
| Poster | 23.6 |
| Friends/Relatives | 67.1 |

## Characteristics of Women Exposed to Media Messages

Table 2. Cumulative Scale of Media Exposure, Kenya 1989

| Score | Media Exposure | Percent |
|---|---|---|
| 0 | No messages reported | 30.3 |
| 1 | Radio only | 41.7 |
| 2 | Radio and print media or poster | 21.0 |
| 3 | All media including television | 7.0 |
| Total | | 100.0 |

It is clear that exposure to the mass media is not randomly distributed throughout the population. Obviously, the print media are only available to women who can read. (In 1989, 34.6 percent of Kenyan married women were illiterate). Listening to the radio or viewing television is associated (though not perfectly) with owning radios and television sets. These constraints imply socioeconomic connections with exposure to the media and thus to media messages on family planning. Residence in a city or region where such programs originate and where newspapers and magazines are more available is also expected to influence media exposure. Religion and ethnicity may capture cultural dimensions related to media exposure. There are also the possible effects of life cycle variables such as age and number of children, which might be expected to influence interest in or receptivity to messages on fertility regulation. In all, we selected a set of seventeen variables which may affect media exposure. These variables will be represented in our regression models by a total of 41 regressors.[23]

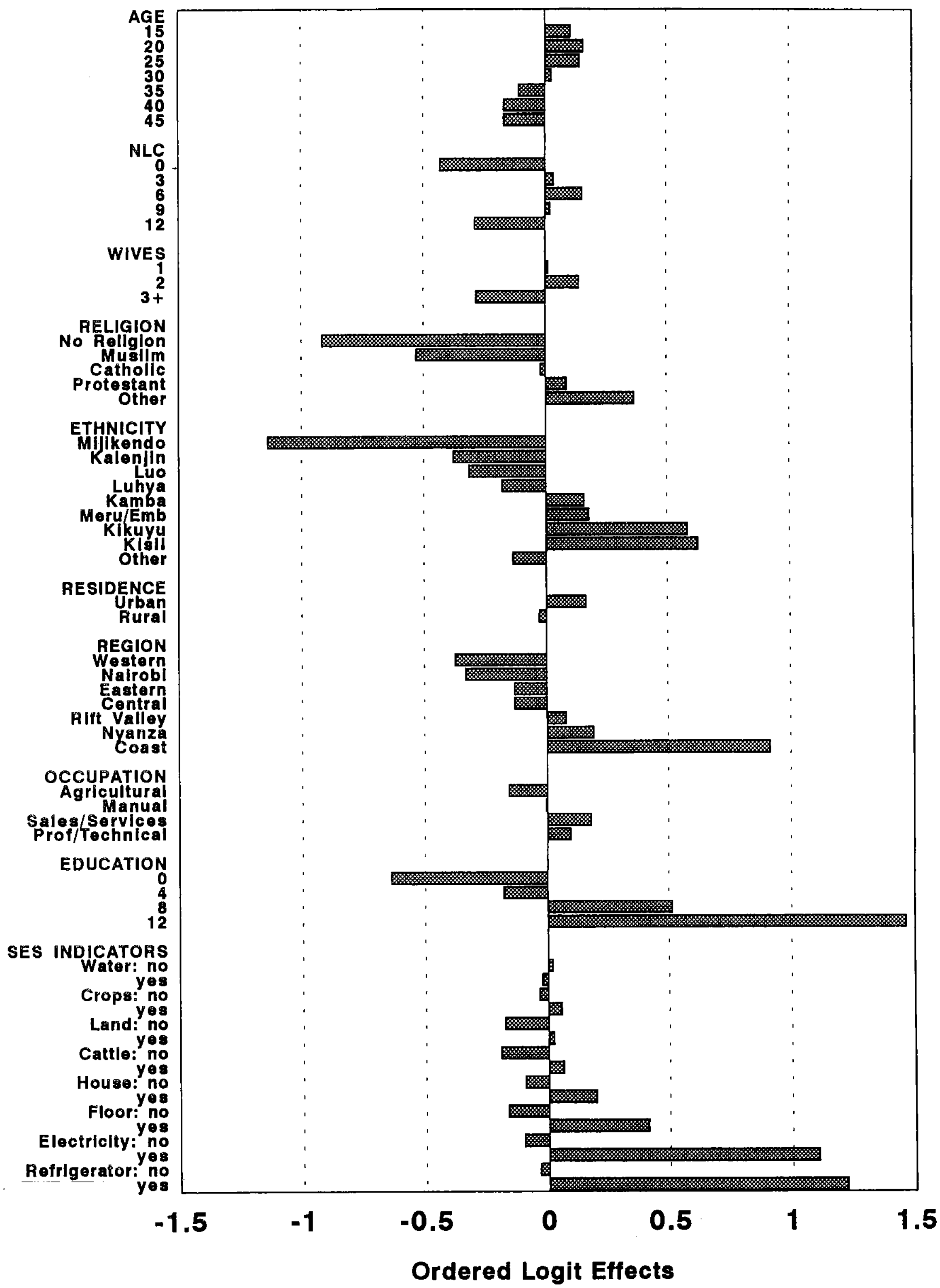
In order to assess the importance of such variables on media exposure score we tried several modelling strategies. The alternatives available for ordered categorical variables such as the media score include multinomial, hierarchical and ordered logit models, as well as ordered probits. The former two models were found to fit the data better than the latter two, but they required nearly three times as many parameters (126 instead of 44 for our full model). In order to take into account parsimony as well as goodness of fit we calculated Akaike's information criterion for each model.[24] On the basis of this criterion, the four models were practically indistinguishable, and we selected the ordered logit model because of its simplicity and to maintain consistency with the remaining analyses. The coefficients in an ordered logit model are similar to those in an ordinary logit model, measuring the effect of each variable on the log-odds, except that the log-odds pertain to the *cumulative* distribution of the response. The results are summarized in Figure 1.[25]

Exposure to media messages increases with age up to age 22 and then declines, in what is probably a generational effect. Similarly, exposure increases with the number of living children up to six but then declines, in an inverted U-shape relationship. The increase is consistent with the notion that women with children are more likely to pay attention to messages dealing with family issues. The decline at very high parities, even after controlling for age, is puzzling. Polygamy is also associated with exposure to the media, with wives whose husbands have at least two other wives being least likely to have heard family planning messages, even after taking into account other demographic and social characteristics. We find large effects of religion and ethnicity on exposure to messages, with women who declare no religion and women from the Milikendo ethnic group being those least exposed to family planning messages. As shown in Figure 1, these cultural effects are substantially larger than the demographic effects of age and number of living children. Type of place of residence has a modest effect in the expected direction, with urban residents more likely to have heard messages from more sources. There are fairly large regional differences, with Nairobi surprisingly at the extreme of least exposure (after adjusting for urban-rural and socioeconomic effects) and the Coast at the extreme of most exposure to messages about family planning. Exposure also increases with husband's occupation as one moves from agricultural to manual to white collar occupations.

Wife's education has the largest effect of all variables studied. The contrast between women with eight years of education and those with no schooling, for example, is 1.15 in the scale of log-odds or 3.16 in terms of an odds ratio. This means that eight years of education are associated with a three-fold increase in the odds of hearing a message on some medium (radio, paper or TV). Since the model restricts the odds ratio to be constant along the media scale, it is also true that the odds of having heard a message on television for a woman with eight years of education are three times as high as for a woman with no education. Finally, we note that some of the socioeconomic indicators are strong predictors of high exposure to media messages, particularly having a covered floor, having electricity and owning a refrigerator. Note, however, that these large effects apply to rather small minorities; for example only three percent of the women in our sample have a refrigerator.

The basic message of this analysis is that exposure to media messages about family planning is affected very much as expected by the usual roster of demographic, socioeconomic and cultural characteristics of the respondents. We therefore need to control for these attributes in our analysis of the association between media exposure and reproductive behavior.

Figure 1. Demographic and Socioeconomic Determinants of Exposure to Family Planning Messages in the Media, Kenya 1989



## Media Messages and Reproductive Behavior

The main analytical focus is on the association between mass media messages on family planning and the practice of contraception and reproductive preferences. We explore this association in terms of several indicators of reproductive behavior.

### Adoption of Contraception

The contraceptive variable can be conceived as a continuum ranging from women who have never heard of any method to those currently using some method, with intermediate gradations in terms of intentions to use and use in the past. The relationship between the media scale and a more or less ordered progression of contraceptive practice is shown in Table 3.[26]

Table 3. Exposure to Media Messages on Family Planning by Contraceptive Knowledge and Use, Kenya 1989

| Contraceptive Status | Media Messages: No Messages | Radio Only | Radio and Print | Radio, Print and TV | Total |
|---|---|---|---|---|---|
| Do not know method | 19.1 | 3.6 | 1.0 | 0.7 | 7.6 |
| Know method, never used, do not intend to use | 28.8 | 23.9 | 13.8 | 7.9 | 22.1 |
| Never used, intend to use | 26.0 | 28.6 | 21.6 | 14.6 | 25.3 |
| Used in past, do not intend to use | 5.6 | 5.9 | 6.5 | 6.9 | 6.0 |
| Used in past, intend to use | 6.3 | 12.0 | 17.7 | 20.2 | 12.0 |
| Currently using | 14.3 | 26.0 | 39.4 | 49.7 | 26.9 |
| Percent total | 100 | 100 | 100 | 100 | 100 |
| Number of women | 1441 | 1948 | 1014 | 375 | 4778 |

The association is quite strong. Knowledge of family planning methods, intentions to use contraception, and the history of contraceptive practice are all related to media exposure in the expected directions. Of course, as the earlier analysis of the characteristics of women reporting these messages indicated, the association may be simply a reflection of the joint connections of media exposure and contraceptive behavior with the family life cycle or with socioeconomic factors. To explore this issue, we now use a series of multivariate models of indicators of reproductive behavior to introduce controls for observed demographic and socioeconomic characteristics of respondents. The models used include ordinary linear regression, logit regression and hierarchical logit regression, as explained below. The results are summarized in Table 4 in terms of the effect of media exposure on each of the indicators studied, before and after adjustment for the controls.

