POLICY RESEARCH WORKING PAPER 2447

Attrition in Longitudinal Household Survey Data

Some Tests for Three Developing-Country Samples

Harold Alderman
Jere R. Behrman
Hans-Peter Kohler
John A. Maluccio
Susan Cotts Watkins

Results from this study of the extent and implications of attrition for three longitudinal household surveys from Bolivia, Kenya, and South Africa suggest that multivariate estimates of behavioral relations may not be biased because of high attrition. This suggests that demographers and other social scientists can proceed with collecting longitudinal data to control for unobserved fixed factors and to capture dynamic relationships.

The World Bank
Development Research Group
Rural Development
September 2000

POLICY RESEARCH WORKING PAPER 2447

Summary findings

For capturing dynamic demographic relationships, longitudinal household data can have considerable advantages over more widely used cross-sectional data. But because the collection of longitudinal data may be difficult and expensive, analysts must assess the magnitudes of the problems specific to longitudinal but not to cross-sectional data.

One problem that concerns many analysts is that sample attrition may make the interpretation of estimates problematic. Such attrition may be especially severe where there is considerable migration between rural and urban areas. And attrition is likely to be selective on such characteristics as schooling, so high attrition is likely to bias estimates.

Alderman, Behrman, Kohler, Maluccio, and Watkins consider the extent and implications of attrition for three longitudinal household surveys from Bolivia, Kenya, and South Africa that report very high annual attrition rates between survey rounds.

Their estimates indicate that:

- The means for a number of critical outcome and family background variables differ significantly between those who are lost to follow-up and those who are re-interviewed.
- A number of family background variables are significant predictors of attrition.
- Nevertheless, the coefficient estimates for standard family background variables in regressions and probit equations for the majority of outcome variables in all three data sets are not significantly affected by attrition.

So attrition is apparently *not* a general problem for obtaining consistent estimates of the coefficients of interest for most of these outcomes. These results, which are very similar to those for industrial countries, suggest that multivariate estimates of behavioral relations may not be biased because of attrition. This would support the collection of longitudinal data.

This paper—a product of Rural Development, Development Research Group—is part of a larger effort in the group to evaluate the impact of Bank-funded investments. The study was funded by the Bank's Research Support Budget under the research project "Evaluation of the Impact of Investments in Early Child Development" (RPO 682-34). Copies of this paper are available free from the World Bank, 1818 H Street NW, Washington, DC 20433. Please contact Patricia Sader, room MC3-556, telephone 202-473-3902, fax 202-522-1153, email address psader@worldbank.org. Policy Research Working Papers are also posted on the Web at www.worldbank.org/research/workingpapers. The authors may be contacted at halderman@worldbank.org, jbehrman@econ.sas.upenn.edu, kohler@demogr.mpd.de, or swatkins@pop.upenn.edu. September 2000. (39 pages)

Attrition in Longitudinal Household Survey Data: Some Tests for Three Developing-Country Samples

Harold Alderman, Jere R. Behrman, Hans-Peter Kohler, John A. Maluccio, Susan Cotts Watkins*

*Alderman is in Development Research Group, World Bank, 1818 H Street NW, Washington D.C. 20433, USA, e-mail: halderman@worldbank.org. Behrman is the W.R. Kenan, Jr. Professor of Economics and Director of the Population Studies Center, McNeil 160, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297, USA, e-mail: jbehrman@econ.sas.upenn.edu.Kohler is Head of Research Group, Max-Planck Institute for Demographic Research, Doberaner Str. 114, 18057 Rostock, Germany, e-mail: kohler@demogr.mpg.de.Maluccio is a Research Fellow, International Food Policy Research Institute, 2033 K Street NW, Washington D.C. 20006, USA, email: j.maluccio@cgiar.org.Watkins is Professor of Sociology, McNeil 113, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6299, USA, e-mail: swatkins@pop.upenn.edu.

This paper is part of three projects: (1)"Evaluation of the Impact of Investments in Early Child Development of Nutrition and Cognitive Development " (World Bank), (2) "Social Interactions and Reproductive Health" (National Institutes of Health-Rockefeller Foundation-USAID), and (3) "1998 KwaZulu-Natal Income Dynamics Study" (a collaborative project of researchers from the International Food Policy Research Institute and the Universities of Natal and Wisconsin).

The authors thank Yingmei Cheng and Alex Weinreb for useful research assistance on the Bolivian and Kenyan components of this paper; members of the PAN staff in Bolivia, particularly Elizabeth Peñaranda, for help in understanding the Bolivian data and how PIDI functions; and Michael Carter, Lawrence Haddad, Julian May, and Duncan Thomas for comments at various points in the analysis of the South African component of the project. The findings, interpretations and conclusions expressed in this paper are entirely those of the authors and do not necessarily represent the views of the various agencies that provided resources for this study.

Attrition in Longitudinal Household Survey Data: Some Tests for Three Developing Country Samples

I. Introduction

Longitudinal household data can have considerable advantages over more widely used cross-sectional data for social science analysis. Longitudinal data permit (1) tracing the dynamics of behaviors, (2) identifying the influence of past behaviors on current behaviors, and (3) controlling for unobserved fixed characteristics in the investigation of the effect of time-varying exogenous variables on endogenous behaviors. These advantages are quite relevant for demographers who study processes that occur over time including the impact of programs on subsequent behavior and who often use time-varying exogenous variables. These advantages are also increasingly appreciated: for example, our review of articles published in Demography shows that 26 articles using longitudinal data appeared between 1980-1989, but 65 articles between 1990-2000.

Unfortunately, the collection of longitudinal data is likely to be difficult and expensive, and some, such as Ashenfelter, Deaton, and Solon (1986), question whether the gains are worth the costs. One problem in particular that has concerned analysts is that sample attrition may lead to selective samples and make the interpretation of estimates problematic. Many analysts share the intuition that attrition is likely to be selective on characteristics such as schooling and thus that high attrition is likely to bias estimates made from longitudinal data. While there has been some work on the effect of attrition on estimates using developed-country samples, little has been done using data from developing countries, where considerable migration between rural and urban areas may make any problems of attrition particularly severe. Table 1 summarizes the attrition rates in a number of longitudinal data sets from

developing countries. While these vary considerably (ranging from 6 to 50 percent between two survey rounds and 1.5 to 20.5 percent per year between survey rounds), attrition is often substantial.

In this paper, we consider the implications of attrition for three of the four longitudinal household surveys from developing countries in Table 1 that report the highest per-year attrition rates between survey rounds: (1) a Bolivian household survey designed to evaluate an early childhood development intervention in poor urban areas, with survey rounds in 1995/1996 and 1998; (2) a Kenyan rural household survey designed to investigate the role of social networks in attitudes and behavior regarding reproductive health (e.g. the use of family planning and prevention against HIV/AIDS), with survey rounds in 1994/1995 and 1996/1997; and (3) a South African (KwaZulu-Natal Province) rural and urban household survey designed for more general purposes with survey rounds in 1993 and 1998. The different aims of the projects and the variety of outcomes measures facilitate generalization, at least for survey areas such as these that are relatively poor and experience considerable mobility.

The next section summarizes theoretical aspects of the effects of attrition on estimates, drawing on recent studies on attrition in longitudinal surveys for developed countries. Section 3 describes the three datasets used in this study while section 4 presents some tests for the implications of attrition between the first and the second rounds of the three surveys. Section 5 summarizes our conclusions.

2. Some Theoretical Aspects of the Effects of Attrition on Estimates

Most studies of attrition we have found are for large longitudinal samples in developed countries, several of which appeared in a special issue of *The Journal of Human Resources* (Spring 1998) on "Attrition in Longitudinal Surveys." The striking result of these studies is that the biases in estimated socioeconomic relations due to attrition are small—despite attrition rates as high as 50 percent and with significant differences between the means of a number of outcome and standard control variables for those lost to follow-up and those who were re-interviewed. For example, Fitzgerald, Gottschalk and Moffitt (1998) observe:

> By 1989 the Michigan Panel Study on Income Dynamics (PSID) had experienced approximately 50 percent sample loss from cumulative attrition from its initial 1968 membership.... (p. 251) We find that while the PSID has been highly selective on many important variables of interest, including those ordinarily regarded as outcome variables, attrition bias nevertheless remains quite small in magnitude. The major reasons for this lack of effect are that the magnitudes of the attrition effect, once properly understood, are quite small (most attrition is random).... (p. 252) Although a sample loss as high as [experienced] must necessarily reduce precision of estimation, there is no necessary relationship between the size of the sample loss from attrition and the existence or magnitude of attrition bias. Even a large amount of attrition causes no bias if it is 'random'.... (p. 256)

The other studies in this volume reach similar conclusions. Lillard and Panis (1998, p. 456 on PSID) indicate that, "While we found significant evidence of selective attrition, it appears that this...introduces only very mild biases in substantive results." Van den Berg and Lindeboom (1998, p. 477 on data from the Netherlands) observe that "...the estimates of the covariate effects in the labor market transition rates do not change a lot when allowing for...relations between labor market duration and attrition. In any standard empirical analyses these covariate effects are the parameters of interest." Zabel (1998, p. 502 on SIPP and PSID) conclude that "It appears that accounting for attrition has little impact on the parameter estimates." Ziliak and Kniesner (1998, p. 507 on PSID) also agree that "...nonrandom attrition is of little concern when estimating [labor relations] because the effect of attrition is absorbed into the fixed effects...." And finally, Falaris and Peters (1998, p. 531 on NLS and PSID)

note that "In general...we find that attrition either has no effect on the regression estimates or only affects the estimates of the intercept...."

Fitzgerald, Gottschalk, and Moffitt (1998) provide a statistical framework for the analysis of attrition bias in which the common distinction between selections on variables observed in the data and variables that are unobserved is used to develop tests for attrition bias and correction factors. While neither type of attrition (on unobservables or on observables) necessarily imposes a bias on estimates, the latter may be more amenable to statistical solutions. This leads to a sequence of tests that we will follow in this study. First, given that there is sample attrition, one determines whether or not there is selection on observables. For this purpose, selection on observables includes selection based on endogenous observables such as lagged dependent variables that are observed prior to attrition (e.g. in the first round of the survey). Even if there is selection on observables, this does not necessarily bias the estimates of interest. Thus, one needs to test for possible attrition bias in the estimates of interest as well.

More formally, assume that what is of interest is a conditional population density $f(y|x)$ where y is a scalar dependent variable and x is a scalar independent variable (for illustration, but in practice the extension to making x a vector is straightforward):

(1) $y = \beta_0 + \beta_1 x + \varepsilon$, y observed if $A = 0$

where A is an attrition indicator equal to 1 if an observation is missing its value of y because of attrition, and equal to zero if an observation is not missing its value of y. Since (1) can be estimated only if $A=0$ (that is, one can only determine $g(y|x, A=0)$), one needs additional information or restrictions to infer f(.) from g(.). These can come from the probability of attrition, $PR(A=0|y, x, z)$, where z is an auxiliary variable (or vector) that is assumed to be observable for all units but not included in x. This implies estimates of the form:

(2) $A^* = \delta_0 + \delta_1 x + \delta_2 z + v$

(3) $A = 1$ if $A^* \geq 0$
$= 0$ if $A^* < 0$

Selection on *unobservables* occurs if z is independent of $\varepsilon|x$ but v is not independent of $\varepsilon|x$. Selection on *observables* is the reverse: it occurs if z is not independent of $\varepsilon|x$ but v is independent of $\varepsilon|x$. Stated alternatively, selection on observables occurs if

(4) $\Pr(A=0|y, x, z) = \Pr(A=0|x, z)$

Selection on unobservables occurs if (4) fails to hold, such that the attrition function cannot be reduced from $\Pr(A=0|y, x, z)$.

