

Measuring Living Standards with Proxy Variables

Mark R. Montgomery
Michele Gragnolati
Kathleen Burke
Edmundo Paredes

1999 No. 129

Measuring Living Standards with Proxy Variables

Mark R. Montgomery

Michele Gragnolati

Kathleen Burke

Edmundo Paredes

Mark R. Montgomery is Senior Associate, Policy Research Division, Population Council and Associate Professor of Economics, State University of New York at Stony Brook. Michele Gragnolati is a Young Professional, World Bank. Kathleen Burke is Assistant Professor, Economics Department, State University of New York at Cortland. Edmundo Paredes is a graduate student in the Department of Economics, SUNY-Stony Brook. We have benefited from comments by Douglas Anderton, Kenneth Bollen, John Bongaarts, John Casterline, Dean Jolliffe, Cynthia Lloyd, and Germán Rodríguez. Salman Zaidi generously provided analysis files for the 1991 Pakistan LSMS and corrected the expenditure data for regional variation in food prices. Contact information: Mark R. Montgomery, Policy Research Division, Population Council, One Dag Hammarskjold Plaza, New York, New York 10017; e-mail: mmontgomery@popcouncil.org.

ABSTRACT

Very few demographic surveys in developing countries gather information on household income or consumption expenditure, despite the theoretical importance of these measures. Consequently, researchers have been forced to rely on ad hoc collections of proxy measures for living standards, and the properties of these proxies have not been systematically analyzed. In this research, we ask what hypotheses can be tested using proxy variables, and evaluate the performance of proxy measures in relation to consumption expenditures per adult, our preferred measure of living standards. We find that the proxy variables commonly employed in demographic research are very weak predictors of consumption per adult, having partial R^2 values that are extremely low. Nevertheless, when other factors are taken into account, we show that tests based on proxy variables are likely to be sufficiently powerful to merit consideration. In an examination of fertility, child mortality, and children's schooling, we compare coefficient estimates based on consumption per adult to alternative estimates based on proxies, and find that the proxy-based estimates provide generally reliable guidance to the sign and magnitude of the preferred estimates.

Demographic theory has long distinguished the effect of income on behavior from that of education. A household's income summarizes its command over resources, including the resources that could promote health, lessen the need for high fertility, or reduce the opportunity costs of children's schooling. Education is closely linked to income; it is also thought to have a separable and distinctive influence on decisionmaking. Schooling can stimulate the development of cognitive abilities and heighten attention to information (LeVine et al. 1991); it can shift the distribution of authority within the household and equip individuals with the social confidence needed to claim extra-household resources (Caldwell 1979); and it can impart specific information that is pertinent to demographic decisions (Elo and Preston 1996). The conceptual distinctions between income and education have become a prominent feature of mortality analyses, figuring, for example, in the study of the early-twentieth-century experience in the United States (Preston and Haines 1991) and in much discussion of mortality in developing countries (Cleland and van Ginneken 1988; Stuebing 1997). Likewise, in the economic and sociological theory that supports fertility research, the separate roles of income and education are an important theme (e.g., LeVine et al. 1991; Schultz 1981).

Unfortunately, very few empirical studies in developing countries are able to do justice to this theme—in these countries, demographic surveys seldom gather data on household income. The collection of accurate income data is a difficult task; it must compete for survey resources against higher-priority modules on health, mortality, fertility, and children's schooling. Household consumption expenditures are preferred to measures of income on some theoretical grounds, and consumption data are somewhat easier to gather; nevertheless, to properly measure consumption is also a costly undertaking. In recognition of these difficulties, most demographic surveys have fallen back on a compromise design. The approach has been to collect a heterogeneous set of simple indicators in the hope that, when taken together, they will make good proxies for living standards. It is the aim of this paper to determine whether the indicators typically collected are, indeed, good proxies.