Table 4. Effects of Exposure to Family Planning Media Messages on Selected Indicators of Reproductive Behavior Unadjusted (U) and Adjusted (A) for Socioeconomic Controls, Kenya 1989

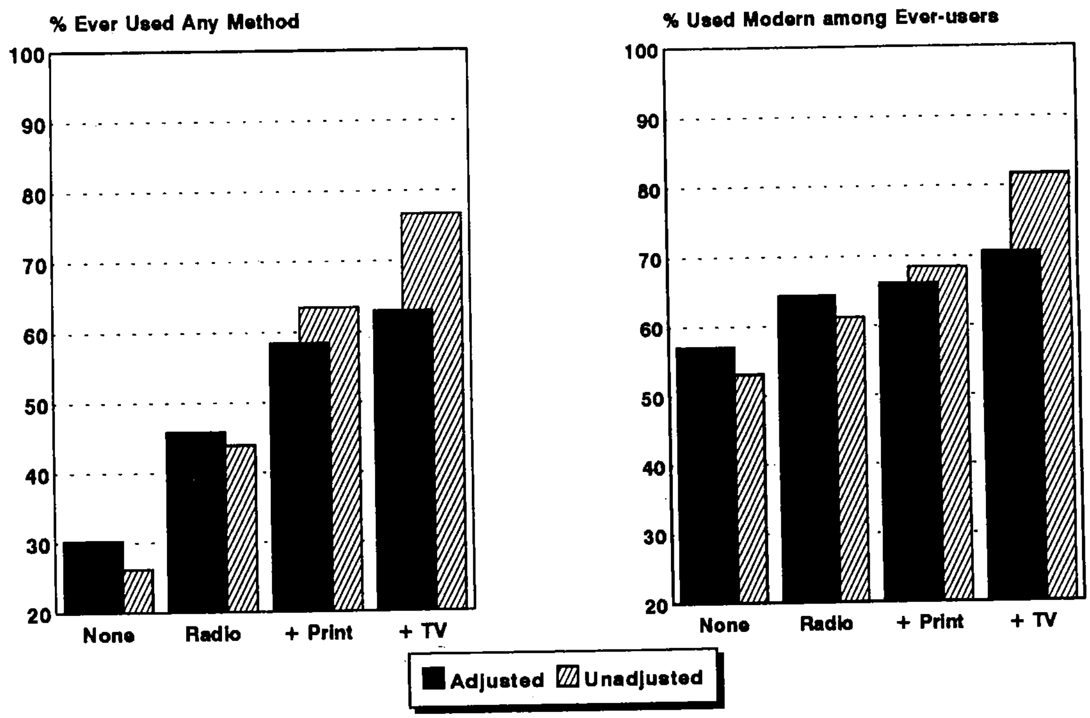
| Variable Indicator | | Family Planning Messages Score | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 | All | Chi-sq or F | df |
| Ever use of Contraception | | | | | | | | |
| % Ever used any method | U | 26.2 | 43.9 | 63.5 | 76.7 | 44.9 | 499.1 | 3 |
| | A | 30.3 | 45.8 | 58.4 | 62.9 | 44.9 | 147.0 | 3 |
| % Used modern among users | U | 53.1 | 61.2 | 68.4 | 81.7 | 64.3 | 64.4 | 3 |
| | A | 57.0 | 64.4 | 66.1 | 70.6 | 64.3 | 7.3 | 3 |
| Current Use of Contraception | | | | | | | | |
| % Currently using any method | U | 14.3 | 26.0 | 39.4 | 49.7 | 26.9 | 284.4 | 3 |
| | A | 18.1 | 27.8 | 35.7 | 33.4 | 26.9 | 66.5 | 3 |
| % Using modern among users | U | 53.2 | 62.9 | 71.8 | 81.2 | 66.4 | 39.4 | 3 |
| | A | 52.6 | 67.2 | 67.6 | 78.5 | 66.4 | 13.6 | 3 |
| Intention to Use in Future | | | | | | | | |
| % Intend to use in future | U | 40.4 | 56.9 | 65.7 | 69.5 | 53.2 | 150.4 | 3 |
| | A | 44.5 | 56.4 | 60.3 | 63.6 | 53.2 | 41.2 | 3 |
| Desire for Future Births | | | | | | | | |
| % Want no more | U | 46.8 | 49.6 | 51.6 | 52.5 | 49.4 | 7.3 | 3 |
| | A | 44.9 | 48.9 | 52.5 | 61.5 | 49.4 | 14.1 | 3 |
| % Spacers among non-limiters | U | 49.7 | 52.5 | 65.4 | 63.5 | 54.9 | 36.9 | 3 |
| | A | 54.2 | 51.1 | 61.5 | 62.1 | 54.9 | 11.6 | 3 |
| Ideal Number of Children | | | | | | | | |
| Mean ideal number | U | 5.97 | 4.98 | 4.55 | 3.82 | 5.11 | 121.8 | 3 |
| | A | 5.50 | 4.96 | 4.98 | 4.69 | 5.11 | 20.4 | 3 |

Adjusted figures are based on linear regression for ideal number of children and nested logit regression models for all other indicators, and have been scaled to reproduce exactly the sample total. The regression models include as controls: age, number of living children, number of wives, region and place of residence, wife's education, husband's occupation, several indicators of household possessions, ethnicity and religion.

**Ever Use of Contraception**

Figure 2 shows the results for ever use of contraception, which has three categories: never used, used traditional methods, and used modern methods. In our analysis we used hierarchical logit models, first contrasting use versus no use and then exploring the choice of modern or traditional method among users. An advantage of this model is that it can be fit simply as a series of two ordinary logistic regression analyses. We find that the proportion of women who have ever used a method increases monotonically as we move up the media scale, from having heard no messages to having heard messages on the radio only, in print and radio; and on TV, print and radio. The differential is slightly attenuated but remains highly significant after adjusting for all other variables in a model involving 41 additional terms. The same is not true, however, of method choice. Exposure to family planning messages apparently increases the likelihood that a user will choose a modern as opposed to a traditional method, but this effect falls just below significance after introducing the demographic and socioeconomic controls.

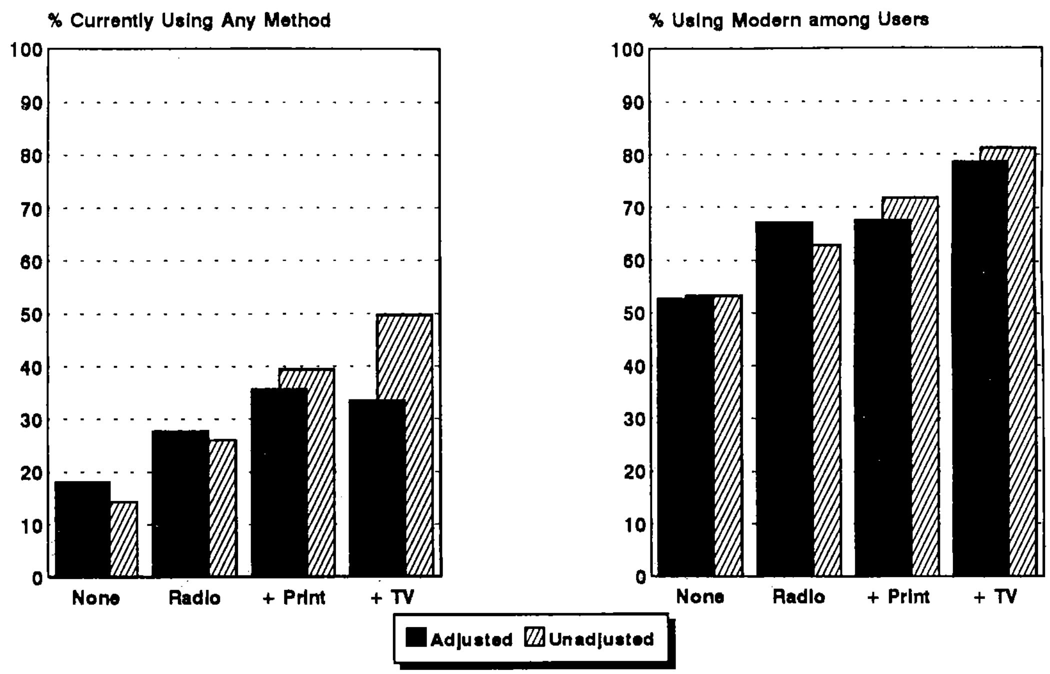
Figure 2. Exposure to Media Messages on Family Planning and Ever Use of Contraception, Kenya 1989




## Current Use of Contraception

Figure 3 shows the results for current contraceptive use. This variable also has three categories: not using, using traditional, and using modern methods, and was modelled using hierarchical logits with the same contrast as ever-use. The contraceptive prevalence rate increases from 14 percent among women who have not heard messages to nearly half of those who have heard messages on TV. The differential is attenuated after introducing the standard set of controls, but remains large and significant. Unlike the case of ever-use, media exposure appears to have an effect on the choice of method among current users: nearly eight of ten users who have heard a family planning message on TV choose a modern method, as opposed to just over half of those who have heard no messages at all. This effect remains significant after adjustment for all other variables.