Selection on unobservables is often presented as dependent on the estimation of the attrition index equation. Identification, however, usually relies on nonlinearities in the index equation or an exclusion restriction, i.e., some z that is not in x. It is difficult to rationalize most such exclusion restrictions because, for example, personal characteristics that affect attrition might also directly affect the outcome variable, i.e., they should be in x. There may be some such identifying variables in the form of variables that are external to individuals and not under their control, such as characteristics of the interviewer in the various rounds (Zabel, 1998). However, in general selection on unobservables presents an obstacle to accurate parameter estimation.[1]

If there is selection on observables, the critical variable is z, a variable that affects attrition propensities and that is also related to the density of y conditional on x. In this sense, z is "endogenous to y." Indeed, a lagged value of y can play the role of z if it is not in the structural relation being estimated but is related to attrition. Two sufficient conditions for the absence of attrition bias due to attrition on observables are either (1) z does not affect A or (2) z is independent of y conditional on x.

---

[1] Fitzgerald, Gottschalk and Moffitt (1998) suggest that indirect tests for selection on unobservables can be made by comparisons with data sets without (or with much less) attrition (e.g., the CPS for the United States), but only very limited possibilities are present for most panels.

Specification tests can be based on either of these two conditions. One test is simply to determine whether candidate variables for z (for example, lagged values of y) significantly affect A. Another test is based on Becketti, Gould, Lillard, and Welch (1988). In the BGLW test, the value of y at the initial wave of the survey (y_0) is regressed on x and on A. The test for attrition is based on the significance of A in that equation. This test is closely related to the test based on regressing A on x and y_0 (which is z in this case); in fact, the two equations are simply inverses of one another (Fitzgerald, Gottschalk, and Moffitt, 1998).

Clearly, if there is no evidence of attrition bias from these specification tests, then one has the desired information on $f(y|x)$. However, Fitzgerald, Gottschalk, and Moffitt (1998) also note that if attrition bias is generated by this type of selection it can be eliminated by the use of weighted least squares (WLS), using weights obtained from estimated equations for the probability of attrition,

$$(5)\ w(z, x) = [(Pr(A=0|z, x))/(Pr(A=0|x)]^{-1}$$

The numerator in relation (5) inside the brackets is the probability of retention in the sample. Because both the weights and the conditional density g are identifiable and estimable functions, the complete population density $f(y|x)$ is estimable, as are its moments such as its expected value. Indeed, Fitzgerald, Gottschalk, and Moffitt (1998) show that a comparison between the WLS and the ordinary least squares (OLS) results provides an additional test for attrition bias.

3. Data and Extent of Attrition

In this section, we describe the three data sets that we use, emphasizing the diverse relations of interest.

3.1 Bolivian Pre-School Program Evaluation Household Survey Data. El Proyecto Integral de Desarrollo Infantil (PIDI) in Bolivia is a targeted urban early child development project expected to

improve the nutritional status and cognitive development of children who participate and to facilitate the labor force participation of their caregivers. PIDI delivers child services through childcare centers located in the homes of local women who have been trained in childcare. The program provides food accounting for 70 percent of the children's nutritional needs, health and nutrition monitoring, and programs to stimulate the children's social and intellectual development. The PIDI program was designed to facilitate ongoing impact evaluation through the collection of panel data.

Eligibility for PIDI at the time of the collection of the first and second rounds of data was based on an assessment of social risk. As a result of this selection, children who attend a PIDI center are, on average, from poorer family backgrounds than children who live in the same communities but who do not attend a PIDI center (see Todd, Behrman, and Cheng, 2000). The first PIDI evaluation data set (Bolivia 1) was collected between November 1995 and May 1996 and consisted of 2,047 households.[2] The follow-up survey (Bolivia 2) was collected in the first half of 1998 and consisted of interviews in the 65 percent of the original 2,047 households that could be located (plus an additional 3,453 households that were not visited in Bolivia 1). The attrition rate of 35 percent for Bolivia 1 is relatively high, which raised concern about whether reliable inferences could be drawn from analysis of Bolivia 2.

3.2 The Kenyan Ideational Change Survey (KDICP). KDICP is a longitudinal survey designed to collect information for the analysis of the roles of informal networks in understanding change in knowledge and behavior related to contraceptive use and AIDS. Four rural sites (sublocations) were

[2] These households were stratified into three subsamples: (P) (40 percent of the total), which is a stratified random sample of households with children attending PIDI in which first the PIDI sites were selected randomly and then children within the sites were selected randomly. (A) (40 percent of the total), which is a stratified random sample (based on the 1992 census) of households with children in the age range served by PIDI living in poor urban communities comparable to those in which PIDI had been established, but in which PIDI programs had not been established as of that time. (B) (20 percent of the total), which is a stratified random sample (based on the 1992 census) of households with at least one child in each household in the age range served by PIDI and living in poor urban communities in which PIDI had been established and within a three block radius of a PIDI but without children attending PIDI.

chosen in Nyanza Province, near Lake Victoria in the southwestern part of Kenya. The sites were chosen to be similar in most respects but to maximize variation on two dimensions: 1) the extent to which social networks were confined to the sublocation versus being geographically extended and 2) the presence or absence of a community-based distribution program aimed at increasing the use of family planning. Villages were selected randomly within each site and interviews were attempted with all ever-married women of childbearing age (15–49) and their husbands. The study consisted of ethnographic interviews, focus groups, and a household survey of approximately 900 women of reproductive age and their husbands that was conducted December 1994–January 1995 (Kenya 1). A second round was conducted in 1996/1997 (Kenya 2). (The surveys are described in detail at www.pop.upenn.edu/networks). The attrition rates between the two surveys were 33 percent for men, 28 percent for women, and 41 percent for couples (Table 1).[3] These rates are comparable to the 35 percent reported for the Bolivian data.

[3] There also is "reverse attrition" in the sense of respondents who were present in Kenya 2 but not in Kenya 1: 12 percent (of the Kenya 2 total) for men, 11 percent for women, and 19 percent for couples.

Table 2 summarizes data on the reported causes of attrition for men and women as obtained generally from other household members for most individuals who were interviewed in Kenya1 but not in Kenya 2.[4] Nyanza Province has a relatively high level of AIDS: Mortality between the surveys accounted for 18.4 percent of the reasons given for men's attrition, but only half as much (9.9 percent) for women . For both men and women the leading explanation was migration, accounting for 58.6 percent of the reasons given for women and 47.8 percent of the reasons given for men. Because this is a patrilocal society, a significant share of this migration (over one-third) for women was associated with divorce or separation, but this was not a major factor for men. Not being found at home after at least three visits by interviewers was the next most common explanation for attrition in Kenya 2, accounting for about one-sixth of the reasons given for both men (17.9 percent) and women (15.8 percent). Explicitly refusing or claiming to be too busy or sick to participate accounted for slightly smaller percentages—15.9 percent for men and 11.4 percent for women (with most of this gender difference accounted by "other," which is 4.4 percent for women but 0.0 percent for men).

3.3 KwaZulu-Natal Income Dynamics Study (KIDS). The first South African national household survey, the 1993 Project for Statistics on Living Standards and Development (PSLSD), was undertaken in the last half of 1993 under the leadership of the South African Labour and Development Research Unit (SALDRU) at the University of Cape Town.[5] Unlike the special purpose household surveys for Bolivia and Kenya described above, the South African survey was a comprehensive household survey similar to a Living Standards Measurement Survey or "LSMS" (Grosh and Muñoz 1996; Deaton 1997; Grosh and Glewwe 2000) and collected a broad array of socioeconomic information from individuals and

[4] These data are not available for 22.4 percent of the men and 21.8 percent of the women interviewed in Kenya 1 but not in Kenya 2.

[5] PSLSD is alternatively referred to as the SALDRU survey, the South African Integrated Household Survey (SAIHS), and the South African Living Standards Measurement Survey (LSMS).

households. Among other things, it included sections on household demographics, household environment, education, food and nonfood expenditures, remittances, employment and income, agricultural activities, health, and anthropometry (weights and heights of children aged six and under). The 1993 sample was selected using a two-stage, self-weighting design. In the first stage, clusters were chosen proportional to population size from census enumerator districts or approximate equivalents when these were unavailable. In the second stage, all households in each chosen cluster were enumerated and then a random sample selected (see PSLSD 1994 for further details).

Since the 1993 survey, South Africa has undergone dramatic political, social, and economic change, beginning with the change of government after the first national democratic elections in 1994. With the aim of addressing a variety of policy research questions concerning how individuals and households were coping during this transition, households surveyed by the PSLSD in South Africa's most populous province, KwaZulu-Natal, were resurveyed from March to June, 1998, for the KIDS (see May, et al., 2000). In this paper, the sample of 1993 PSLSD households in KwaZulu-Natal is referred to as South Africa 1 and those re-interviewed in 1998 for KIDS, South Africa 2.

An important aspect of the South Africa resurvey—differentiating it further from the Bolivian and Kenyan longitudinal surveys—is that, when possible, the interviewer teams tracked, followed, and re-interviewed households that had moved.[6] Hence, in the South Africa survey migration does not imply automatic attrition from the sample. In addition to reducing the level of attrition and allowing analysis of migration behavior, tracking and following plausibly reduced biases introduced by attrition, a claim that is evaluated below.

In 1993, the KwaZulu-Natal sample contained 1,393 households (215 Indian and 1,178 African).

---

[6] In practice certain key individuals in the household were pre-designated for tracking if they had moved; in some cases this led to split households in 1998, but that does not affect this analysis which, except for the attrition indicator, uses only 1993 data (May et al., 2000).

Of the target sample, 1,171 households (84 percent) with at least one 1993 member were successfully re-interviewed in 1998 (Maluccio 2000). There were four one- and two-person households whose members had all died over the period. As in most surveys in developing countries, refusal rates are low: only nine re-contacted households refused an interview. The remaining households that could not be followed-up were either verified as having moved but could not be tracked (81 or 5.8 percent) or left no trace (128 or 9.2 percent). Had 63 movers not been followed, only 79 percent of the target households would have been re-interviewed. Put another way, the tracking procedures yielded a 25 percent reduction in the number of households that were lost to follow-up.

Re-interview rates were slightly higher in urban than in rural areas, reflecting the 89 percent success rate in re-contacting urban Africans (294 households). Offsetting that success was a follow-up rate of 78 percent (215 households) for Indian households, all of which were urban. The follow-up rate for rural Africans was 84 percent (884 households), reflecting the rate for the overall sample. There were no major differences between the rural and urban samples, and we therefore pooled them in the analysis below.