To evaluate the performance of such indicators, we require data sets that include both the indicators themselves and the economic variables that they are meant to represent, that is, household consumption expenditures or incomes. The

World Bank's Living Standards Measurement Study (LSMS) surveys provide this information. We use recent LSMS surveys from five countries that span the major regions and per capita income levels of the developing world: Ghana (1987–89), Jamaica (1989), Pakistan (1991), Peru (1994), and Tanzania (1993–94). These surveys are complemented by a 1995 survey from rural Guatemala, the Encuesta Guatemalteca de Salud Familiar (EGSF), which employed a simplified, lower-cost procedure for collecting household consumption data.¹ Of course, six countries do not make a fully general sample, but it is reasonable to suppose that this group encompasses much of the relevant variation.

The plan of the paper is as follows. In the first section, we review current practice in demographic research in the selection of proxy variables for living standards. Here, and throughout the paper, we limit attention to studies of developing countries. As will be seen, numerous approaches have been explored in the recent literature, and no “best practice” approach has yet emerged. Indeed, one finds surprisingly little discussion of the methodological issues, with the result that living standards measures seem to have been selected on an ad hoc, study-specific basis.

Our approach is not to examine all of the proxy variable specifications that have appeared in the literature, but rather to give systematic attention to a representative selection. Drawing upon the LSMS and EGSF surveys, we assemble a set of about a dozen variables that are commonly used as proxies for living standards. Each of these variables is available in the prototype Demographic and Health Survey (DHS) questionnaire, and each could therefore be considered for inclusion in a demographic analysis. Three specifications using these proxy variables are examined in the remainder of the paper.

The necessary statistical context is provided in the second section, where we take up the question of causal modeling with proxy variables. In a causal model, the use of proxies introduces a form of specification error. The consequences are not unlike those of measurement error (Fuller 1987; Griliches 1974), but proxy variables present new and distinct issues. This section delineates the theoretical criteria against which such proxies can be judged. We summarize findings from the literature that do not appear to be widely known or fully appreciated in demographic empirical work.

In the third section, we employ these criteria to evaluate the empirical association between the proxy variables and the living standards measure that we take them to represent. We compare the proxies to household consumption expenditures per adult, which is our preferred measure of living standards. The choice of consumption expenditures is defended in this section and the alternatives to it are discussed. As will be shown, even the best-performing proxy variables and indices are only weakly associated with consumption per adult. Nevertheless, we find that they can be effective in testing one important hypothesis about the influence of living standards on demographic behavior.

The fourth section returns to the larger theoretical question: What is the distinctive contribution made by education to demographic behavior, as compared to the contribution of consumption? To address the question, we present empirical models of three important demographic domains—lifetime fertility, child survival, and children’s schooling. First, we estimate models that incorporate consumption per adult and then compare the results to those from alternative specifications based on its proxies. The central issue in the comparison is how use of proxies affects the estimated contribution of mother’s education and the significance of consumption expenditures. In the final section, we set out the main findings of the research, addressing both methodology and the theoretical debate.

CURRENT RESEARCH PRACTICE

In Table A-1 (see Appendix A), we present a sample of recent empirical analyses for developing countries that employ measures of living standards. The entries refer to articles published in either *Demography* or *Population Studies* from 1990 to 1996. The DHS are prominently represented among data sources, but a number of other surveys also make an appearance. As can be seen in the last column of the table, only three of the analyses use a measure of income or consumption as such (Lloyd and Gage-Brandon 1994; Sastry 1996; Stewart et al. 1991). The diversity of alternative specifications is striking. To judge from this list, there would seem to be no generally accepted method of constructing an index of living standards from the ingredients available in demographic surveys.

If there is a common aspect to the specifications, it is the reliance on at least one of three sets of measures: access to water and the nature of toilet facilities,

indicators of housing quality, and ownership of selected consumer durables. The LSMS and EGSF surveys contain data on these variables comparable to those gathered in a generic DHS survey. Table 1 describes the index elements available in both the LSMS/EGSF and DHS surveys.²

Some researchers might object to the inclusion of electricity-dependent consumer items on the grounds that access to electricity is a community rather than a household characteristic (Knodel and Wongsith 1991), but, given the possibilities afforded by batteries, generators, and electrical line taps, we think it is reasonable to include them. After having considered whether to include measures of the occupation or education of the spouse, we have elected not to include these measures. In several of our sample countries—most notably Ghana, Jamaica, and Peru—there is considerable marriage dissolution, informal unions are important, and spouses need not co-reside. These demographic difficulties are not handled consistently in the LSMS surveys, and to restrict the analysis to currently married, spouse-present women would reduce the LSMS sample sizes and might bring about sample selection or endogeneity bias.