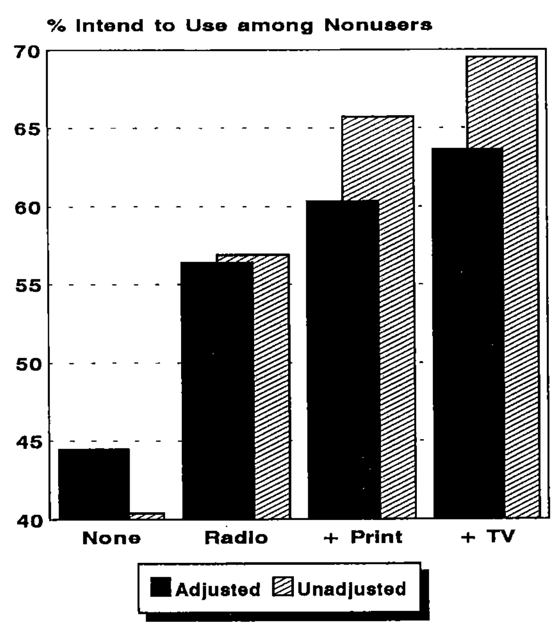
Figure 3. Exposure to Media Messages on Family Planning and Current Use of Contraception, Kenya 1989




## Intentions to Use in Future

We now turn our attention to women who are not using contraception, a substantial majority of our sample. Following exploratory analyses where we used different categorizations of this variable, we settled on a simple contrast between those who intend to use in the future and the remainder (including those undecided). The logistic regression analysis summarized in Figure 4 shows a strong association between exposure to family planning messages and declaring an intention to use contraception in future, even after partialling out the confounding influences of all other variables.

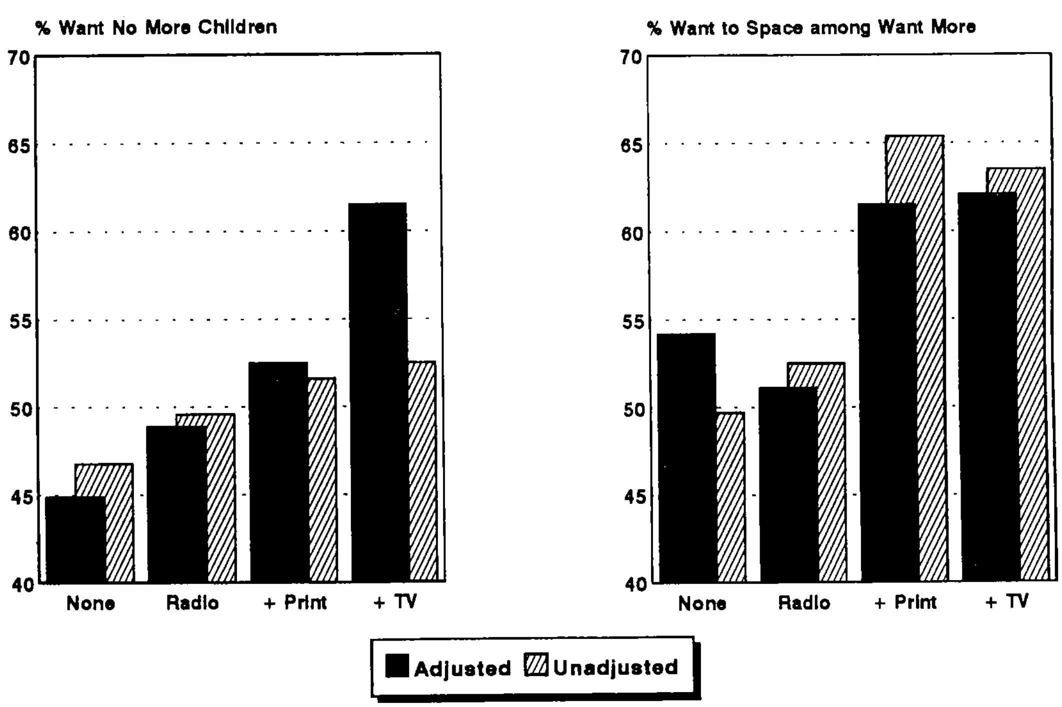
Figure 4. Exposure to Media Messages on Family Planning and Intention to Use, Kenya 1989



**Desire for Future Births**

The next variable on our list deals with fertility preferences. We have distinguished three types of women: those who would like to have a birth as soon as possible, those who would prefer to wait two years (hereafter called *spacers*) and those who would like to have no more children (hereafter called *limiters*). After exploring different ways of setting up the two contrasts required for a variable with three categories, we decided first to compare limiters with the rest, and then look at spacing among women who want more children. An interesting feature of the results shown in Figure 5 is that the effect of exposure to media messages on desire to limit fertility does not become evident until demographic and socioeconomic controls are introduced. The explanation is that women who have been exposed to family planning messages in more media are disproportionately young and at low parities, and hence unlikely to have reached a stage in their reproductive careers where they would want to limit births. After controlling for the number of living children the effect becomes immediately obvious, albeit not of a magnitude comparable to the effect of media on ever use of contraception. The second contrast shows that even among women who do not want to limit their fertility, exposure to media messages is associated with an increased desire to space births.

Figure 5. Exposure to Media Messages on Family Planning and Fertility Preferences, Kenya 1989

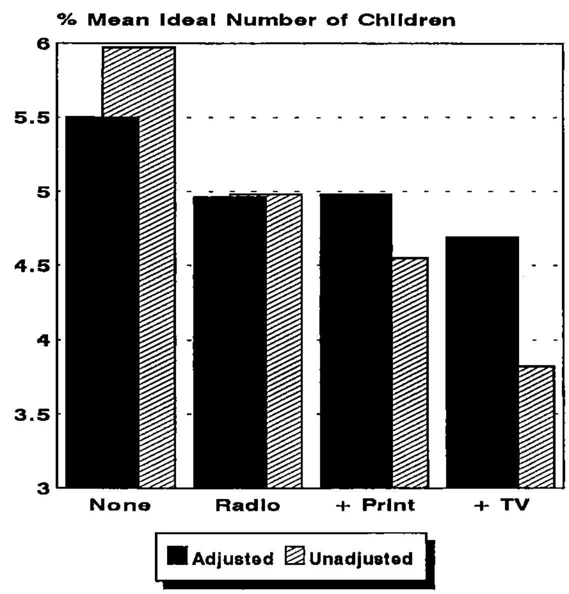


### Ideal Number of Children

The final indicator of reproductive behavior is the declared ideal number of children, which was analyzed using ordinary linear models.[27] As shown on Figure 6, the ideal number of children declines monotonically as one moves up the media scale. The difference between the extremes remains a significant eight-tenths of a child after controlling for all available observed characteristics.

## The Problem of Selectivity

Figure 6. Exposure to Media Messages on Family Planning and Ideal Number of Children, Kenya 1989




We have repeatedly called attention to the issue of the direction of causation between exposure to mass media and reproductive behavior. All we have been able to establish empirically in this analysis is that there is a strong correlation between reporting having heard or seen messages on family planning and, for example, the use of contraception. We have been able to exclude the possibility that the relationship is due to the joint association with life cycle, residential or socio-economic variables. The strong association that persists does not prove that these information and motivational efforts have had the intended effects on reproductive behavior, but such a presumption would certainly have been seriously undermined had the expected association been absent. However, we do not have any information on the time sequences of exposure and use (at least in these Kenya data) except for the general information that one of the main radio efforts occurred in the year prior to the survey. But it remains entirely possible that women who had already used contraception might simply be more sensitive to media messages on the subject than women who have not used a method. The anecdotal evidence from clinics and from some evaluation surveys strongly supports the inference that media messages have a significant impact on both motivation to limit fertility and on information about the availability of supplies. It would be reassuring, however, if there were internal evidence from the survey that would increase our confidence in the direction of the association.

One approach is to classify the sample by characteristics which are functionally related to contraceptive behavior and compare the groups on exposure to the media. For example, married and unmarried women or sexually active or inactive women should be equally exposed to media messages (at the same age). If married or sexually active women (with the appropriate age controls) report such messages more frequently than unmarried or sexually inactive women, one would be concerned about the force of selectivity. Of course, if there are no differences between these groups, we still cannot definitively conclude that selectivity is absent; the test may simply be inadequate. For what it is worth, such an analysis failed to reveal any significant or even patterned differences.

## Summary

This analysis has attempted to determine the case that can be made for the proposition that mass media efforts can influence reproductive behavior. Based on data from Kenya collected in 1989 as part of the Demographic and Health Surveys project, we have demonstrated the presence of a strong statistical association between reports of having heard or seen messages about family planning on radio, in newspapers or magazines or posters, or on television and various measures of reproductive behavior, in general the use of contraception and reproductive preferences. These associations are seen to persist even when a variety of life cycle, residential, and socioeconomic controls are imposed. There is a nagging problem of selectivity inherent in this kind of cross-sectional study, namely trying to elucidate the directions of causation; it could be that women who are already using contraception would simply be more sensitized to any exposure to media messages on the subject of family planning. The internal probes for such selectivity do not support this concern, but only a "before and after" design would satisfactorily resolve the issue. Based on the persistence of such strong relationships and in the light of anecdotal and other evidence from Kenya and elsewhere, we believe that the mass media can have an important effect on reproductive behavior.

## NOTES

1. William J. McGuire, "The Myth of Massive Media Impact: Savagings and Salvagings," in George Comstock, ed., *Public Communication and Behavior*, Academic Press, 1986, pp. 173-257.

2. Richard H. Gilluly and Sidney H. Moore, "Radio — Spreading the Word on Family Planning," *Population Reports* J(32):J884-886, September-October 1986. See also Cathleen A. Church, "Lights! Camera! Action! Promoting Family Planning with TV, Video and Film," *Population Reports* J(38), December 1989.

3. S.S. Lieberman, R. Gillespie, and M. Loghmoni, "The Isfahan Communications Project," *Studies in Family Planning* 4(4):73-100, April 1973.

4. Bogue, Donald J., Amy Ong Tsui, and Delia Barcelona, *Communicating Family Planning to Egypt*, A Report on the State Information Service Communication Campaign to Promote Contraceptive Adoption, 1980-1982, U.S. Agency for International Development/Cairo, October 1982.

5. "Awareness is Not Enough," *People* 13(2), 1986.

6. Phyllis T. Piotrow, et al., "Mass Media Family Planning Promotion in Three Nigerian Cities," *Studies in Family Planning* 21(5):272, September/October 1990.

7. Reported in Robert E. Lande and Judith S. Geller, "Paying for Family Planning," *Population Reports* J(39):16, November 1991.