The discussion of attrition between South Africa 1 and South Africa 2 to this point has focused on attrition at the household level. For an analysis of individual level outcomes, however, measuring attrition at the individual level is more appropriate. Because a household was considered to be found if at least one 1993 member was re-interviewed, individual-level attrition for the entire sample is necessarily higher than household attrition (although this need not be the case for subsamples of individuals). Focusing on the sample of children aged 6–72 months for whom there is complete information on height,

weight, and age in 1993 (N=916), for example, 78 percent were re-interviewed as resident or nonresident household members in 1998, indicating one-third more attrition than at the household level.[7]

4. Some Attrition Tests for the Bolivian, Kenyan, and South African Samples

As noted, the attrition rates for the three samples considered here are considerable—35 percent for the Bolivian sample, from 28 percent for women to 41 percent for couples in the Kenyan sample, and from 16 percent for households to 22 percent for preschool children in the South African sample. However, studies for developed countries suggest that while attrition of this magnitude may be selective, it need not significantly affect estimated multivariate relations. To test this, we conducted three sets of tests of attrition as it relates to observed variables in the data, using some of the tests presented by Fitzgerald, Gottschalk, and Moffitt (1998). We begin with a comparison of means, since the intuition that attrition is likely to bias estimates is often made on the basis of such univariate comparisons. We then estimate probits for the probability of attrition in order to ask what variables predict attrition and compare univariate and multivariate estimates. Lastly, we test whether coefficient estimates differ for the two subsamples, one that is lost to follow-up and one that is re-interviewed.

4.1 Comparison of Means for Major Outcome and Control Variables. First, we compared means for major outcome and control variables measured in the first rounds of the respective data sets for those

[7] There are 1,029 African and Indian children in KwaZulu-Natal in 1993 with complete height, weight, and age information but the following are dropped from the analysis: 26 because the absolute value of at least one of the three height-for-weight z scores, weight-for-age z scores, or weight-for-height z scores exceeded 9.9; 47 who were less than 6 months old; and 30 who were more than 72 months old. If only those re-interviewed as residents (living in the household more than 15 out of the past 30 days) are considered, attrition rises to 31 percent, but the

subsequently lost to follow-up versus those who were re-interviewed (Table 3). Major variables are defined with respect to the interests of the project for which the data were collected.

Bolivia: A number of means for those lost to follow-up differ statistically from those for who eventually were re-interviewed: rates of severe stunting, moderate wasting, the fraction reporting that they mainly spoke Quechua at home, weight-for-age, gross motor ability test scores, fine motor ability test scores, language-audition test scores, personal-social test scores, mother's age, father's age, home ownership, fraction with both parents present, number of rooms in the home, number of siblings, ownership of durables, mother having job, and household income. All of these observable characteristics distinguish the two subsamples at least at the 10 percent significance level, and show that in the first round of the data (Bolivia 1) children who were worse off in terms of these measures were more likely to be lost to follow-up before the second round than those who would eventually be re-interviewed. Among the fourteen predetermined parental and household level variables in Table 3, eleven differ significantly for the two groups at least at the 10 percent significance level. Thus, both in terms of child development outcome variables and family background variables, attrition seems to be systematically more likely for children who are worse off. Such systematic differences, together with the high attrition rates, may cause concern about what can be inferred with confidence from these longitudinal data.

Kenya: For the Kenyan data, both males and females lost to follow-up have higher schooling, more languages, and are more likely to have heard radio messages about contraception and lived in households with males who received salaries. They are also younger and have fewer children than those who were re-interviewed. For a few variables the means differ significantly between these two subsamples for men but not for women (ever-use of contraceptives, residence in the sublocation of

---

results reported on here are qualitatively the same.

Owich) or for women but not for men (want no more children, visited by community-based distribution agent, speaks Luo only, belongs to credit group or to clan welfare society, residence in the sublocation of Wakula South). On the other hand, the means do not differ for the subsamples for either men or women for a number of characteristics (currently using contraceptives, heard about family planning at clinic, discussed family planning with others, number of partners in networks, primary schooling, lived outside of province, polygamous household).

Therefore, it appears that attrition is selective in terms of some "modern" characteristics (including some of the outcome variables that these data were designed to analyze) with selectivity related more to women's characteristics. But the means for many characteristics—including those for most of the indicators of social interaction, the impact of which is central to the project for which these data were gathered—do not differ significantly between those lost to follow-up and those re-interviewed.

South Africa: Because the South African survey is a comprehensive household survey with a large number of variables, for comparability this attrition study examined a set of variables similar to those considered for Bolivia, i.e., child nutritional status as measured by anthropometrics and a health indicator, whether the child was sick in the last two weeks, as well as a set of predetermined family background characteristics.

There are no significant differences in child nutritional status and health outcome variables between the two groups. This is not the case for the predetermined family background variables, however, where there are a number of significant differences in means. Those who were re-interviewed are significantly more likely to be African rather than Indian, have lower income, lower expenditures, less educated household heads, and fewer durable assets. Of course, since these background variables themselves tend to be highly correlated (in particular race with education, income, and assets), it is not surprising that they show similar patterns in the comparisons of means. In sum, while there are no

apparent differences in the child outcome variables, children from better off or Indian households were more likely to be lost to follow-up.

4.2 Probits for Probability of Attrition. We start with a parsimonious specification of probits for the probability of attrition in which only one outcome variable at a time is included; we then include all outcome variables plus predetermined family background variables (Table 4). The dependent variable in these probits is whether attrition occurred between the survey rounds (1=yes; 0=no). Chi^2 tests presented at the bottom of the table test the significance of the overall relations.

Bolivia: The Chi^2 tests indicate that if only one of the outcome variables at a time is included in these probits, the probit is significant at the 5 percent level only for severe stunting—that is, a child who is severely stunted is more likely to be lost to follow-up. For moderate and severe low weight-for-age and the four test scores, the probits are significant at the 10 percent level, suggesting that poor childhood development increases the probability of attrition. When all of the family background variables and all childhood development indicators are included in the analysis, however, among the childhood development indicators only moderate stunting is significantly nonzero, even at the 10 percent level, with a negative sign. That 1 in 11 of the childhood development indicators has a significant coefficient estimate at the 10 percent level in the multivariate analysis is what one would expect to occur by chance, even if none of the childhood development indicator coefficients were truly significant predictors of attrition. Moreover, the one childhood development outcome variable that has a significantly nonzero coefficient estimate in Table 4 in the multivariate analysis does *not* show significant differences in the comparison of means in Table 3.

The comparisons of means for childhood development outcomes between subsamples of those lost to follow-up and those who were re-interviewed , therefore, may be quite misleading regarding the extent of significant associations of these childhood development indicators with sample attrition once

family background characteristics are controlled. The comparisons in Table 3 indicate that there is selective attrition with regard to childhood development indicators, with those children who are worse off in round 1 significantly more likely to be lost to follow-up. But the multivariate estimates present a different picture: they indicate that the extent of significant associations for the child development outcomes in probits for predicting attrition is about what would be expected by chance. Thus, conditional on controls for observed family background characteristics, attrition is *not* predicted by child development indicators for round 1. (Of course, there may be multicollinearity among the child development indicators that disguises their significance.)

If the predetermined family background variables in Bolivia 1 are included alone or with all of the early childhood development indicators, the probits are significantly nonzero at very high levels. Some family background variables are significantly (at least at the 10 percent level) associated with higher probability of attrition: older and less-schooled fathers, speaking mainly Quechua in the household, not owning the home, having more rooms in the house, having fewer siblings, having fewer durables, father having permanent or no (rather than a temporary) job, and mother having no or a temporary (rather than a permanent) job, with some significant differences also among the urban areas included in the program. The majority of these significant coefficient estimates are consistent with what might be predicted from the significant differences in the means in Table 3, reinforcing the observation that attrition tends to be selectively greater among children from worse-off family backgrounds.

But some of these significant coefficient estimates are opposite in sign from what might be expected from the comparisons of the means in Table 3, suggesting the opposite relation to attrition if there are multivariate controls for standard background variables other than what appear in the comparisons of means. Specifically, the comparisons in Table 3 suggest that attrition is significantly more likely if fathers are younger, the house has fewer rooms, and there are fewer siblings—but all three

of these signs are reversed with significant coefficient estimates in the multivariate analyses of Table 4. Moreover, two variables that are not significantly different for the two subsamples in Table 3 have significant coefficient estimates in Table 4, i.e., father's schooling and father having a temporary job, both of which are estimated to significantly reduce attrition probabilities in Table 4. Finally, both mother's age and household income have means that are significantly different between the subsamples in the univariate comparisons in Table 3, but do not have coefficient estimates that are significantly nonzero, even at the 10 percent level, once there is control for other family background characteristics in Table 4.

Thus, exactly which family background characteristics predict attrition with multivariate controls and what the directions of those effects are cannot be inferred simply by examining the significance of means in univariate comparisons between the subsamples. While the patterns in Tables 3 and 4 suggest that worse-off family background is associated with greater attrition, the multivariate estimates are less supportive of this conclusion.

Kenya: Since there are gender differences in the probit estimates of the probability of attrition, we report separately for men and women. For men, we find that when the five outcomes are included singly, only the number of surviving children is significantly related to attrition at the 5 percent level; one other--ever-used family planning--is significantly related to attrition at the 10 percent level. If other variables are included as right-side variables, among the five fertility related outcomes none is significantly nonzero at the 5 percent level, and only not wanting more children is significantly related to attrition at the 10 percent level. A Chi^2 test for the joint significance of these five variables rejects such significance ($p=0.52$). Among the control variables only age is significant, but not schooling, language, household characteristics, past residence in Nairobi or Mombasa, or current sublocation of residence. A

Chi^2 test for the joint significance of all the right-side variables rejects such significance at the 5 percent level ($p=0.068$).

For women, we find that two of the lagged outcome variables, wanting no more children and the number of surviving children, are individually significant (and negative). When all the lagged outcome variables and the predetermined variables are included, only the latter (number of surviving children) remains significant. However, in contrast to the results for men, Chi^2 tests for the joint significance of the five fertility related outcome variables and for the entire set of right-side variables indicate significance ($p=0.0000$ in both cases).

Thus, for the Kenyan data, there is no significant association between attrition, most of the outcome variables, and most of the major control variables. However, gender does matter in these multivariate analyses : there is a significant negative association between attrition and number of surviving children for women but not for men.

South Africa: Probit estimates for the probability of attrition reveal little evidence that the outcome variables are associated with attrition, paralleling the results of the mean comparisons in Section 4.1. When only one outcome variable at a time is included, none is significant at conventional levels. When all are included in at once, the outcome variables are both individually and jointly insignificant.

The conditional influence of the predetermined variables differs from the mean comparisons but confirms that some of them are significant predictors of attrition even though the overall relation is insignificant. Children in households with older heads and more assets (number of rooms and durables are jointly significant) are more likely to have been lost to follow-up. Conditional on these assets, however, household ownership made it less likely that there was attrition, probably due to homeowners having deeper roots or higher moving costs. After controlling for these factors, race is no longer

associated with attrition.