Table 2 indicates the percentages of households—strictly speaking, these are households with at least one woman in the age range 15–49—having the index items for the six sample countries. We will consider three standard of living indices (SLI) formed from these components. The first (SLI-1) is a simple summation of the number of items present. (The maximum will differ slightly from one country to the next.) A second measure (SLI-2) is specified with dummy variables for each distinct value of SLI-1, with the lowest value (or set of values) treated as the omitted category. This specification is useful in allowing the sum of the index items to exert a nonlinear influence. The third specification (SLI-3) treats each index item as a separate variable, thereby introducing a set of 10 to 12 dummy variables.

CAUSAL MODELING WITH PROXY VARIABLES

What are the implications of using such proxy measures in an otherwise well-specified causal model? In this section, we explore the issues with the aid of large-sample statistical theory and Monte Carlo experiments. The theoretical results are presented here without proof, but the interested reader can find the proofs in

Table 1 Components of a living standards index

	Ghana 1987-89	Guatemala 1995	Jamaica 1989	Pakistan 1991	Peru 1994	Tanzania 1993-94
Access to clean water	x	x	x	x	x	x
Water on premises	x	x	x	x	x	x
Time to water < 30 min	x	x	x	x	x	x
Toilet facility	x	x	x	x	x	x
Flush toilet	x	x	x	x	x	x
Nondirt flooring	x	x		x	x	x
Electricity	x	x	x	x	x	x
Radio	x	x	x	x	x	x
TV	x	x	x		x	x
Refrigerator	x	x	x		x	x
Bicycle	x	x	x	x	x	x
Motorcycle	x	x	x	x	x	
Car	x	x	x	x	x	

Table 2 Percentages of households with index items

	Ghana	Guatemala	Jamaica	Pakistan	Peru	Tanzania
Access to clean water	24.0	54.6	79.3	43.2	71.7	44.1
Water on premises	12.6	65.0	60.5	38.5	66.8	19.8
Time to water < 30 min	67.3	92.4	26.0	94.7	97.6	64.5
Toilet facility	72.5	16.4	97.6	67.4	60.9	95.8
Flush toilet	6.0	8.3	25.2	40.7	51.5	8.4
Nondirt flooring	77.7	39.9		51.6	60.6	50.6
Electricity	25.4	48.8	65.3	80.1	71.5	26.3
Radio	26.0	73.4	70.0	20.3	86.9	43.1
TV	7.6	33.4	51.9		72.9	3.2
Refrigerator	7.0	7.4	45.4		42.2	7.7
Bicycle	12.8	33.5	12.5	36.2	24.9	1.7
Motorcycle	0.6	2.1	2.0	10.6	3.4	
Car	3.0	7.9	8.2	2.9	6.8	

Appendix B. To bring the key issues to the forefront, we consider the case of linear regression.

Let the correctly specified causal model be given as in equation (1),

$$Y = X\beta + Z\delta + \epsilon, \quad (1)$$

where we have written the equation in matrix form assuming a sample of n observations. The dependent variable Y is an indicator of demographic behavior. The covariates X represent the measured influences on behavior, among which education would be included. The covariate Z is the household's standard of living, expressed in its level of consumption per adult. We will assume that both X and Z are statistically exogenous to the regression disturbance term ϵ , so that ordinary least squares applied to equation (1) would produce consistent estimates of the β , δ , and σ_ϵ^2 parameters as well as their standard errors.³

The causal model of equation (1) cannot be estimated as it stands, because no data on Z are available. However, the researcher has access to a proxy variable P , or to a set of $l > 1$ such proxies, that could be inserted in the equation in place of the unmeasured Z variable. (In the cases that we examine, P will be represented by one of the SLI indices.) What would be gained by such a substitution?