8. Felipe Korzenny, G. Blake Armstrong and Tatiana Galvan, "Mass Communication, Cosmopolite Channels and Family Planning Among Villagers in Mexico," *Development and Change* 14:237-253, Sage: London, Beverly Hills and New Delhi, 1983.

9. *Ibid*, p. 249.

10. Jane T. Bertrand, et al. "Family Planning Communications and Contraceptive Use in Guatemala, El Salvador and Panama," *Studies in Family Planning* 13(67):190-199, June/July 1982.

11. Because of the inherent difficulties in achieving pure randomization with human subjects, the experimental design is far from a panacea. See Lawrence Wallack, "Mass Media Campaigns: The Odds Against Finding Change," *Health Education Quarterly* 8(3):209-261, Fall 1981.

12. Such a question or questions were included in Bolivia, Colombia, the Dominican Republic, Ecuador, Egypt, Ghana, Guatemala, Indonesia, Kenya, Morocco, Peru, Sri Lanka, Thailand, Trinidad & Tobago, Tunisia and Zimbabwe.

13. José G. Rimon and Cheryl Lettenmeier, "Strategic Options for IEC Interventions in Kenya," mimeo, Population Communication Services, Johns Hopkins University, March 29, 1990.

14. Radio broadcasting began in Kenya in 1927, preceded only by South Africa in 1924. See Graham Mytton, *Mass Communication in Africa*, Edward Arnold Publish., London, 1983, p. 52.

15. Sheila Rule, "To Cut Rates, Kenya Turns to TV Show," *New York Times*, June 14, 1987.

16. Population Communications International, 1990 Annual Report, p. 1.

17. Alamin Mazrui, Salma Mazrui, Negeta Kabiri and Ngala Mutunga, "A Formative Survey of 'Ushikwapo Shikamona'," mimeo, National Council for Population and Development, October 1988.

18. *Ibid*, p. 59.

19. Population Communications International, *op cit.*, p. 2.

20. A.W. Inambao and Gary L. Lewis, "The Kenya Family Planning IE&C Survey," Center for African Family Studies Research Unit, mimeo, 1989.

21. In Merton and Lazarsfeld's terms, this is part of the concept of "supplementation," one of the necessary conditions for effective mass communication. See Paul F. Lazarsfeld and Robert K. Merton, "Mass Communication, Popular Taste and Organized Social Action," reprinted in W. Schramm (ed.), *Communications*, Urbana IL: University of Illinois Press, 1975.

22. See for example Guttman, L. "The basis for scalogram analysis", Pages 60-90 in S.A. Stouffer *et al.* *Measurement and prediction*, Princeton: Princeton University Press, 1950.

23. The seventeen variables and the way in which they were included in the model are as follows. Age and number of living children were modelled using regression splines; specifically, we used natural cubic splines with interior knots at the terciles. The distribution of the number of wives has a rather long tail; we decided to group it into three categories (1,2, and 3+) and represent these with two dummy variables. Religion is treated as a factor with five categories, including no religion, Catholic, Protestant, Muslims and others. Ethnicity has nine categories represented by eight dummy variables, after combining a very small group of Somalis with the original "other" category. Region and place of residence where handled in terms of an urban/rural dichotomy and a set of seven regions, including Nairobi, Central, Coast, Eastern, Nyanza, Rift Valley and Western. There is some evidence that the effect of urban residence may vary by region, but unfortunately the sample size does not allow reliable estimation of such interactions. Husband's occupation was coded using four categories: agricultural, manual, sales and services, and professional, technical and clerical. Wife's education was measured in years and modelled using a natural cubic spline with interior knots at the terciles. The final set of variables is a set of socioeconomic indicators, dealing with the possession of a house, running water, floor covering, electricity, a refrigerator, land, cattle and crops. We gave some thought to the idea of combining these indicators in a socioeconomic index, but in the end preferred to maximize their potential explanatory power by using a separate dummy variable for each indicator.

24. The Akaike information criterion (AIC) basically adds twice the number of parameters to the model's deviance, see Akaike, H. "Information theory and an extension of the maximum likelihood principle", pages 267-281 in B.N.Petrov and F. Csàki, Editors, *Second International Symposium on Information Theory*. Budapest: Akademia Kiadó, 1973.

25. We are grateful to Thomas Espenshade for a suggestion leading to Figure 1. The ordered logit effects shown have been standardized to have a weighted sample average of zero for each variable, thereby making the appearance of the graph independent of the choice of reference cell.

26. This analysis was suggested by a similar typology in Bertrand *et al.*, *op. cit.*, pp. 195-196. The apparent anomaly of small numbers of women who never heard of particular contraceptive methods who report family planning messages from the media could arise from the content of some messages about the general advantages of smaller families.

27. Values in excess of 12 and non-numeric responses such as "it's up to God" were assigned a value of 12. Altogether, 5.5 percent of the responses were so assigned.

# Appendix: Methodology

## Appendix: Methodology

In this appendix we explain some of the more arcane aspects of the methodology used in the analysis, in the hope that the details will be useful to researchers attempting to replicate our results or conducting studies along similar lines. We start with a brief discussion of the construction of the media scale, which required examining a four-way cross-tabulation. We then move on to the use of regression splines to handle continuous predictors, describing how one can use standard software to fit cubic regression splines and natural splines. Next we describe the calculation of the ordered logit effects shown in Figure 1, explaining our rationale for centering the estimates. Finally, we describe the computation of the adjusted and unadjusted effects of exposure to media messages shown in Table 4, which play a central role in the analysis.

There are several statistical packages that can be used to fit the models used in this study. We used Stata (Computing Resource Center, Santa Monica, California), a fast and user-friendly package that includes facilities for data manipulation and for fitting linear regression, logistic regression, ordered logit and multinomial logit models. We use Stata notation to specify a couple of recode instructions at the end of Section 2 below, but the expressions are easily translated into other languages. Similarly, we show the actual raw output from four Stata runs in Tables A.3, A.4, A.6 and A.7, so that the interested reader can see *all* the steps involved in obtaining our results; the output from other packages may be presented in a somewhat different format, and perhaps using different conventions, but the general ideas will be the same.

### 1. The Media Scale

Questions about exposure to media messages on family planning were not part of the core questionnaire in DHS-I. They were, however, included in one form or another in 15 countries other than Kenya: Bolivia, Colombia, Dominican Republic, Ecuador, Egypt, Ghana, Guatemala, Indonesia, Morocco, Peru, Sri Lanka, Thailand, Trinidad and Tobago, Tunisia and Zimbabwe. The most common form was a single question on whether a woman had heard or seen messages on the radio or television in the past month. In some countries the reference was the last six months, or whether she had ever heard or seen such messages in the media.

The questions in the Kenya survey are among the most detailed. They include whether the woman had heard or seen family planning messages in four media: radio, newspapers, posters and television. The implications of these intercountry differences in the coverage of the media topic is that the media exposure variable will have to be treated as a dichotomy in most of the countries. In Kenya, however, a richer analysis is possible.

The obvious first step in examining how to handle multiple indicators is to cross-tabulate them. Most statistical packages will present a four-way table as an unwieldy series of two-way tables (row and columns) for each combination of categories of the other two variables (layer and panels). We used the standard trick of combining three of the variables using a three-digit code and then tabulating this combined variable against the fourth. In the case at hand, all variables have been coded (or recoded) as 1 (yes), 2 (no) or 9 (not stated), so we calculated

$$\text{combine} = 100 * \text{radio} + 10 * \text{paper} + \text{poster}$$

and then tabulated the combined index against television to obtain Table A.1

Table A.1 Joint distribution of exposure to family planning (FP) messages through radio, newspapers, posters and television, Kenya 1989

| Heard through radio, newspapers, posters | Exposed to FP messages on television | | | |
|---|---|---|---|---|
| | 1 | 2 | 9 | Total |
| 111 | 271 | 368 | 0 | 639 |
| 112 | 83 | 265 | 0 | 348 |
| 119 | 0 | 2 | 0 | 2 |
| 121 | 21 | 330 | 0 | 351 |
| 122 | 35 | 1759 | 1 | 1795 |
| 129 | 1 | 0 | 0 | 1 |
| 191 | 0 | 2 | 0 | 2 |
| 199 | 0 | 0 | 2 | 2 |
| 211 | 7 | 36 | 2 | 45 |
| 212 | 2 | 22 | 0 | 24 |
| 221 | 2 | 119 | 0 | 121 |
| 222 | 6 | 1431 | 0 | 1437 |
| 229 | 0 | 2 | 0 | 2 |
| 291 | 0 | 1 | 0 | 1 |
| 999 | 0 | 0 | 8 | 8 |
| Total | 428 | 337 | 13 | 4778 |

This table shows that most of the responses fall into a few patterns. We decided to combine papers and posters in a single category representing print media, and merged the few cases with missing data in each indicator into the "no" category for that variable. Calculating a new index

$$\text{index} = 100 * \text{radio} + 10 * \text{print} + \text{TV}$$

and running a marginal frequency we obtain the "3-way" Table A.2.

Note that most of the responses fall into one of the patterns 222, 122, 112 and 111, effectively forming what's known as a cumulative scale (Guttman, 1950). In view of this result we constructed our media scale simply counting the number of "yes" answers to exposure to radio, print and TV. A nice property of the resulting score is that the actual response pattern can be reproduced from the total score for a large fraction of the cases. For example, a score of 2 tends to mean that the woman has been exposed to messages on the radio and print but not on TV (which is in fact the case for 967 of the 1014 women with scores of 2).

## 2. Regression Splines

The next problem we consider is how to code the regressors or predictor variables. Many of these are categorical variables, and were represented using binary or dummy variables in the standard way. For example religion has five categories: no religion, Catholic, Protestant, Muslim and other, and was represented using four dummy variables with "no religion" as the omitted category. Other variables handled in this fashion were ethnicity, region, place of residence, husband's occupation and a host of dichotomous socioeconomic indicators.