4.3 Do Those Lost to Follow-up have Different Coefficient Estimates than Those Re-interviewed? Our aim here is to determine whether those who subsequently leave the sample differ in their initial behavioral relationships. We conduct the BGLW tests, in which the value of an outcome variable at the initial wave of the survey is regressed on predetermined variables for the initial survey wave and on subsequent attrition. In short, the test is whether the coefficients of the predetermined variables and the constant differ for those respondents who are subsequently lost to follow- up versus those who are re-interviewed. Table 5 presents these multivariate regression and probit estimates for the same outcome variables considered above, with the same family background variables among the right-side variables. The first part of the table gives the coefficient estimates for the family background variables for the subsample of those who were re-interviewed. At the bottom of the table are the F or Chi^2 tests for whether there are significant differences between the two subsamples and tests for (i) all of the slope coefficients and constant and (ii) all of the slope coefficients (but not the constant).

Bolivia: F tests indicate that all of the eleven estimated equations for childhood development indicators are statistically significant at the 0.00 percent level. These estimates indicate a number of associations that are consistent with widely held perceptions about child development. For example, household income is significantly positively associated with height-for-age and significantly negatively associated with severe stunting; mother's schooling is significantly positively associated with height-for-age and weight-for-age, though significantly negatively associated with gross motor ability; and ownership of consumer durables is significantly positively associated with height-for-age, gross motor ability, fine motor ability, language-audition, and personal-social test scores, but significantly negatively associated with severe wasting.

There are, however, no significant differences at the 5 percent level[8] between the set of coefficients for the subsample of those lost to follow-up versus the subsample of those re-interviewed for over half of the indicators of child development: height-for-age, moderate stunting, gross motor ability tests, fine motor ability tests, language-audition tests, and personal-social tests. The second set of tests, further, indicates that there are no significant differences at the 10 percent level for severe stunting. These estimates for the anthropometric indicators related to stunting and for the four cognitive development test scores, therefore, suggest that the coefficient estimates of standard family background variables are *not* significantly affected by sample attrition.

The results differ sharply, however, for the anthropometric indicators related to wasting. Both tests for these four child outcome variables indicate that the coefficient estimates for observed family background variables do differ significantly at the 5 percent level (and for all but weight-for-age at the 1 percent level) between the two subsamples. For these outcomes, therefore, it is important to control for the attrition in the analysis, e.g., as with the matching methods used in Todd, Behrman, and Cheng 2000.

Kenya: We conduct BGLW tests with Kenya 1 contraceptive use (ever or currently), want no more children, number of surviving children, and family planning network size as the dependent variables. The right-side variables again include a fairly standard set of control variables, i.e., age, schooling, wealth indicators, language indicators, and location of residence. Tests for the significance of the differences in the slope coefficients in all cases for both men and women fail to reject equality of all the coefficients between the subsamples of those lost to follow-up and those re-interviewed. Tests for the joint significance of the differences in the slope coefficients and intercepts in all cases fail to reject equality of all the coefficients and of an additive variable for attrition (with the exception at the 5 percent

[8] This is true at the 10 percent level as well for all of these except for the fine motor ability test score.

level of number of surviving children and at the 10 percent level for currently using contraceptives, both only for women and in both of which cases the constant differs between the subsamples, but not the slope coefficient estimates).

Thus there is no significant effect on the slope coefficients of attrition for either men or women, and but limited evidence of a significant effect on the constants for women.

South Africa: The evidence for South Africa presented earlier in Sections 4.1 and 4.2 suggests that the amount of attrition bias resulting from selection on observables is not significant. The BGLW tests largely confirm this, although there are some exceptions.For the first three anthropometric outcomes, the attrition interactions are not jointly significant although in the case of height-for-age the joint test on all interacted coefficients approaches significance at the 10 percent level ($p=0.104$) when the constant is not considered. The overall fit for the stunting and wasting probits is much better than for the regressions in the first three columns: all four relationships are significant at the 5 percent level. The attrition interaction terms are significant only in the case of moderate stunting, indicating the possibility of attrition bias in this relationship. On the other hand, attrition does not appear to have any association with severe stunting or moderate and severe wasting. If the child was sick in the last two weeks (the last column) the results for the full set of interactions suggest attrition bias is present.

As described in Section 3, one important difference in the South African sample relative to the others is that, when possible, households that had moved were followed. These households are included in the analysis presented above. What would happen if they were excluded? Re-estimating the equations in Table 5 categorizing those who had moved but were interviewed as if they had been lost to follow-up leads to a somewhat stronger, but still fairly weak, rejection of the null hypothesis that there are no

differences in coefficients across the two groups (results not shown). In every case the p value for both F-tests declines; for height-for-age and severe stunting this decline is enough for the tests to become significant at the 10 percent level. It appears that the investment made in following movers had some payoff in terms of reduced attrition bias for this set of relationships, though these alternative estimates still do not indicate very high probabilities of attrition bias.

5. Conclusions

Our conclusions are similar in some respects to those of Fitzgerald, Gottschalk, and Moffitt (1998) for the Panel Study of Income Dynamics in the United States that is summarized in Section 2 but differ in other respects:

(a) The means for a number of critical child development outcome and family background variables do differ significantly between the subsample of those lost to follow-up between two rounds of a survey and those who were re-interviewed. For the Bolivian PIDI data, there is a definite tendency for those lost to follow-up to have poorer child development outcomes and family background than those who were re-interviewed. In the poor urban communities on which PIDI concentrates, it appears that worst-off households are most mobile and thus most difficult to follow over time. This is similar to the U.S. results. It contrasts, however, with the Kenyan rural data and the South African rural and urban data, where households and individuals with better backgrounds, (e.g., more schooling, more likely to speak English), are most mobile and thus hardest to follow over time. For the Kenyan data, this may be the case because better-off individuals tend to migrate from the poor rural sample areas to urban areas. For the South African data, however, this result is for both rural and urban areas, so it does not reflect selective migration from rural to urban areas by those who are better off.

(b) Neither family background variables nor outcome variables measured in the first of two

surveys reliably predict attrition in multivariate probits. Some of the Bolivia 1 family background variables, but not the Bolivia 1 child outcome variables, are significant predictors of attrition. The result for the child outcome variables is similar to that for the outcome variables in the Kenyan case. But the significance of a number of background variables in predicting attrition in the Bolivian data, while similar to the U.S. results, again contrasts with the limited significance of such background variables in predicting attrition in the Kenyan and South African data. There are some gender differences in the Kenyan data, with attrition for women being more associated with their observed characteristics than is attrition for men. For South Africa, the overall probit relation does not significantly predict attrition, even though some individual variables appear to predict greater attrition of children—older household heads, more nonhousing assets, and lack of home ownership.

(c) *Attrition does not generally significantly affect the estimates of the association between family background variables and outcome variables.* The coefficient estimates for standard family background variables in regressions and probit equations for the majority of the Bolivian child development outcome variables—including all of those related to stunting and to the test scores for gross and fine motor ability, language/auditory and personal/social interactions—are not affected significantly by attrition. The coefficients on standard variables in equations with the major outcome and family planning social network variables in the Kenyan data also are unaffected by attrition and—in contrast to the Fitzgerald, Gottschalk, and Moffitt (1998) study—the constants also do not differ (with the possible exceptions of number of surviving children and of currently using contraceptives for which cases the constants differ at the 10 percent level for women). For six of the seven child anthropometric measures in the South African data, moreover, there are no significant effects of attrition on the coefficient estimates of the standard variables nor, again, of the constants. Therefore, attrition apparently is *not* a general problem for obtaining consistent estimates of the coefficients of interest for most of the child development

outcomes in the Bolivian data, for the fertility/social network outcomes in the Kenyan data, and for some of the anthropometric indicators in the South African data. These results are very similar to the results for the outcome measures for similar analyses with longitudinal U.S. data and suggest that despite suggestions of systematic attrition from univariate comparisons between those lost to follow-up and those re-interviewed, multivariate estimates of behavioral relations of interest may not be biased due to attrition.

It should be noted that for some outcomes the results differ strikingly and suggest that attrition bias will sometimes be a problem in multivariate estimates of behavioral relations that do not control for attrition. Among the particular outcomes that we consider in all three samples, there are significant interactions of attrition with the sets of standard variables that we consider in 6 out of 29, or 21 percent, of the cases, higher than the 5 percent that would be expected by chance at the 5 percent significance level. Attrition selection bias appears to be model specific: changing outcome variables may change the diagnosis even within the same data set. Thus, as a general observation, analysts should assess the problem for the particular model and the particular data they are using.

Nevertheless, the basic point remains: in contrast to often-expressed concerns about attrition, for many estimates the coefficients on standard variables in equations are unaffected by attrition. This is the case for longitudinal samples for developed countries, and we have shown it to be the case for longitudinal samples in developing countries as well, using a wide variety of outcome variables. Thus, even when attrition is fairly high, as it is in the samples we used, attrition apparently is *not* a general and pervasive problem for obtaining consistent estimates. This suggests that demographers, as well as other social scientists, proceed with greater confidence in their growing attempts to use longitudinal data to control for unobserved fixed factors and to capture dynamic relationships.

References

Alderman, Harold and Jere R. Behrman, 1999, "Attrition in the Bolivian Early Childhood Development Project and Some Tests of the Implications of Attrition," Philadelphia: University of Pennsylvania, mimeo.

Ashenfelter, Orley, Angus Deaton, and Gary Solon, 1986, "Collecting Panel Data in Developing Countries: Does it Make Sense?" LSMS Working Paper 23, Washington, DC: The World Bank.

Becketti, Sean, William Gould, Lee Lillard, and Finis Welch, 1988, "The Panel Study of Income Dynamics after Fourteen Years: An Evaluation," *Journal of Labor Economics* 6: 472-92.

Behrman, Jere R., Hans-Peter Kohler, and Susan Cotts Watkins, 1999, "Attrition in Longitudinal Socioeconomic Data from Sub-Saharan Africa: The Kenyan Ideational Change and Diffusion Data," Philadelphia: University of Pennsylvania, mimeo.

van den Berg, Gerard J. and Maarten Lindeboom, 1998, "Attrition in Panel Survey Data and the Estimation of Multi-State Labor Market Models," *The Journal of Human Resources* 33(2): 458-478.

Deaton, Angus, 1997, *The Analysis of Household Surveys*, London: The Johns Hopkins University Press.

Evans-Pritchard, E.E., 1949, "Luo Tribes and Clans," *Journal of the Rhodes-Livingston Institute* 24 (40): 125–8.

Falaris, Evangelos M. and H. Elizabeth Peters, 1998, "Survey Attrition and Schooling Choices," *The Journal of Human Resources* 33 (2): 531–554.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt, 1998, "An Analysis of Sample Attrition in Panel Data," *The Journal of Human Resources* 33 (2): 251–99.

Foster, Andrew and Mark R. Rosenzweig, 1995, "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy* 103 (6): 1176–1209.

Grosh, Margaret and Paul Glewwe, eds., 2000, *Designing Household Survey Questionnaires for Developing Countries: Lessons from Ten Years of LSMS Experience*, Oxford, UK: Oxford University Press for the World Bank

Grosh, Margaret E. and Juan Muñoz, 1996, "A manual for Planning and Implementing the Living Standards Measurement Study Survey," Living Standards Measurement Study Working Paper No. 126, Washington DC: The World Bank.

Lillard, Lee A. and Constantijn W.A. Panis, 1998, "Panel Attrition from the Panel Study of Income Dynamics," *The Journal of Human Resources* 33 (2): 437–57.