Large-sample Biases

In a situation such as this, misspecification of the causal model is all but inevitable. To ignore the proxies P would bring on one form of specification error, but an analogous form of error would be risked by including them. If the proxies are excluded and equation (2) is estimated,

$$Y = X\beta + u, \quad (2)$$

with $u = Z\delta + \epsilon$, the estimated coefficient $\hat{\beta}$ will diverge from the causal parameter β unless X and Z happen to be orthogonal.⁴ If the proxies P are included, we have, instead, equation (3),

$$Y = X\beta + Pd + v, \quad (3)$$

with $v = Z\delta - Pd + \epsilon$. Once again, $\hat{\beta}$ estimated from this equation will diverge from the true β parameter. Furthermore, the \hat{d} coefficients on the proxies P will

not converge to quantities that shed light on the true δ parameter, at least in the general case.

We suspect that when faced with such a dilemma, most researchers would follow intuition and include the proxies. Surely, most would reason, it must be better to control for consumption Z , however imperfectly, than to omit all measures of it. The proxies P presumably contain some information about Z , and one might at least hope to reduce the inconsistency of $\hat{\beta}$ by incorporating P in the model.

This line of thinking is appealing and can be shown to be correct in special cases (see Appendix B and Aigner 1974; Maddala 1977; Wickens 1972). Surprisingly, however, it is not correct in general. Proxy controls for Z have the potential to reduce inconsistency in $\hat{\beta}$, but there can be no guarantee of such reduction even in very large samples.

Testing Hypotheses About Consumption

When estimators are inconsistent, as \hat{d} is for δ , this usually invalidates hypothesis tests. For a certain kind of hypothesis, however, one can formulate a perfectly valid test by making use of the proxy variables. We refer to a test that is focused on the relevance of consumption, that is, a test for $\delta = 0$. The test procedure is no more than an ordinary χ^2 test applied to \hat{d} , the proxy variable coefficients, with degrees of freedom equal to l , the number of such proxy variables. (Another label for the χ^2 test is the Wald test. An F test would be equivalent to the χ^2 test in large samples.) Although one might hope to know more about the role of consumption than whether its coefficient is zero, even this information would make a contribution to the demographic debates.

At first glance, this result about testing might be thought curious and even disconcerting. The χ^2 test statistic is calculated from \hat{d} and from data on Y , X , and the P proxies. Because the consumption variable Z is missing, it cannot itself enter the calculation. Yet, the quality of the test must somehow depend on the nature of the relationship between Z and its proxies.

Consider an extreme case in which the proxies P are utterly uninformative about consumption. If P contains no information on Z , then a rejection of the null hypothesis using a test statistic based on P cannot be meaningful. A meaningful test for $\delta = 0$ should reject the null with high probability when the null is false,

and the likelihood of rejection should increase with the difference between the true δ and the $\delta = 0$ value specified by the null. As the case of uninformative proxies suggests, the strength of the relationship between Z and P must manifest itself in the power of the testing procedure. We show in Appendix B that the stronger the relationship between the proxies P and Z , the greater the power of the χ^2 test. The appropriate measure of association between Z and P is a partial R^2 . This is the R^2 taken from a regression of Z on X and P ; it reveals the degree to which the proxies P explain Z net of the X covariates.

In addition to the partial R^2 , a factor that comes into play only when proxy variables are employed, three other factors determine the power of the χ^2 test. These are the size of the δ coefficient relative to the standard error σ_ϵ , the empirical variability of consumption Z net of other covariates X , and the sample size. These three factors would determine test power even if Z were actually available, and, as shown in Appendix B, they continue to exert influence when Z is missing.