**Table A.2. Joint distribution of exposure to family planning (FP) messages through radio, print media (newspapers and posters), and television, Kenya 1989**

| Heard through radio, print, TV | Freq. | Percent | Cum. |
|---|---|---|---|
| 111 | 375 | 7.85 | 7.85 |
| 112 | 967 | 20.24 | 28.09 |
| 121 | 36 | 0.75 | 28.84 |
| 122 | 1762 | 36.88 | 65.72 |
| 211 | 11 | 0.23 | 65.95 |
| 212 | 180 | 3.77 | 69.72 |
| 221 | 6 | 0.13 | 69.84 |
| 222 | 1441 | 30.16 | 100.00 |
| Total | 4778 | 100.00 | |

Number of wives is an example of a quantitative variable where most of the responses fall in a couple of categories (namely 1 or 2) but there is also a long tail, with one of the husbands having as many as 10 wives. Since the tail of the distribution is very thin, we decided to group this variable in three categories: 1, 2 and 3 or more wives, and coded it using two dummy variables.

This leaves three other quantitative variables: age, number of living children and woman's education, which was measured in completed years of schooling. Treating the actual value of one of these variables as a predictor in the model assumes that its effect is linear. Although this may indeed be the case, we wanted a more flexible approach that, ideally, would reduce to the simpler case if appropriate. An often-used alternative is to group the variable into categories and model it using a set of dummy variables. This approach is flexible, but grouping naturally throws away some information.

The alternative we adopted was to use regression *splines*. A spline is basically a piece-wise polynomial. Picture a standard set of coordinates axes and imagine putting a few marks along the x-axis, thereby dividing the range of the variable into segments. The markings are called *knots*. A spline consists of a series of polynomials, one for each segment, which are joined smoothly at the knots. One can construct splines of any degree, say linear, quadratic, cubic or higher. A linear spline with a single knot at a point $c$, say, consists of two straight lines which intersect at $c$, and is a convenient way of writing a "two-regime" model, where the slope changes along the way. The requirement that the pieces should join smoothly at the knots translates into constraints on their derivatives. Specifically, for a spline of degree $d$ we require the two polynomials meeting at each knot and their first $d-1$

derivatives to agree exactly at the knot. Cubic splines have been found to have pleasant properties including visual smoothness, and can be used to approximate almost any function by judicious choice of knots.

A cubic spline $S(x)$ with $k$ knots placed at points $\xi_1 < \xi_2 < \ldots < \xi_k$ can be written as follows:

$$S(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{j=1}^{k} \gamma_j (x-\xi_j)_+^3 \qquad (1)$$

where the $(\ )_+$ notation after the summation means that we use the cube of the difference between $x$ and each knot $\xi_j$ when $x$ is above the knot and 0 when $x$ is below the knot, i.e.

$$(x-\xi_j)_+^3 = \begin{cases} (x-\xi_j)^3, & x > \xi_j \\ 0, & x \le \xi_j \end{cases} \qquad (2)$$

This representation of the spline as a linear function of $x$ and the $k$ variables $z_j = (x-\xi_j)_+^3$ with unknown coefficients (called the power-series representation) suggests fitting the spline to data by OLS or logit regression using $x$ and $z_1,\ldots,z_k$ as predictors. The result is called a regression spline. A word of warning: the auxiliary variables $z_j$ are highly collinear, and a special-purpose package would certainly use a numerically more stable representation of splines (such as B-splines). In our experience, however, the simpler representation just described works reasonably well for a moderate number of knots, say half-a-dozen or so, and can be used with standard software.

Splines are much better behaved than polynomials at the extremes of the range, but one still worries about the use of cubic approximations in the tails of a distribution, particularly if the tails are thin. A solution to this problem is the *natural* cubic spline, which is a cubic spline restricted to be linear outside the range of the knots. Note that requiring the spline to be linear outside the range of the knots actually imposes some constraints inside the range, because of the requirement that the pieces join smoothly. It can be shown that a natural cubic spline can be represented as

$$S(x) = \beta_0 + \beta_1 x + \sum_{j=1}^{k} \gamma_j (x-\xi_j)_+^3 \qquad (3)$$

where the $\gamma_j$ coefficients are subject to two constraints: $\sum_{j=1}^{k} \gamma_j = 0$ and $\sum_{j=1}^{k} \gamma_j \xi_j = 0$. Note that we drop two terms from the general representation (1) and add two constraints, saving a total of four parameters. One usually adds a knot at the minimum and one at the maximum, however, so the effective savings is only two parameters.

The easiest way to handle the two constraints on the $\gamma$ coefficients is to reparametrize the model eliminating redundant parameters. In the present case, solving for $\gamma_k$ and $\gamma_{k-1}$ in terms of the remaining *k-2* parameters shows that (3) can be fit by regression using as predictors $x$ and $z_1^*, \dots, z_{k-2}^*$, where

$$z_j^* = z_j - \frac{\xi_j - \xi_k}{\xi_{k-1} - \xi_k} z_{k-1} + \frac{\xi_j - \xi_{k-1}}{\xi_{k-1} - \xi_k} z_k \tag{4}$$

Again, a special-purpose package would use a numerically more stable basis for natural cubic splines, but we have found the basis given at (3) to be easy to use and quite adequate when the number of knots is moderate.

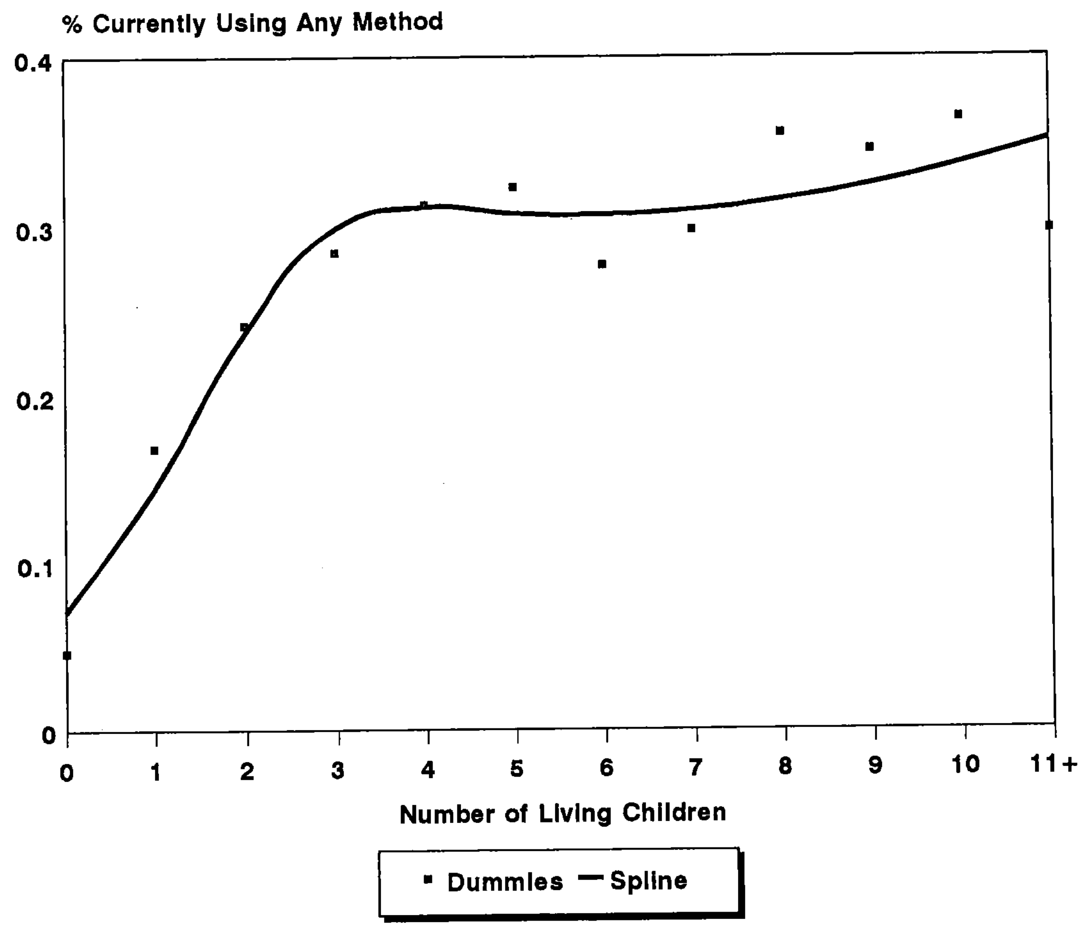
In the case of age, number of living children and education, we found that natural cubic splines with interior knots placed approximately at the terciles of the distribution provided a nice representation of the effects. In the case of age, for example, the knots were placed at 14.5, 25.5, 34.5 and 49.5 years (we used ages ending in 0.5 so that none of the knots coincided with an actual data value, but this is not necessary). To fit a spline on age, we use as predictors age itself (representing the linear part of the spline) and the following two auxiliary variates (written in Stata notation), based on formula (4).

```
age1 = (age-14.5)^3 - (age>34.5)*(age-34.5)^3*(49.5-14.5)/15
age2 = (age>25.5)*(age-25.5)^3 - (age>34.5)*(age-34.5)^3*(49.5-25.5)/15
```

In Stata, the notation age > 25.5 results in the value 1 if age exceeds 25.5 and 0 otherwise. Multiplying the result by (age-25.5)^3 then gives (age-25.5)^3 when age exceeds 25.5 and 0 otherwise. The spline for number of living children had knots at -0.5, 2.5, 4.5 and 16.5. The spline for years of education had knots at -0.5, 1.5, 6.5 and 18.5 (with 99s treated as no education).

In order to show how good a job these splines do, we ran a couple of logistic regressions of current contraceptive use on number of living children. In the first run we treated surviving children using a separate dummy variable for each value up to 10 and another for 11 or more (the categories 12 to 16 have too few cases, 28 in all, to support separate estimation). In the second run we used the natural cubic spline just described. Figure A.1 compares the two fits. As can be seen, the spline provides an excellent fit with only three parameters other than the constant, compared to 10 for the dummy variables.