Maluccio, John A., 2000, "Attrition in the KwaZulu-Natal Income Dynamics Study, 1993–1998," Food Consumption and Nutrition Division Discussion Paper (forthcoming), Washington, DC: IFPRI.

May, Julian, Michael R. Carter, Lawrence Haddad, and John A. Maluccio, 2000, "KwaZulu-Natal Income Dynamics Study 1993–1998: A Longitudinal Household Database for South African Policy Analysis," *Development Southern Africa* (forthcoming).

PSLSD, 1994, Project for Statistics on Living Standards and Development: South Africans Rich and Poor: Baseline Household Statistics, South African Labour and Development Research Unit, University of Cape Town, South Africa.

Renne, Elisha P., 1997, "Considering Questionnaire Responses: An Analysis of Survey Interactions," Paper presented at the annual meeting of the African Studies Association, Columbus, Ohio, 13–16 November 1997.

Smith, James P. and Duncan Thomas, 1997, "Migration in Retrospect: Remembrances of Things Past," Santa Monica, CA: Rand Labor and Population Program, Working Paper Series 97-06.

Thomas, Duncan, Elizabeth Frankenberg, and James P. Smith, 1999, "Lost But Not Forgotten: Attrition in the Indonesian Family Life Survey," RAND Labor and Population Program Working Paper Series 99-01, Santa Monica, CA: RAND.

Todd, Petra, Jere R. Behrman, and Yingmei Cheng, 2000, "Evaluating Preschool Programs when Length of Exposure to the Program Varies: A Nonparametric Approach," Philadelphia: University of Pennsylvania, mimeo.

Zabel, Jeffrey E., 1998, "An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Behavior," *The Journal of Human Resources* 33 (2): 479–506.

Ziliak, James P. and Thomas J. Kniesner, 1998, "The Importance of Sample Attrition in Life Cycle Labor Supply Estimation," *The Journal of Human Resources* 33 (2): 507–3

Table 1. Attrition Rates for Longitudinal Household Survey Data in Developing Countries Listed in Order of Attrition Rates Per Year			
Country, Time Period/Interval Between Rounds (in rough order of attrition rates per year)	**Attrition Rate between Rounds (%)**	**Attrition Rate per Year (%)**	**Source**
Bolivia (urban), 1995/6to 1998 (two-year interval)	35	17.5	Present study (also see Alderman and Behrman 1999)
Kenya (rural, South Nyanza Province), 1994/5 to 1996/7 (two year interval) – couples – men – women	41 33 28	20.5 16.5 14.0	Present study (also see Behrman, Kohler, and Watkins 1999)
Nigeria (five year interval)	50	10.0	Renne (1997)
South Africa (KwaZulu-Natal) 1993 to 1998 (five year interval) –households –pre-school children	16 22	3.2 4.4	Present study (also see Maluccio 2000)
India (rural) 1970/71 to 1981/2 (11-year interval)	33	3.0	Foster and Rosenzweig 1995
Malaysia (12 year interval)	25	2.1	Smith and Thomas 1997
Indonesia 1993 to 1997 (four-year interval)	6	1.5	Thomas, Frankenberg, and Smith 1999

Table 2. Reported Reasons for Men's and Women's Attrition in Kenyan (KICS) Survey				
	Men		**Women**	
Reason for attrition:	*N*	*%*	*N*	*%*
i. Working, moved to, or visiting outside Nyanza Province	45	*22.4*	21	*10.3*
ii. Working, moved to, or visiting elsewhere in Nyanza Province				
iii. Not home	51	*25.4*	56	*27.6*
iv. Refused				
v. Sick or busy	36	*17.9*	32	*15.8*
vi. Deceased	26	*12.9*	20	*9.9*
vii. Separated, divorced, then moved away	6	*3.0*	3	*1.5*
viii. Other	37	*18.4*	20	*9.9*
	n/a	*n/a*	42	*20.7*
	0	*0.0*	11	*4.4*
	201		**205**	

Table 3. Bolivia. T-tests for Differences in Means in Bolivia 1 Data for Attritors versus Nonattritors [a]

Variables	Means for Nonattritors (Standard Deviation)	Means for Attritors (Standard Deviation)	Difference in Means (T test)
Early Child Development Outcome Variables			
Height-for-age	18.0 (22.5)	17.4 (22.1)	0.65 (0.72)
Weight-for-age	32.2 (26.5)	30.3 (25.8)	1.91** (1.81)
Weight-for-Height	58.1 (26.5)	56.9 (27.2)	1.21 (1.10)
Moderate Stunting	0.639 (0.480)	0.631 (0.483)	0.008 (0.43)
Severe Stunting	0.279 (0.449)	0.323 (0.468)	-0.0437* (-2.37)
Moderate Wasting	0.365 (0.482)	0.400 (0.490)	-0.035** (-1.79)
Severe Wasting	0.0796 (0.271)	0.0946 (0.293)	-0.0150 (-1.30)
Gross Motor Ability	20.8 (7.81)	20.3 (7.67)	0.5136** (1.65)
Fine Motor Ability	19.4 (7.28)	19.0 (7.19)	0.480** (1.65)
Language-Audition	19.2 (7.62)	18.6 (7.44)	0.569** (1.88)
Personal-Social	19.9 (8.02)	19.4 (8.06)	0.534** (1.65)
Predetermined Family Background Variables			
Mother's age	29.8 (6.45)	28.7 (6.44)	1.07* (4.10)
Father's age	33.0 (7.70)	32.2 (8.03)	0.85* (2.66)
Mother's schooling	3.0 (1.5)	3.0 (1.5)	-0.06 (-0.9113)
Father's schooling	3.6 (1.4)	3.6 (1.4)	-0.02 (-0.42)
Quecha mainly	.00099 (0.0315)	0.0114 (0.106)	-0.00414* (-2.85)
Amarya mainly	.00396 (0.0628)	0.00456 (0.0675)	-0.000605 (-0.23)
Home ownership	0.428 (0.495)	0.215 (0.411)	0.213* (12.02)
Number of rooms in the house	1.50 (1.05)	1.40 (1.00)	0.100* (4.17)
Both parents present	0.841 (0.366)	0.775 (0.418)	0.0656* (4.54)
Number of siblings	2.37 (1.80)	2.05 (1.59)	0.322* (4.80)
Ownership of durables	6.30 (2.11)	5.92 (1.92)	0.375* (4.69)
Job of mother	2.26 (0.91)	2.08 (0.91)	0.174* (4.73)
Job of father	2.70 (0.54)	2.70 (0.55)	-0.006 (-0.28)
Household income	922 (755)	868 (638)	55* (2.68)

[a] Value of two-sample t test with unequal variances given in parentheses in last column. * indicates significance at 5 percent level and ** at 10 percent.

Notes: (1) Stunting and wasting refer to the Z scores for height and weight based on NCHS/CDC/WHO standards. "Moderate" refers to being more than one standard deviation below the means and "severe" refers to being more than two standard deviations below the means. (2) Ownership of durables measures number of durables owned out of 15 asked. (3) Job of mother/job of father: 1=no job; 2=temporary job; 3=permanent job.

Table 3. Kenya. T-tests for Differences in Means in Kenya 1 Data for Attritors versus Nonattritors [a]						
	Men			Women		
Variables	Means for Nonattritors (S. D.)	Means for Attritors (S. D.)	Difference in Means (T test)	Means for Nonattritors (S. D.)	Means for Attritors (S. D.)	Difference in Means (T test)
Fertility-Related Outcome Variables						
Currently using contraceptives	0.196 (0.017)	0.229 (0.031)	-0.033 (-0.95)	0.126 (0.012)	0.103 (0.021)	0.024 (0.91)
Ever used contraceptives	0.233 (0.018)	0.311 (0.052)	-0.077** (-1.79)	0.238 (0.016)	0.196 (0.027)	0.042 (1.25)
Want no more children	0.208 (0.017)	0.237 (0.031)	-0.029 (-0.83)	0.351 (0.018)	0.220 (0.037)	0.132* (3.59)
Number of surviving children	4.76 (0.171)	3.94 (0.277)	0.817* (2.46)	3.88 (0.089)	2.78 (0.138)	1.10 (5.90*)
Family Planning Program Variables						
Visited by community-based distribution agent	0.156 (0.015)	0.132 (0.025)	0.024 (0.78)	0.163 (0.014)	0.113 (0.022)	0.050** (1.75)
Heard family planning message on radio	0.931 (0.011)	0.963 (0.013)	-0.037** (-1.86)	0.870 (0.916)	0.916 (0.019)	-0.046** (-1.79)
Heard about family planning at clinic	0.495 (0.021)	0.513 (0.036)	-0.018 (-0.42)	0.851 (0.013)	0.828 (0.027)	0.023 (0.80)
Discussed with others family planning lecture heard at clinic	0.679 (0.029)	0.691 (0.047)	-0.012 (-0.21)	0.629 (0.070)	0.661 (0.037)	-0.032 (-0.76)
Number of Network Partners in Network for						
Family planning	3.7 (0.20)	4.0 (0.35)	-0.3 (-0.86)	2.9 (0.11)	3.1 (0.20)	-.18 (-0.78)
Wealth flows	5.0 (0.21)	5.0 (0.36)	-0.04 (-0.10)	2.8 (0.12)	2.4 (0.21)	0.38 (1.45)
Reproductive Health	-	-	-	3.2 (0.16)	2.8 (0.23)	0.38 (1.19)
Knows secret contraceptive user	0.637 (0.069)	0.558 (0.095)	0.079 (0.60)	0.408 (0.02)	0.377 (0.03)	0.030 (0.77)
Control Variables:						
Age (years)	40.1 (0.52)	36.8 (0.78)	3.3* (3.24)	29.7 (0.332)	26.3 (0.488)	3.4* (5.04)
Education						
No schooling	0.112 (0.013)	0.063 (0.018)	0.049** (1.94)	0.214 (0.015)	0.141 (0.024)	0.072** (2.30)
Some primary schooling	0.577 (0.021)	0.537 (0.036)	0.040 (0.96)	0.669 (0.018)	0.668 (0.033)	0.001 (0.03)
Secondary schooling	0.298 (0.019)	0.379 (0.035)	-0.081* (-2.06)	0.117 (0.012)	0.190 (0.027)	-0.074* (-2.75)
Language						
Luo only	0.796 (0.017)	0.805 (0.029)	-0.010 (-0.28)	0.422 (0.018)	0.327 (0.033)	0.095** (2.46)
English	0.443 (0.021)	0.532 (0.036)	-0.089* (-2.11)	0.178 (0.014)	0.263 (0.031)	-0.086* (-2.73)
Swahili	0.655 (0.020)	0.726 (0.032)	-0.072** (-1.82)	0.396 (0.018)	0.517 (0.035)	-0.121* (-3.11)