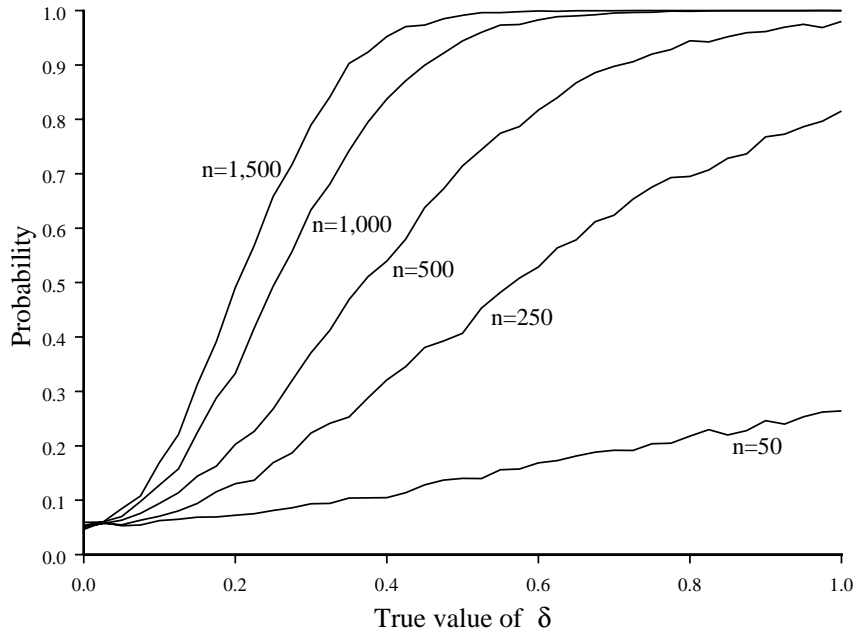
To sum up, for the narrow purpose of testing the hypothesis $\delta = 0$, it makes good sense to use proxy variables provided that they are arguably highly correlated, net of X , with the unobserved consumption variable. The test will have lower power than if Z itself were available, and the partial R^2 summarizes concisely the loss of information.

A Monte Carlo Illustration

An example may help to clarify these points. Consider the simple linear model $Y = Z\delta + \epsilon$, in which Z is a single explanatory variable and both Z and ϵ are standard normal. Suppose that a single proxy variable P is available; it is also standard normal but correlated with Z , with ρ being the correlation coefficient. Since there are no X variables in the model, the partial R^2 value is simply ρ^2 , the square of the raw correlation coefficient.

One can show that the regression coefficient \hat{d} , which is drawn from the misspecified regression $Y = Pd + v$, converges not to δ , but rather to $\rho\delta$ as sample size grows. The regression-based variance estimator s^2 , which should converge to $\sigma_\epsilon^2 = 1$, converges instead to $\delta^2(1 - \rho^2) + 1$. Whereas \hat{d} is a downwardly biased estimator of δ , the estimator s^2 is upwardly biased for σ_ϵ^2 .

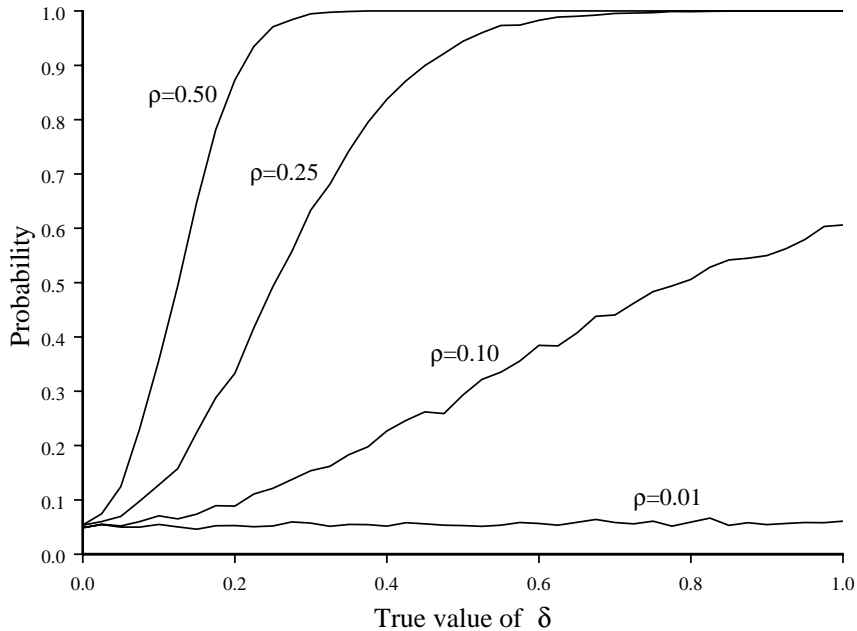
Figure 1 Probability of rejecting the null $\delta = 0$, given $\rho = 0.25$