Figure A.1 Contraceptive Use by Number of Living Children, Kenya 1989




### 3. Profiling Estimated Effects

The next problem we consider is how to present in a concise fashion the results of fitting a regression model. We will use as illustration the ordered logit model used to study the effects of various socioeconomic factors on media exposure. Table A.3 shows the raw output from Stata.

Consider first a categorical variable such as urban/rural. The results show that the cumulative log-odds are 0.195 higher for urban than for rural. This means that the odds of having been exposed to family planning messages in any media are 21.5 percent higher in urban than in rural areas (exp(0.195)=1.25). Due to the nature of the model, the odds of having heard messages on television, as opposed to no messages or messages in other media, are also 21.5 percent higher in urban than in rural areas. A similar analysis can be done for religion, or for any of the other categorical variables in the model.

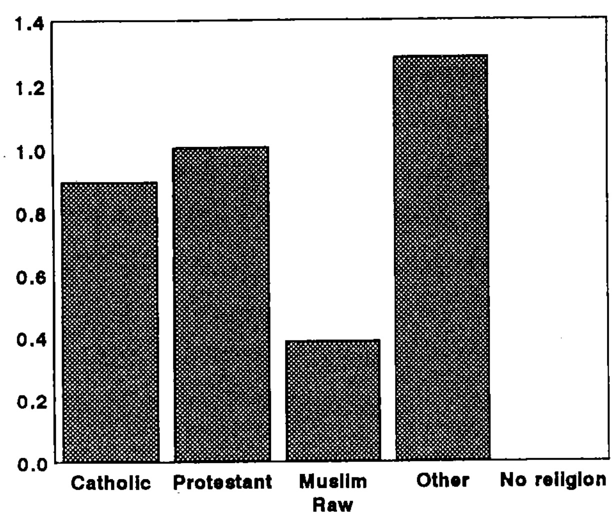
Table A.3 Raw output from ordered logit regression of media exposure on demographic and socioeconomic controls, Kenya 1989

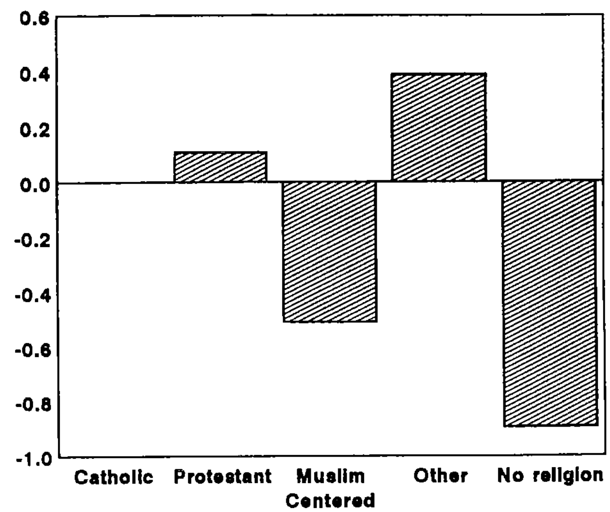
| Ordered Logit Estimates | | | | | Number of obs | = 4778 |
|---|---|---|---|---|---|---|
| | | | | | chi2(41) | = 1801.89 |
| | | | | | Prob > chi2 | = 0.0000 |
| Log Likelihood = -5024.937 | | | | | Pseudo R2 | = 0.1520 |

| media | Coefficient | Std Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|
| age | .0132188 | .0269118 | 0.491 | 0.623 | -.0395409 | .0659784 |
| age1 | -.0000804 | .0000626 | -1.284 | 0.199 | -.0002031 | .0000424 |
| age2 | .0002502 | .0001814 | 1.380 | 0.168 | -.0001054 | .0006058 |
| nlc | .195167 | .0747318 | 2.612 | 0.009 | .0486579 | .341676 |
| nlc1 | -.0027961 | .0023531 | -1.188 | 0.235 | -.0074093 | .0018171 |
| nlc2 | .0043557 | .0068129 | 0.639 | 0.523 | -.0090008 | .0177121 |
| wives2 | .1264686 | .0857213 | 1.475 | 0.140 | -.041585 | .2945222 |
| wives3 | -.2964672 | .1003914 | -2.953 | 0.003 | -.4932809 | -.0996535 |
| catholic | .8942021 | .2035052 | 4.394 | 0.000 | .4952372 | 1.293167 |
| protest | 1.002053 | .2015926 | 4.971 | 0.000 | .6068374 | 1.397268 |
| muslim | .3840043 | .241467 | 1.590 | 0.112 | -.0893834 | .8573919 |
| other | 1.281331 | .2964551 | 4.322 | 0.000 | .7001409 | 1.862521 |
| kalenjin | -.243824 | .1649933 | -1.478 | 0.140 | -.5672875 | .0796395 |
| kamba | .2949322 | .2445462 | 1.206 | 0.228 | -.1844921 | .7743564 |
| kikuyu | .7188005 | .1600552 | 4.491 | 0.000 | .4050179 | 1.032583 |
| kisii | .7628913 | .2003868 | 3.807 | 0.000 | .3700401 | 1.155743 |
| luhya | -.0437183 | .14708 | -0.297 | 0.766 | -.3320635 | .2446268 |
| luo | -.1783672 | .166886 | -1.069 | 0.285 | -.5055413 | .1488069 |
| meru | .3149812 | .2714243 | 1.160 | 0.246 | -.2171367 | .847099 |
| miji | -.9973431 | .238377 | -4.184 | 0.000 | -1.464673 | -.5300134 |
| urban | .1949038 | .1260986 | 1.546 | 0.122 | -.0523079 | .4421156 |
| nairobi | .0434619 | .1865189 | 0.233 | 0.816 | -.3222017 | .4091256 |
| central | .2432455 | .1627498 | 1.495 | 0.135 | -.0758198 | .5623108 |
| coast | 1.290781 | .216068 | 5.974 | 0.000 | .867187 | 1.714374 |
| eastern | .2424092 | .244783 | 0.990 | 0.322 | -.2374794 | .7222977 |
| nyanza | .5708148 | .1568515 | 3.639 | 0.000 | .263313 | .8783166 |
| rift | .4567668 | .1252518 | 3.647 | 0.000 | .211215 | .7023186 |
| manual | .1521976 | .0846532 | 1.798 | 0.072 | -.0137621 | .3181573 |
| saleserv | .339913 | .0769614 | 4.417 | 0.000 | .1890329 | .4907931 |
| proftech | .2540553 | .0865157 | 2.937 | 0.003 | .0844443 | .4236663 |
| edu | .227278 | .0652733 | 3.482 | 0.001 | .099312 | .355244 |
| edu1 | -.0078935 | .0038994 | -2.024 | 0.043 | -.0155381 | -.0002489 |
| edu2 | .0170322 | .0076666 | 2.222 | 0.026 | .0020021 | .0320622 |
| house | .2948472 | .0833012 | 3.540 | 0.000 | .1315382 | .4581562 |
| water | -.0424625 | .0652097 | -0.651 | 0.515 | -.1703038 | .0853787 |
| floor | .5785271 | .0922242 | 6.273 | 0.000 | .3977248 | .7593293 |
| electr | 1.20774 | .1392964 | 8.670 | 0.000 | .9346544 | 1.480826 |
| refrig | 1.255448 | .2347547 | 5.348 | 0.000 | .7952202 | 1.715677 |
| land | .2017926 | .1093598 | 1.845 | 0.065 | -.0126035 | .4161887 |
| cattle | .2578997 | .0784973 | 3.285 | 0.001 | .1040084 | .4117909 |
| crops | .0897655 | .0668189 | 1.343 | 0.179 | -.0412305 | .2207615 |
| _cut1 | 2.396834 | .5972088 | | | (Ancillary parameters) | |
| _cut2 | 4.684426 | .6005912 | | | | |
| _cut3 | 6.884129 | .604998 | | | | |

One way to represent these effects is by using a bar chart. The problem with this strategy, however, is that the appearance of the chart depends very much on the choice of reference category. Take religion, for example. Figure A.2 shows the graphs calculated using no religion as the baseline, and also using Catholics as the baseline. The greater black mass on the first graph may exaggerate the magnitude of the religion effect, particularly if put side by side with other variables. Thus, the appearance of Figure 1 in the text could be altered substantially by the choice of reference category for each variable.

Figure A.2. Religion effects using different reference cells, Kenya 1989




The solution adopted here was to *center* all the effects before plotting. Using religion as an example, we first calculated for each woman in the sample the religion effect, using zero for women with no religion. Then we calculated the weighted average of these effects (using the sample weights) and subtracted the mean from the raw effect. The new effects, which are centered on zero, are the values actually plotted in Figure 1.

There is also a serendipitous consequence of using weighted means, which is particularly appealing in the case of dummy variables. Consider as an example the possession of a refrigerator, which has a very large effect of 1.255 on the log-odds. This coefficient means that the odds of having heard messages of any kind are 2.5 times higher for women with refrigerators than for women without (exp(1.255) = 3.509). One could represent this effect as a single bar of length 1.255. The centered plot uses two bars instead, representing deviations from the average effect. Since women with refrigerators represent only 2.8 percent of the total sample, the average effect is 0.035. The deviations from this average are 1.22 for women with refrigerators and -0.035 for women without refrigerators. The fact that the two bars in Figure 1 are very disproportionate shows immediately that the effect of having a refrigerator is large but affects very few people. In contrast, the effect of having covered floors is smaller but affects a substantial number of women, as shown by the fact that the bars for covered floors in Figure 1 are more closely balanced.