Lived:						
–outside of province	0.591 (0.021)	0.653 (0.035)	0.061 (1.49)	0.370 (0.018)	0.371 (0.034)	-0.001 (-0.02)
–in Nairobi or Mombasa	0.336 (0.020)	0.400 (0.036)	-0.064 (-1.58)	0.214 (0.015)	0.205 (0.028)	0.009 (0.29)
Belongs to credit group	0.257 (0.019)	0.242 (0.031)	0.015 (0.40)	0.351 (0.018)	0.288 (0.032)	0.064** (1.70)
Belong to clan welfare society	0.868 (0.014)	0.905 (0.021)	-0.037 (-1.35)	0.747 (0.016)	0.644 (0.034)	0.103* (2.93)
Women sells on market	–	–	–	0.464 (0.019)	0.444 (0.035)	0.020 (0.51)
Household characteristics						
Polygamous household	0.293 (0.019)	0.238 (0.031)	0.055 (1.45)	0.350 (0.018)	0.371 (0.034)	-0.021 (-0.56)
Self/Husband receives monthly Salary	0.170 (0.016)	0.255 (0.032)	-0.085* (-2.56)	0.334 (0.019)	0.402 (0.037)	-0.068** (-1.66)
Husband interviewed	–	–	–	0.765 (0.016)	0.752 (0.029)	0.013 (0.41)
Household has radio	–	–	–	0.492 (0.019)	0.546 (0.035)	-0.055 (-1.38)
House has metal roof	0.173 (0.016)	0.189 (0.029)	-0.016 (-0.51)	0.201 (0.015)	0.187 (0.027)	0.014 (0.45)
Sublocation of residence						
Gwassi	0.278 (0.019)	0.216 (0.030)	0.063** (1.69)	0.213 (0.015)	0.210 (0.029)	0.003 (0.08)
Kawadhgone	0.230 (0.018)	0.237 (0.031)	-0.007 (-0.20)	0.240 (0.015)	0.205 (0.028)	0.035 (1.06)
Oyugis	0.259 (0.019)	0.300 (0.033)	-0.041 (-1.11)	0.286 (0.017)	0.263 (0.031)	0.023 (0.63)
Ugina	0.233 (0.018)	0.247 (0.032)	-0.014 (-0.39)	0.261 (0.016)	0.322 (0.033)	-0.061** (-1.72)

[a] Value of two-sample t test with unequal variances given in parentheses in third and sixth columns. * indicates significance at 0.05 level, and ** at 0.10 level.

Table 3. South Africa. T-tests for Differences in Means in South Africa 1 Data for Attritors versus Nonattritors						
	Nonattritors		Attritors		Difference	
	Means	(S.D.)	Means	(S.D.)	In Means	(T-test)
Early Child Nutritional Status and Health Outcome Variables						
Height-for-age	0.377	(0.008)	0.377	(0.016)	0.000	(1.00)
Weight-for-age	5.369	(0.107)	5.281	(0.195)	0.088	(0.69)
Weight-for-height	14.83	(0.099)	14.74	(0.198)	0.090	(0.68)
Height-for-age z-score	-1.171	(0.073)	-1.338	(0.143)	0.167	(1.04)
Weight-for-age z-score	-0.621	(0.058)	-0.742	(0.106)	0.122	(1.00)
Weight-for-height z-score	0.179	(0.070)	0.113	(0.136)	0.066	(0.43)
Moderate stunting	0.539	(0.015)	0.534	(0.035)	0.005	(0.13)
Severe stunting	0.275	(0.017)	0.284	(0.032)	-0.009	(-0.25)
Moderate wasting	0.389	(0.018)	0.441	(0.035)	-0.052	(-1.32)
Severe wasting	0.185	(0.015)	0.172	(0.026)	0.014	(0.46)
Sick in last 2 weeks	0.104	(0.011)	0.098	(0.021)	0.006	(0.25)
Predetermined Family Background Variables						
Age in months	37.36	(0.671)	37.51	(1.260)	-0.146	(-0.10)
Fraction male	0.501	(0.019)	0.490	(0.035)	0.011	(0.28)
Fraction African	0.912	(0.011)	0.863	(0.024)	0.049**	(1.85)
Household size	8.817	(0.144)	8.500	(0.289)	0.317	(0.98)
Total monthly expenditures	1473.3	(30.19)	1545.4	(65.47)	-72.1	(-1.00)
Per capita monthly exp.	194.2	(5.55)	219.3	(12.91)	-25.1**	(-1.79)
Total monthly income	1160.6	(45.02)	1396.3	(97.41)	-235.7*	(-2.20)
Per capita monthly income	156.8	(7.88)	215.8	(20.86)	-59.1*	(-2.65)
Household Head Age	51.75	(0.515)	52.98	(1.076)	-1.235	(-1.03)
Household head education	2.978	(0.123)	3.453	(0.250)	-0.476**	(-1.70)
Household head male	0.698	(0.017)	0.711	(0.032)	-0.013	(-0.35)
Own house	0.886	(0.012)	0.843	(0.026)	0.043	(1.53)
Number of rooms	4.949	(0.099)	5.377	(0.211)	-0.428**	(-1.84)
Number of durables	3.132	(0.081)	3.608	(0.146)	-0.476*	(-2.85)
Urban	0.278	(0.017)	0.294	(0.032)	-0.016	(-0.44)
In former Natal	0.160	(0.014)	0.225	(0.029)	-0.065*	(-2.02)
Notes: (1) Value of two-sample t test with unequal variances given in parentheses in last column. * indicates significance at 5 percent level and ** at 10 percent level. (2) Height-for-age in cm/years. Weight-for-age in kg/years. (3) Stunting and wasting based height-for-age and weight-for-age z-scores calculated based on NCHS/CDC/WHO standards. "Moderate" refers to being more than one standard deviation below the means and "severe" more than 2 standard deviations below mean.						

Table 4. Probits for Predicting Attrition between Rounds 1 and 2 for Bolivian, Kenyan and South African Data [a]

Bolivia			Kenyan Men			Kenyan Women		South Africa		
Outcome Variables	**Outcome Var. One at a Time**	**All Outcome Var + Predet. Var.[b]**	**Outcome Variables**	**Outcome Var. One at a Time**	**All Outcome Var + Predet. Var.[c]**	**Outcome Var. One at a Time**	**All Outcome Var + Predet. Var.[d]**	**Outcome Variables**	**Outcome Var. One at a Time**	**All Outcome Var + Predet. Var.[e]**
Height-for-age	-.0015 (-0.83)	-.0002 (-0.04)	Currently contracepting	0.118 (0.95)	-0.065 (0.34)	-0.134 (0.92)	0.004 (0.02)	Height-for-age	-0.001 (0.01)	1.376 (1.44)
Weight-for-height	-.0015 (-0.99)	.0032 (0.80)	Ever used contraceptives	0.162** (1.67)	-0.103 (0.70)	-0.142 (1.26)	-0.036 (0.28)	Weight-for-height	-0.007 (0.37)	0.042 (1.08)
Weight-for-age	-.003** (-1.74)	-.0037 (-0.78)	Want no more children	0.099 (0.83)	0.245** (1.69)	-0.374* (3.60)	-0.010 (0.07)	Weight-for-age	-0.006 (0.42)	-1.355 (0.04)
Moderate wasting	.148** (1.78)	.1003 (0.70)	No. surviving children	-0.033* (2.46)	-0.017 (0.78)	-0.139* (5.82)	-0.136* (3.73)	Moderate wasting	0.125 (1.19)	0.279 (1.62)
Severe wasting	.191 (1.35)	.1353 (0.70)	No. family planning network partners	-0.009 (0.85)	0.003 (0.22)	0.012 (0.78)	-0.010 (0.56)	Severe wasting	-0.055 (0.47)	-0.119 (0.81)
Moderate stunting	-.0315 (-0.38)	-.291** (-1.93)						Moderate stunting	-0.012 (0.13)	-0.040 (0.38)
Severe stunting	.2110* (2.41)	.2066 (1.51)						Severe stunting	0.026 (0.22)	0.056 (1.62)
Bulk motor ability	-.009 (-1.64)	.0123 (0.59)						Sick in last two weeks	-0.038 (0.23)	-0.055 (0.32)
Fine motor ability	-.009 (-1.63)	-.0073 (-0.35)								
Language-audition	-.010** (-1.84)	-.0059 (-0.27)								
Personal-social	-.008 (-1.64)	-.0014 (-0.07)								
Constant		.75** (1.72)			-0.239 (0.70)		-0.097 (0.29)			-1.271 (-0.93)
Chi^2 test [prob > Chi^2]	f	300.22 [0.00]		g	25.13 [0.068]	h	54.49 [0.000]		i	24.63 [0.22]

[a] Absolute value of z test in parentheses beneath point estimates: * indicates significance at 5 percent level, ** indicates significance at 10 percent level.

[b] Predetermined variables for Bolivian households that are: (a) significant at 5 percent level (with sign in parentheses)—father's age(+); Quecha only (+); ownership of house (-); number of durables owned (-); Oruro (-), Postosi (-), Santa Cruz (-) relative to La Paz; mother's job permanent relative to no job (-); (b) significant at the ten percent level – father's schooling (-), number of rooms in the house (+), number of siblings of child (-); father's job temporary relative to no job (-); (c) not significant even at the ten percent level – mother's age, mother's schooling, Amarya only, El Alto, Cochabama, Tarija relative to La Paz; father's job permanent relative to no job; mother's job temporary relative to no job; household income.

[c] Predetermined variables for Kenyan men that are: (a) significant at the five percent level (with sign in parentheses)—men's age; (b) not significant even at the ten percent level – primary schooling; secondary schooling; Luo only; English; lived in Nairobi or Mombasa; polygamous household; earns a monthly salary; sublocation of residence.

[d] Predetermined variables for Kenyan women that are: (a) significant at the five percent level (with sign in parentheses)—husband interviewed (-); (b) significant at the ten percent level—resided in Oyugnis relative to Ugina (-) (c) not significant even at the ten percent level—primary schooling; secondary schooling; Luo only; English; lived in Nairobi or Mombasa; polygamous household; household has radio; household has metal roof; other sublocation of residence.

[e] Predetermined variables for South African households that are: (a) significant at the five percent level (with sign in parentheses)—age of household head(+); (b) significant at the ten percent level—own on home (-); (c) not significant even at the ten percent level—male respondent; African respondent; household size; ln total monthly expenditures; household head schooling; household head sex; number of rooms; number of durables; urban; former Natal.

[f] For Bolivian data, Probability .> Chi^2 (a) at the five percent level—severe stunting; (b) at the ten percent level—weight-for-age, moderate wasting, language-auditory.

[g] For Kenyan men, Probability .> Chi^2 (a) at the five percent level—number of surviving children; (b) at the ten percent level—ever-used contraceptives.

[h] For Kenyan women, Probability .> Chi^2 (a) at the five percent level—want no more children, number of surviving children.