Let us now turn our attention to variables modelled using splines. We follow basically the same procedure: we multiply the estimated coefficients by the variables representing the spline to obtain a prediction for each individual. For example for age we calculate

agefit = 0.132188 * age - 0.0000804 * age1 + 0.0002502 * age2

At this point one could (and should!) plot the age fit versus age to see how it looks. The result should be a continuous, smooth, and hopefully reasonable-looking curve. The next step is to center the effect, subtracting the weighted mean. The centered plot, suitably scaled and rotated, could be pasted directly into Figure 1. In order to make the figure stylistically homogeneous, however, we decided to evaluate the spline at typical values of $x$, (for example ages 15, 20, 25, 30, 35, 40, 45 and 50) and plot these using bars.

The final result is a plot that we believe fairly represents the effects of all variables on comparable scales.

## 4. Adjusted and Unadjusted Effects

A substantial part of our analysis focuses on the comparison of the effects of exposure to media messages on some measure of reproductive behavior *before* and *after* adjustment for the effects of demographic and socioeconomic variables.

We illustrate the methodology with the analysis of the effects of exposure to the mass media on current contraceptive use. Recall that we distinguished three categories of contraceptive use: not using, using a traditional method and using a modern method, and that we decided to analyze this variable in terms of two dichotomies representing behavioral choices: first we study the determinants of contraceptive use, and then we focus on users and study the determinants of method choice. Each of these contrasts is analyzed using a logistic regression model. Consider first the analysis of contraceptive use among all women. We start by fitting a model that has the media exposure variable as the sole predictor. The raw results appear in Table A.4.

This model serves to estimate the gross or *unadjusted* effect of media exposure. The first thing to note in the output is the likelihood ratio test for the unadjusted effect, which has a highly significant chi-square value of 284.4 with 3 degrees of freedom, as reported in Table 4. The next step is to predict the probability of using contraception for

Table A.4 Raw output from logistic regression of current contraceptive use on media scale, Kenya 1989

| Logit Estimates | | | | | Number of obs = 4778 | |
|---|---|---|---|---|---|---|
| | | | | | chi2(3) = 284.40 | |
| | | | | | Prob > chi2 = 0.0000 | |
| Log Likelihood = -2640.1751 | | | | | Pseudo R2 = 0.0511 | |
| useany | Coefficient | Std. Err. | t | P>\|t\| | [95% Conf. Interval] | |
| media1 | .740122 | .0907168 | 8.159 | 0.000 | .5622753 | .9179688 |
| media2 | 1.355501 | .0989456 | 13.699 | 0.000 | 1.161522 | 1.54948 |
| media3 | 1.77461 | .1327146 | 13.372 | 0.000 | 1.514428 | 2.034792 |
| _cons | -1.787711 | .0749457 | -23.853 | 0.000 | -1.934639 | -1.640783 |

each category of mass media exposure. The steps involved are adding the constant to the parameter estimates and calculating antilogits, and are done automatically by some packages (including Stata). The results are the unadjusted proportions shown in the column labelled "Unadjusted" in Table A.5 (and in Table 4 in the text). A nice property of these estimates is that the weighted average of the predicted proportions, calculated using the sample weights, agrees exactly with the overall proportion using contraception, the 27 percent shown in Table 4. In fact, the unadjusted proportions could have been obtained directly from a simple crosstabulation of contraceptive use by media exposure. Actually reproducing these numbers from the output, however, may prove instructive.

Table A.5 Calculation of adjusted media effects, Kenya 1989

| | | | Adjusted by method | | |
|---|---|---|---|---|---|
| Media score | % | Unadjusted | 1 | 2 | 3 |
| 0 | 0.3034 | 0.1434 | 0.1439 | 0.1898 | 0.1807 |
| 1 | 0.4167 | 0.2597 | 0.2269 | 0.2724 | 0.2782 |
| 2 | 0.2101 | 0.3936 | 0.2972 | 0.3355 | 0.3570 |
| 3 | 0.0698 | 0.4967 | 0.2766 | 0.3175 | 0.3342 |
| All | 1.0000 | 0.2691 | 0.2200 | 0.2637 | 0.2691 |
| | | | Parameter Estimates | | |
| constant | | -1.7877 | -1.7837[a] | -1.4513[b] | -1.5115[c] |
| media1 | | 0.7401 | 0.5579 | 0.4688 | 0.5579 |
| media2 | | 1.3555 | 0.9231 | 0.7681 | 0.9231 |
| media3 | | 1.7746 | 0.8221 | 0.6858 | 0.8221 |
| | | | Predicted Logits | | |
| 0 | | -1.7877 | -1.7837 | -1.4513 | -1.5115 |
| 1 | | -1.0476 | -1.2257 | -0.98254 | -0.9536 |
| 2 | | -0.4322 | -0.8606 | -0.68319 | -0.5884 |
| 3 | | -0.0131 | -0.9616 | -0.7655 | -0.6894 |

[a] Obtained setting all demographic and socioeconomic variables to their means.
[b] The values in this column are implied parameter estimates calculated *from* the proportions at the top, which were generated as explained in the text.
[c] Calculated so as to reproduce exactly the overall proportions using contraception.

The next step is to fit a model with both the media exposure variable and all demographic and socioeconomic controls. Table A.6 shows the raw output from Stata.

Table A.6 Raw output from logistic regression of current contraceptive use on media scale and demographic and socioeconomic cohorts, Kenya 1989

| Logit Estimates | | | | | Number of obs = 4778<br>chi2(44) = 1015.97<br>Prob > chi2 = 0.0000 | |
|---|---|---|---|---|---|---|
| Log Likelihood = -2274.3877 | | | | | Pseudo R2 = 0.1826 | |
| useany | Coefficient | Std. Err. | t | P>\|t\| | [95% Conf. Interval] | |
| media1 | .5579431 | .0999202 | 5.584 | 0.000 | .362053 | .7538333 |
| media2 | .9230868 | .1166211 | 7.915 | 0.000 | .6944552 | 1.151718 |
| media3 | .8220695 | .1734543 | 4.739 | 0.000 | .4820185 | 1.162121 |
| age | -.1245641 | .0415901 | -2.995 | 0.003 | -.2061 | -.0430282 |
| age1 | .0003806 | .0000919 | 4.141 | 0.000 | .0002004 | .0005608 |
| age2 | -.0012089 | .0002602 | -4.646 | 0.000 | -.001719 | -.0006989 |
| nlc | .9718393 | .1290221 | 7.532 | 0.000 | .7188959 | 1.224783 |
| nlc1 | -.0214653 | .0037004 | -5.801 | 0.000 | -.0287199 | -.0142108 |
| nlc2 | .0551004 | .0103682 | 5.314 | 0.000 | .0347739 | .075427 |
| wives2 | -.156915 | .1233447 | -1.272 | 0.203 | -.3987281 | .0848981 |
| wives3 | -.0023307 | .1414902 | -0.016 | 0.987 | -.2797174 | .275056 |
| catholic | .2965212 | .2981532 | 0.995 | 0.320 | -.2879978 | .8810403 |
| protest | .4141929 | .2952476 | 1.403 | 0.161 | -.1646298 | .9930155 |
| muslim | -.0896059 | .3614763 | -0.248 | 0.804 | -.7982676 | .6190557 |
| other | -.1227712 | .4274395 | -0.287 | 0.774 | -.9607515 | .715209 |
| kalenjin | -.3292647 | .2251427 | -1.462 | 0.144 | -.7706492 | .1121197 |
| kamba | .627231 | .3019089 | 2.078 | 0.038 | .0353491 | 1.219113 |
| kikuyu | .9409074 | .2040773 | 4.611 | 0.000 | .5408209 | 1.340994 |
| kisii | .2356939 | .2740531 | 0.860 | 0.390 | -.3015778 | .7729655 |
| luhya | .0011949 | .2093971 | 0.006 | 0.995 | -.4093209 | .4117107 |
| luo | -.4319167 | .23459 | -1.841 | 0.066 | -.8918223 | .0279888 |
| meru | .5735075 | .3309134 | 1.733 | 0.083 | -.0752367 | 1.222252 |
| miji | .1210063 | .3200831 | 0.378 | 0.705 | -.5065055 | .7485181 |
| urban | -.0356117 | .171943 | -0.207 | 0.836 | -.3727 | .3014767 |
| nairobi | .2369617 | .250119 | 0.947 | 0.343 | -.2533878 | .7273112 |
| central | .1981964 | .2135736 | 0.928 | 0.353 | -.2205072 | .6169001 |
| coast | .238009 | .2839511 | 0.838 | 0.402 | -.3186672 | .7946852 |
| eastern | .7894128 | .3054885 | 2.584 | 0.010 | .1905132 | 1.388312 |
| nyanza | .1410944 | .2294991 | 0.615 | 0.539 | -.3088307 | .5910194 |
| rift | .8504383 | .178055 | 4.776 | 0.000 | .5013677 | 1.199509 |
| manual | .0084505 | .1148461 | 0.074 | 0.941 | -.2167013 | .2336022 |
| saleserv | .3091284 | .101574 | 3.043 | 0.002 | .1099961 | .5082606 |
| proftech | .1698001 | .1142128 | 1.487 | 0.137 | -.0541101 | .3937104 |
| edu | .0216958 | .0819779 | 0.265 | 0.791 | -.139019 | .1824106 |
| edu1 | .002241 | .0048512 | 0.462 | 0.644 | -.0072696 | .0117516 |
| edu2 | -.0036832 | .009537 | -0.386 | 0.699 | -.0223803 | .0150138 |
| house | .2437392 | .1060121 | 2.299 | 0.022 | .0359061 | .4515723 |
| water | .2530343 | .0847937 | 2.984 | 0.003 | .0867993 | .4192694 |
| floor | .1006173 | .1177406 | 0.855 | 0.393 | -.130209 | .3314436 |
| electr | .4169832 | .1715525 | 2.431 | 0.015 | .0806604 | .753306 |
| refrig | .6816207 | .253941 | 2.684 | 0.007 | .1837781 | 1.179463 |
| land | .2719102 | .1404806 | 1.936 | 0.053 | -.0034971 | .5473175 |
| cattle | -.2263562 | .1041131 | -2.174 | 0.030 | -.4304664 | -.022246 |
| crops | -.0088574 | .0888804 | -0.100 | 0.921 | -.1831044 | .1653896 |
| _cons | -2.929137 | .9089653 | -3.222 | 0.001 | -4.711132 | -1.147143 |

Calculation of adjusted proportions using contraception involves using the model to predict the probability of using contraception in each category of exposure to mass media *at constant values of the other variables*. The problem, of course, is what values of the other variables one should use. A common strategy is to set all other variables equal to their means. With linear models this works fine because the mean of the adjusted values equals the observed mean. With logit and other non-linear models, however, the mean of the adjusted values does not equal the observed mean. In the case at hand, following this strategy leads to the adjusted proportions using contraception shown in the column labelled "Method 1" in Table A.5. They show, as one would expect, smaller differences after adjustment. Note, however, that the overall adjusted proportion using contraception, calculated as the weighted average of the proportions in the four categories, is 22 percent, which is substantially *less* than the observed proportion of 27 percent. We should stress that there is nothing wrong with this result —it simply means that if women in the four categories of exposure to the media were otherwise equal to the average, the overall proportion using contraception would be lower than observed— but many analysts find it disturbing.

A second strategy seeks to preserve individual variation. Instead of setting everyone equal to the mean, we let each person retain all their individual characteristics except exposure to the mass media, which is set successively to 0, 1, 2 or 3. In other words, we come up with four predictions for each person, stating what their probability of using contraception would be if they all were in a given exposure category but could keep all their other attributes intact. This method relies on the fact that the likelihood equations in logistic regression (in fact, in any generalized linear model with canonical link) equate observed and predicted values of the minimal sufficient statistics, which include the mean in models with a constant. Thus, we are guaranteed that the average prediction *at the observed* X's equals the average response. In the present case, however, we are predicting using the observed values for all the X's except one, namely exposure to the media. Thus, there is no guarantee that we will reproduce the observed mean. In practice, however, this method usually comes up closer to the observed mean than the previous method. In the case at hand, the method gives the adjusted proportions in the column labelled "Method 2" of table A.5, which average to 26 percent.[1]

The third strategy is the method used in this study. It has the advantages of being exact and much easier to implement than the previous two methods. We start from the observation that under the model, the predicted logit for the $j$-th category of media exposure has the form

$$logit(\pi_j) = \alpha + \hat{\beta}_j$$

where $\alpha$ is a constant depending on the values at which the other $X$'s have been set, and the $\hat{\beta}_j$ are the media

[1] Method 2 doesn't always come this close. In the first draft of Table A.5 we had misaligned the group sizes, putting 0.3034 in category 0 and moving the other proportions down by one. The overall proportion using was then 29 percent. Methods 1-3 gave 21, 26 and 29 percent, respectively.

effects, with $\hat{\beta}_0 = 0$. For example, when we set all other $X$'s to their means, the constant $\alpha$ is $-1.7837$, as shown in Table A.5. The basic idea is to leave implicit the values of the $X$'s and simply solve for the value of $\alpha$ which would lead to predictions which average *exactly* to the observed sample mean. The calculation is easily done using a spreadsheet package such as Quattro Pro, which has a "solve for" option tailored for this precise purpose. In the case at hand the "magic" constant turns out to be -1.5115, as shown in Table A.5. Using this value to make predictions for the four groups results in the adjusted proportions shown in Table A.5 under "Method 3", which average to 27 percent as desired.

One further point worthy of note is that the predictions based on Methods 1 and 3, although different, satisfy exactly the model, in the sense that the difference in the logits of the adjusted proportions for groups 1 and 0, for example, is exactly equal to the parameter estimate of 0.5579. This should not be surprising since the proportions were calculated directly from the parameter estimates. Method 2, however, does some averaging *after* the predictions, and as a result ends up with values which do not satisfy the model exactly. The difference between the logits of the adjusted proportions for media 1 and 0, for example, is 0.4688 instead of 0.5579.

In order to complete the job of reproducing the results in Table 4 in the text we would need to run one further model, namely a logistic regression of contraceptive use on the demographic and socio-economic controls *excluding* the media score. The raw output is shown in Table A.7.

The relevant figure in the output is the log-likelihood, which is -2307.64. As shown in Table A.6, the model with the same demographic and socioeconomic controls *plus* the media scale had a log-likelihood of -2274.39. Twice the difference in log-likelihoods gives the likelihood ratio test for the *net* effect of mass media after controlling for the other variables, which is 66.5 as reported in Table 4. (This number can also be obtained as the difference in the model chi-squared statistics from Tables A.6 and A.7). This completes the work required to obtain two of the lines in Table 4.

So far we have dealt with only one of the two equations in the analysis of contraceptive use. Fortunately, the same procedure can be used for the second equation with only one obvious change: the weight of each media category should be based on the weighted number of *users* rather than the total number of women in that category. The question arises as to whether one should use the *observed* or the *adjusted* number of users (with the adjusted number based, of course, on Method 3). We have used the observed number because we prefer to average both observed and adjusted values using the same weights, but recognize that a case could be made for using weights based on the predicted numbers of users.

Table A.7 Raw output from logistic regression of current contraceptive use on demographic and socioeconomic cohorts, Kenya 1989

| Logit Estimates | | | | | Number of obs = 4778 | |
|---|---|---|---|---|---|---|
| | | | | | chi2(41) = 949.47 | |
| | | | | | Prob > chi2 = 0.0000 | |
| Log Likelihood = -2307.6384 | | | | | Pseudo R2 = 0.1706 | |
| useany | Coefficient | Std. Err. | t | P>\|t\| | [95% Conf. Interval] | |
| age | -.1183375 | .0413198 | -2.864 | 0.004 | -.1993434 | -.0373315 |
| age1 | .0003548 | .0000911 | 3.894 | 0.000 | .0001762 | .0005334 |
| age2 | -.0011276 | .0002577 | -4.375 | 0.000 | -.0016328 | -.0006224 |
| nlc | .9909138 | .1289966 | 7.682 | 0.000 | .7380206 | 1.243807 |
| nlc1 | -.02159 | .0036891 | -5.852 | 0.000 | -.0288223 | -.0143578 |
| nlc2 | .0549401 | .0103261 | 5.320 | 0.000 | .0346961 | .0751842 |
| wives2 | -.1359105 | .1224089 | -1.110 | 0.267 | -.3758889 | .1040679 |
| wives3 | -.0513315 | .1399945 | -0.367 | 0.714 | -.3257857 | .2231228 |
| catholic | .4455228 | .2954499 | 1.508 | 0.132 | -.1336963 | 1.024742 |
| protest | .5748277 | .2925938 | 1.965 | 0.050 | .0012079 | 1.148448 |
| muslim | -.0116279 | .3580936 | -0.032 | 0.974 | -.7136579 | .6904021 |
| other | .0456595 | .423957 | 0.108 | 0.914 | -.7854934 | .8768123 |
| kalenjin | -.3934019 | .222992 | -1.764 | 0.078 | -.83057 | .0437662 |
| kamba | .6428863 | .2983039 | 2.155 | 0.031 | .058072 | 1.227701 |
| kikuyu | 1.012563 | .202596 | 4.998 | 0.000 | .6153804 | 1.409745 |
| kisii | .3088902 | .2719222 | 1.136 | 0.256 | -.2242038 | .8419841 |
| luhya | -.0379531 | .2050803 | -0.185 | 0.853 | -.4400058 | .3640995 |
| luo | -.4760873 | .2327901 | -2.045 | 0.041 | -.9324641 | -.0197106 |
| meru | .5817939 | .3271141 | 1.779 | 0.075 | -.0595018 | 1.22309 |
| miji | .0072395 | .3205433 | 0.023 | 0.982 | -.6211744 | .6356534 |
| urban | -.0291266 | .1718619 | -0.169 | 0.865 | -.3660558 | .3078027 |
| nairobi | .2684436 | .2483468 | 1.081 | 0.280 | -.2184317 | .7553188 |
| central | .2184304 | .2110942 | 1.035 | 0.301 | -.1954124 | .6322732 |
| coast | .3921981 | .2827312 | 1.387 | 0.165 | -.1620865 | .9464827 |
| eastern | .8136472 | .3017653 | 2.696 | 0.007 | .2220469 | 1.405247 |
| nyanza | .2470158 | .2262902 | 1.092 | 0.275 | -.1966182 | .6906498 |
| rift | .9118729 | .1752166 | 5.204 | 0.000 | .5683669 | 1.255379 |
| manual | .0244203 | .1136164 | 0.215 | 0.830 | -.1983207 | .2471613 |
| saleserv | .3603676 | .1004873 | 3.586 | 0.000 | .1633657 | .5573694 |
| proftech | .205673 | .1127285 | 1.824 | 0.068 | -.0153272 | .4266732 |
| edu | .0576601 | .080854 | 0.713 | 0.476 | -.1008513 | .2161715 |
| edu1 | .0013603 | .004789 | 0.284 | 0.776 | -.0080284 | .010749 |
| edu2 | -.0019959 | .0094147 | -0.212 | 0.832 | -.0204531 | .0164612 |
| house | .2702839 | .1050053 | 2.574 | 0.010 | .0644247 | .4761431 |
| water | .2323795 | .0838189 | 2.772 | 0.006 | .0680555 | .3967035 |
| floor | .1963068 | .1155713 | 1.699 | 0.089 | -.0302666 | .4228802 |
| electr | .5217825 | .1678962 | 3.108 | 0.002 | .1926278 | .8509372 |
| refrig | .6816892 | .2506509 | 2.720 | 0.007 | .190297 | 1.173081 |
| land | .2966581 | .1387396 | 2.138 | 0.033 | .024664 | .5686522 |
| cattle | -.1872603 | .1027881 | -1.822 | 0.069 | -.3887726 | .0142521 |
| crops | .0078114 | .0880053 | 0.089 | 0.929 | -.1647198 | .1803426 |
| _cons | -2.971589 | .9020009 | -3.294 | 0.001 | -4.73993 | -1.203248 |