[i] For South African data, Probability .> Chi^2 (a) at the five percent level—none; (b) at the ten percent level—none

Table 5. Bolivia (1). Testing Impact of Attrition between Bolivia 1 and Bolivia 2 on Coefficient Estimates of Family Background Variables in Early Childhood Development Anthropometric Outcomes [a]

	Ordinary Least Squares Regressions for				Probits for		
Right-Side Variables	**Height for age**	**Weight for age**	**Weight for height**	**Moderate Stunting**	**Severe Stunting**	**Moderate Wasting**	**Severe Wasting**
Predetermined Family Background Variables							
Mother's age	-0.0369 (-0.31)	0.162 (1.13)	0.214 (1.46)	-0.00933 (-0.79)	-.00363 (-0.27)	-0.00352 (-0.29)	0.0142 (0.67)
Father's age	0.222* (2.29)	0.130 (1.13)	-0.072 (-0.61)	-0.00558 (-0.58)	-0.0165 (-1.50)	-.0209* (-2.08)	-0.0186 (-1.06)
Mother's schooling	0.998* (2.40)	1.51* (3.05)	0.611 (1.20)	---	---	---	—
Father's schooling	-0.143 (-0.34)	-0.407 (-0.82)	-0.534 (-1.05)	---	---	---	-0.106 (-1.37)
Quecha mainly	-3.58 (-0.23)	-7.23 (-0.40)	-1.05 (-0.06)	16.4* (21.42)	-0.667 (-0.46)	17.3* (25.26)	—
Amarya mainly	-0.010 (-0.00)	-3.19 (-0.35)	-7.47 (-0.79)	-0.755 (-1.00)	0.476 (0.65)	0.313 (0.43)	—
Ownership of house	-1.37 (-1.20)	-1.07 (-0.79)	0.075 (0.05)	0.0537 (0.46)	0.0183 (0.15)	-0.0225 (-0.20)	—
Number of rooms in the house	1.48* (2.44)	1.15 (1.59)	0.108 (0.15)	-0.0523 (-0.86)	-0.0591 (-0.83)	-0.0127 (-0.21)	-0.0269 (-0.23)
Number of siblings	-1.76* (-5.08)	-1.50* (-3.63)	0.133 (0.31)	0.182* (4.99)	0.242* (6.42)	0.104* (3.00)	—
Ownership of durables	0.946* (3.28)	0.535 (1.56)	-0.246 (-0.70)	—	—	---	-0.172* (-3.13)
El Alto	0.036 (0.03)	-0.135 (-0.08)	2.149 (1.182)	.262** (1.70)	0.343* (2.22)	-0.0610 (-0.42)	-0.150 (-0.54)
Cochabama	4.63* (2.94)	-2.17 (-1.16)	-6.01* (-3.12)	—	---	0.130 (0.84)	—
Oruro	-4.43* (-2.10)	-6.89* (-2.75)	1.12 (0.44)	0.526* (2.29)	0.551* (2.56)	0.509* (2.53)	0.676* (2.10)
Potosi	-0.869 (-0.43)	-10.0* (-4.16)	-11.93* (-4.83)	0.229 (1.08)	0.481* (2.34)	0.936* (4.78)	—
Tarija	6.65* (3.18)	14.35* (5.76)	12.4* (4.83)	-0.189 (-0.91)	-0.0944 (-0.41)	-0.723* (-3.10)	—
Santa Cruz	9.65* (6.28)	5.02* (2.74)	-2.27 (-1.21)	-0.748* (-4.92)	-0.673* (-3.67)	-0.346* (-2.21)	-0.372 (-1.26)
Job of father is temporary	-4.77** (-1.79)	-7.29* (-2.30)	-3.85 (-1.18)	0.411 (1.57)	0.6766** (2.06)	0.372 (1.35)	—
Job of father is permanent	-4.38** (-1.73)	-6.38* (-2.12)	-2.88 (-0.93)	0.393 (1.59)	0.679* (2.14)	0.282 (1.07)	0.0729 (0.16)
Job of mother is temporary	-4.80* (-2.84)	-3.53** (-1.75)	2.63 (1.27)	0.544* (3.04)	0.692* (3.90)	0.268** (1.61)	0.0967 (0.33)
Job of mother is permanent	-3.23* (-2.91)	-1.92 (-1.46)	2.37** (1.75)	0.250* (2.26)	0.390* (3.07)	0.226* (2.01)	0.0356 (0.18)

Household income	.00121** (1.62)	.000558 (0.63)	-.000538 (-0.59)	-0.000065 (-0.86)	-0.000164** (-1.64)	-0.0000262 (-0.33)	-0.0000376 (-0.25)
constant	10.28* (2.51)	27.19 *(5.58)	57.91*(11.58)	0.845*(2.07)	-0.901**(-1.87)	-0.00232(-0.01)	-1.39**(-1.91)
F test for overall relation [probability > F test]	7.11* [0.0000]	5.58 * [0.0000]	4.02* [0.0000]	257.80* [0.0000]	278.38* [0.0000]	179.06* [0.0000]	98.91* [0.0000]
F Tests for attrition [probability > F]							
1. joint effect of attrition on constant and all estimates	1.32 [0.1428]	1.88* [0.0070]	1.58* [0.0385]	22.68 [0.3614]	35.34** [0.0357]	44.86* [0.0018]	261.66* [0.0000]
2. joint effect of attrition on all coefficient estimates but not on constant	1.37 [0.1169]	1.90* [0.0068]	1.63* [0.0315]	22.49 [0.3147]	29.18 [0.1097]	42.17* [0.0026]	253.89* [0.0000]
* t test indicates that significantly nonzero at 5 percent level. ** t test indicates that significantly nonzero at 10 percent level.							

Table 5. Bolivia (2). Multivariate Ordinary Least Squares Regressions for Testing Impact of Attrition between Bolivia 1 and Bolivia 2 on Coefficient Estimates of Family Background Variables in Child Test Scores [a]				
Right-Side Variables	**Gross Motor Ability**	**Fine Motor Ability**	**Language-Auditory**	**Personal-Social**
Predetermined Family Background Variables				
Mother's age	0.204* (4.84)	0.189* (4.80)	0.203* (4.96)	0.199* (4.57)
Father's age	-0.00767 (-0.23)	0.00268 (0.08)	0.0118 (0.36)	0.00547 (0.16)
Mother's schooling	-0.257** (-1.75)	-0.127 (-0.93)	-0.0290 (-0.20)	-0.167 (-1.10)
Father's schooling	0.236** (1.61)	0.219 (1.60)	0.159 (1.12)	0.209 (1.38)
Quecha mainly	2.85 (0.53)	2.88 (0.57)	3.32 (0.63)	4.28 (0.77)
Amarya mainly	-4.01 (-1.47)	-3.05 (-1.19)	-3.091 (-1.17)	-2.91 (-1.03)
Ownership of house	-0.167 (-0.41)	0.137 (0.36)	-0.123 (-0.31)	—
Number of rooms in the house	-0.0260 (-0.12)	0.0373 (0.19)	-0.0751 (-0.36)	0.0433 (0.20)
Number of siblings	-0.0370 (-0.30)	-0.139 (-1.21)	-0.00220 (-0.02)	-0.103 (-0.81)
Ownership of durables	0.335* (3.30)	0.278* (2.92)	0.395* (4.00)	0.403* (3.84)
El Alto	1.70* (3.26)	1.49* (3.07)	1.87* (3.71)	1.84* (3.43)
Cochabama	0.569 (1.03)	-0.254 (-0.49)	0.156 (0.29)	0.675 (1.18)
Oruro	.537 (0.72)	-0.337 (-0.49)	0.761 (1.06)	0.401 (0.52)
Potosi	-1.08 (-1.51)	-1.23** (-1.85)	-0.720 (-1.04)	-1.07 (-1.45)
Tarija	4.01* (5.43)	2.64* (3.83)	3.31* (4.63)	3.68* (4.83)
Santa Cruz	2.05* (3.79)	1.09* (2.16)	1.63* (3.10)	—
Job of father is temporary	---	-1.79** (-2.05)	-1.77** (-1.95)	-1.69** (-1.75)
Job of father is permanent	-2.35* (-2.64)	-2.03* (-2.44)	-2.09* (-2.42)	-2.02* (-2.20)
Job of mother is temporary	2.20* (3.69)	1.92* (3.45)	---	2.17* (3.53)
Job of mother is permanent	0.948* (2.43)	0.900* (2.45)	0.844* (2.22)	1.06* (2.63)
Household income	.000068 (0.26)	.0000878 (0.36)	-0.0000282 (-0.11)	-0.0000404 (-0.15)

Constant	13.4* (9.28)	12.47 * (9.25)	10.28* (7.35)	11.4* (7.62)
F test for overall relation [probability > F test]	5.38* [0.0000]	5.21* [0.0000]	5.80* [0.0000]	5.39* [0.0000]
F Tests for Attrition [probability > F]				
1. joint effect of attrition on all estimates, including constant	1.31 [0.1461]	1.45** [0.0772]	1.34 [0.1277]	1.38 [0.1055]
2. joint effect of attrition on all coefficients but not on constant	1.37 [0.1160]	1.51** [0.0594]	1.40 [0.1013]	1.44** [0.0824]

Table 5. Kenya. *Multivariate Probits/Regressions for Testing* Impact of Attrition for Men and Women between Kenya 1 and Kenya 2 on Key Fertility-Related Outcome Variables [a]

Right-Side Variables	Men					Women				
	Probits			OLS Regressions		Probits			OLS Regressions	
	Currently using contra-ceptives	Ever used contra-ceptives	Want no more children	Number of surviving children	Family planning social network size	Currently using contra-ceptives	Ever used contra-ceptives	Want no more children	Number of surviving children	Family planning social network size
Control Variables										
Age (years)	0.004 (0.74)	0.009 (1.62)	0.013* (8.58)	0.200* (20.26)	0.015 (0.86)	0.014* (2.03)	0.023* (3.68)	0.079* (11.80)	0.161* (20.82)	0.025* (1.97)
Education (relative to no schooling)										
Primary schooling	0.075 (0.36)	-0.048 (0.26)	0.133 (0.69)	0.955* (2.85)	1.202* (2.08)	0.122 (0.72)	0.094 (0.66)	-0.004 (0.03)	-0.440* (2.66)	0.957* (3.41)
Secondary schooling	0.310 (1.22)	0.122 (0.55)	0.197 (0.81)	0.736** (1.77)	2.247* (3.12)	0.125 (0.47)	0.279 (1.23)	-0.107 (0.46)	-0.447 (1.60)	1.786* (3.83)
Language										
Luo only	0.372** (1.87)	0.368* (2.37)	0.142 (0.89)	-0.180 (0.66)	0.815** (1.74)	-0.268** (1.86)	-0.236* (1.95)	-0.228** (1.88)	-0.142 (1.00)	-0.395** (1.68)
English	-0.037 (0.24)	-0.048 (0.33)	0.074 (0.46)	0.325 (1.20)	0.243 (0.52)	0.264 (1.41)	0.265 (1.59)	-0.002 (0.01)	-0.334 (1.59)	0.125 (0.36)
Lived in Nairobi or Mombasa	0.130 (1.12)	0.221* (2.02)	0.324* (2.74)	0.086 (0.41)	0.258 (0.71)	0.311* (2.33)	0.356* (3.05)	0.240* (2.01)	0.144 (0.97)	-0.066 (0.26)
Women sell in market	–	–	–	–	–	0.254* (2.02)	0.147 (1.34)	-0.119 (1.07)	0.032 (0.24)	0.180 (0.83)
Household characteristics										
Polygamous household	0.091 (0.65)	-0.025 (0.19)	-0.296* (2.10)	2.386* (9.69)	0.017 (0.04)	-0.161 (1.28)	-0.104 (0.97)	0.187** (1.79)	-0.201 (1.57)	-0.089 (0.42)
Earns a monthly salary	0.058 (0.38)	0.302* (2.16)	0.251 (1.63)	0.312 (1.13)	0.953* (2.00)	–	–	–		
Husband interviewed	–	–	–	–	–	0.211 (1.51)	-0.108 (0.94)	-0.113 (0.99)	-0.147 (1.05)	0.101 (0.44)

Household has radio	–	–	–	–	--	-0.019 (0.16)	-0.005 (0.05)	0.046 (0.44)	-0.106 (0.85)	0.270 (1.31)
Household has metal roof	–	–	–	–	--	0.003 (0.019)	0.253* (2.00)	0.173 (1.39)	0.810* (5.15)	0.142 (0.53)
Sublocation of residence (relative to Ugina)										
Gwassi	-0.639* (3.42)	-0.571* (3.50)	-0.630* (3.42)	-0.032 (0.11)	-0.323 (0.66)	-0.441* (2.37)	-0.645* (4.10)	0.169 (1.13)	0.357* (2.03)	-0.668* (2.29)
Kawadhgone	0.145 (0.88)	0.015 (0.09)	0.153 (0.93)	0.165 (0.57)	-0.182 (0.36)	-0.170 (0.99)	-0.260** (1.79)	0.130 (0.85)	0.240 (1.34)	0.496** (1.68)
Oyugis	0.256 (1.62)	0.239** (1.67)	0.328* (2.10)	0.229 (0.82)	-0.392 (0.81)	0.013 (0.08)	-0.179 (1.26)	0.437* (2.93)	0.218 (1.23)	1.537* (5.22)
Constant	-1.53* (4.38)	-1.43* (4.67)	-3.34* (9.31)	-4.96* (8.94)	0.970 (1.02)	-1.85* (5.50)	-1.34* (4.71)	-3.03* (10.01)	-0.90* (2.57)	1.87* (3.23)
Chi squared test for overall relation [probability > Chi squared]	48.87* [0.0000]	58.21* [0.0000]	134.25* [0.0000]			44.22* [0.0001]	86.05* [0.0000]	234.12* [0.0000]		
R squared F test [probability > F]				0.560 82.81* [0.0000]	0.057 3.98* [0.0005]				0.469 50.36* [0.0000]	0.082 5.48* [0.0000]
Tests for Attrition										
Effect of attrition on constant	0.027 (0.21)	0.046 (0.38)	0.150 (1.13)	-0.065 (0.29)	0.166 (0.42)	0.126** (1.90)	-0.162 (1.31)	-0.189 (1.50)	-0.549* (3.77)	0.057 (0.24)
Chi squared test for joint effect of attrition on constant and all coefficient estimates [probability > Chi squared] (F tests for regressions)	12.11 [0.437]	11.27 [0.506]	16.79 [0.158]	1.11 [0.352]	0.71 [0.725]	10.85 [0.763]	12.60 [0.633]	10.68 [0.775]	2.08* [0.009]	0.82 [0.657]
Chi squared test for joint effect of attrition on all coefficient estimates but not on constant [probability > Chi squared] (F tests for regressions)	11.90 [0.371]	11.04 [0.440]	15.27 [0.171]	1.20 [0.284]	0.67 [0.781]	10.74 [0.706]	11.58 [0.640]	9.20 [0.818]	1.05 [0.397]	0.87 [0.588]

[a] Absolute value of z test (for probits) and t tests (for regressions) in parentheses beneath point estimates: * indicates significance at the 5 percent level, and ** at the ten percent level.

Table 5. South Africa. Multivariate Regressions/Probits for Testing Impact of Attrition between South Africa 1 and South Africa 2 on Child Nutritional Status and Health								
	Height-for-age	**Weight-for-age**	**Weight-for-height**	**Moderate stunting**	**Severe stunting**	**Moderate wasting**	**Severe wasting**	**Sick in past 2 weeks**
Control Variables								
Respondent male	0.019	0.243	-0.028	0.116	0.132	0.160	0.114	0.032
	(1.08)	(1.05)	(0.14)	(1.21)	(1.28)	(1.50)	(1.00)	(0.21)
Respondent African	0.007	0.451	1.001*	0.044	0.069	-0.888*	0.288	-0.125
	(0.17)	(0.80)	(2.47)	(0.15)	(0.18)	(2.75)	(0.89)	(0.33)
Household size	0.002	-0.013	-0.083*	0.007	-0.020	0.020	0.013	-0.042
	(0.55)	(0.24)	(2.24)	(0.40)	(0.80)	(0.98)	(0.56)	(1.48)
Log total monthly expenditures	0.000	0.092	0.228	-0.159	-0.215	-0.200	0.044	-0.023
	(0.01)	(0.34)	(0.95)	(1.31)	(1.40)	(1.28)	(0.30)	(0.13)
Household head age	0.000	0.005	0.005	-0.003	0.005	0.002	0.002	-0.012
	(0.09)	(0.40)	(0.46)	(0.79)	(1.01)	(0.34)	(0.43)	(1.68)
Household head schooling	-0.002	-0.050	-0.032	-0.017	0.014	0.011	0.002	0.016
	(0.68)	(1.03)	(0.76)	(0.84)	(0.55)	(0.53)	(0.08)	(0.54)
Household head male	-0.015	-0.317	-0.202	-0.029	0.004	0.154	0.221	-0.062
	(0.87)	(1.42)	(0.82)	(0.26)	(0.03)	(1.28)	(1.76)	(0.40)
Own house	0.024	-0.130	-0.813*	0.090	0.431	0.560*	0.554*	0.025
	(0.78)	(0.33)	(2.93)	(0.55)	(1.88)	(3.13)	(2.60)	(0.09)
Number of rooms	0.001	0.054	0.083	-0.012	0.018	-0.044	-0.057*	-0.056
	(0.30)	(1.25)	(1.62)	(0.61)	(0.75)	(1.69)	(2.40)	(1.49)
Number of durables	-0.001	0.020	0.093	-0.040	-0.050	-0.063	-0.048	-0.011
	(0.26)	(0.27)	(1.29)	(1.04)	(1.06)	(1.61)	(1.04)	(0.21)
Urban	0.008	-0.126	-0.536	-0.185	-0.115	0.161	0.317	0.347
	(0.38)	(0.47)	(1.37)	(1.02)	(0.50)	(0.88)	(1.56)	(1.42)
Former Natal	0.027	0.424	0.277	-0.250	-0.296	-0.420	-0.184	-0.306
	(0.72)	(0.86)	(0.97)	(1.41)	(0.93)	(1.42)	(0.69)	(1.19)
Constant	0.327*	4.160*	13.150*	1.517	0.404	1.406	-2.082*	0.116
	(2.30)	(2.18)	(8.47)	(1.68)	(0.35)	(1.32)	(1.92)	(0.09)
F-test overall (Cols 1-3)	1.80*	2.00*	1.43	113.27*	86.29*	51.43*	49.34*	6842.91*
Chi-2 test overall (Cols 4-7) [p-value]	[0.03]	[0.01]	[0.12]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Tests for Attrition								
Effect of attrition on constant	0.462	5.504	1.818	-5.314*	-3.746	-2.970	0.103	
	(1.61)	(1.61)	(0.32)	(2.38)	(1.38)	(1.40)	(0.04)	n/a
	[011]	[0.11]	[0.75]	[0.02]	[0.17]	[0.16]	[0.97]	
Joint effect of attrition on constant and all estimates - p-value [p-value]	1.52	1.32	0.88	30.26*	16.81	10.31	5.82	n/a
	[0.13]	[0.22]	[0.58]	[0.00]	[0.21]	[0.67]	[0.95]	
Joint effect of attrition on all estimates but constant [p-value]	1.64	1.43	0.91	30.20*	16.56	6.49	5.82	4187.32*
	[0.10]	[0.18]	[0.54]	[0.00]	[0.17]	[0.89]	[0.92]	[0.00]

Notes: Columns 1–3 ordinary least squares and columns 4–7 probit estimation. All estimated allowing for clustering at community level and with robust standard errors to account for multiple observations on same households within communities. Absolute value of t tests (for regressions) and z test (for probits) in parentheses: * indicates significance at the 5 percent level, and ** at the ten percent level. P-values of tests in brackets.

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS2429	Corruption, Composition of Capital Flows, and Currency Crises	Shang-Jin Wei	August 2000	H. Sladovich 37698
WPS2430	Financial Structure and Bank Profitability	Asli Demirgüç-Kunt Harry Huizinga	August 2000	K. Labrie 31001
WPS2431	Inside the Crisis: An Empirical Analysis of Banking Systems in Distress	Asli Demirgüç-Kunt Enrica Detragiache Poonam Gupta	August 2000	K. Labrie 31001
WPS2432	Funding Growth in Bank-Based and Market-Based Financial Systems: Evidence from Firm-Level Data	Asli Demirgüç-Kunt Vojislav Maksimovic	August 2000	K. Labrie 31001
WPS2433	External Interventions and the Duration of Civil Wars	Ibrahim A. Elbadawi Nicholas Sambanis	September 2000	H. Sladovich 37698
WPS2434	Socioeconomic Inequalities in Child Malnutrition in the Developing World	Adam Wagstaff Naoko Watanabe	September 2000	A. Maranon 38009
WPS2435	The Evolution of Poverty during the Crisis in Indonesia, 1996–99	Asep Suryahadi Sudarno Sumarto Yusuf Suharso Lant Pritchett	September 2000	P. Sader 33902
WPS2436	Safety Nets and Safety Ropes: Who Benefited from Two Indonesian Crisis Programs—the "Poor" or the "Shocked?"	Sudarno Sumarto Asep Suryahadi Lant Pritchett	September 2000	P. Sader 33902
WPS2437	Quantifying Vulnerability to Poverty: A Proposed Measure, Applied to Indonesia	Lant Pritchett Asep Suryahadi Sudarno Sumarto	September 2000	P. Sader 33902
WPS2438	Measurements of Poverty in Indonesia: 1996, 1999, and Beyond	Menno Pradhan Asep Suryahadi Sudarno Sumarto Lant Pritchett	September 2000	P. Sader 33902
WPS2439	State-Community Synergies in Development: Laying the Basis for Collective Action	Monica Das Gupta Helene Grandvoinnet Mattia Romani	September 2000	M. Das Gupta 31983
WPS2440	Lessons from Uganda on Strategies to Fight Poverty	John Mackinnon Ritva Reinikka	September 2000	H. Sladovich 37698
WPS2441	Controlling the Fiscal Costs of Banking Crises	Patrick Honohan Daniela Klingebiel	September 2000	A. Yaptenco 38526

Policy Research Working Paper Series

	Title	Author	Date	Contact for paper
WPS2442	A Firms's-Eye View of Policy and Fiscal Reforms in Cameroon	Bernard Gauthier Isidro Soloaga James Tybout	September 2000	L. Tabada 36896
WPS2443	The Politics of Economic Policy Reform in Developing Countries	Richard H. Adams Jr.	September 2000	M. Coleridge-Taylor 33704
WPS2444	Seize the State, Seize the Day: State Capture, Corruption, and Influence in Transition	Joel S. Hellman Geraint Jones Daniel Kaufmann	September 2000	D. Billups 35818
WPS2445	Subsidies in Chilean Public Utilities	Pablo Serra	September 2000	G. Chenet-Smith 36370
WPS2446	Forecasting the Demand for Privatized Transport: What Economic Regulators Should Know, and Why	Lourdes Trujillo Emile Quinet Antonio Estache	September 2000	G. Chenet-Smith 36370