In spite of these biases, the proxy variables estimator \hat{d} can be useful in testing the null hypothesis $\delta = 0$. To illustrate this, we present the results of a Monte Carlo experiment. Figure 1 depicts the probability of rejecting the null $\delta = 0$ for a range of values of δ and sample sizes.⁵

As would be expected, the probability of rejecting $\delta = 0$ rises with the true δ value for all sample sizes. (As discussed in Appendix B, it is actually the changing ratio of δ to σ_ϵ that generates the upward slope of these curves; here, $\sigma_\epsilon = 1$.) With the true value of δ held constant, the rejection probability increases with the sample size. As is evident in the figure, for sample sizes in the range typically seen in demographic applications ($n > 500$), the probability of rejecting the null is reasonably high, with a likelihood exceeding 0.7 where the value of δ exceeds one-half. Of course, the actual value of δ is unknown. As the figure shows, it is much harder to discriminate between the zero value for δ specified in the null hypothesis and non-zero but small values of the δ parameter. Even in the largest

Figure 2 Probability of rejecting the null $\delta = 0$, for sample size $n = 1,000$



samples considered here, the null is rejected less than 25 percent of the time for true values of δ that are less than one-tenth of σ_ϵ . In short, if consumption is thought to have an effect on behavior, but one expects the effect to be relatively small, then the χ^2 test will be unlikely to reject the hypothesis of no effect whatsoever.

As we emphasized above, these rejection probabilities depend on the value of the partial R^2 between Z and the proxy P , which is equal to ρ^2 in this instance. Figure 2 shows how the likelihood of rejection varies with ρ for a sample size of 1,000. If the correlation between P and Z is fairly high (the $\rho = 0.50$ case), so is the probability of rejecting the null. Weaker correlations are associated with decidedly lower probabilities of rejection; in the extreme (represented here by $\rho = 0.01$), the null is rejected only 5 percent of the time. In such extreme cases, the test's failure to reject the null provides no useful information about the true value of the δ parameter. Because \hat{d} converges to $\rho\delta$, this estimator is also a misleading guide to δ 's true value.

As we have seen in this example, failure to reject can reflect no more than the inadequacies of the proxy variables. To know what message the χ^2 test is likely to convey, we clearly must have some prior knowledge of the correlation between Z and these proxies.

Revisiting the Assumptions

The theoretical results discussed above have been secured with a minimum of assumptions about the relationship between Z and the proxy variables. One crucial assumption should be emphasized, however, because it bears on the distinction between the proxy variables and the other X covariates included in the causal model.

The distinguishing feature of a proxy variable is that, according to the theory that informs the causal model, it does not directly enter that model. Its role is simply to stand in for Z , the missing variable that belongs in the causal specification. If Z were available, the proxies P would have no role to play and P would therefore be excluded from the causal model.

Because such exclusion restrictions are justified mainly by theory, the distinction between the X and P variables will inevitably be controversial, a matter on which different researchers, equipped with different theories, might well disagree. Returning to the measures listed in Table 1, we note that several could play dual roles: They could be proxy measures for consumption and, at the same time, have their own direct causal influences on behavior. The access to water measures, for instance, could exert an important causal influence on child mortality. In a mortality analysis, such variables would be properly treated as belonging to the set of X covariates. In an analysis of children's schooling or fertility, however, theory might suggest that the water measures have no direct influence. Depending on the context, then, a given variable might be assigned either to the P or the X category.

When the data include the consumption variable Z , the researcher will have some freedom to test for the direct causal contribution of other variables. When Z is missing, however, it becomes much more difficult to determine whether to place a variable in the P or X category. An alternative approach to ours is to think of consumption Z as being an unobserved latent construct and to use multiple-indicator or related techniques to model its association with the observed indica-

tors (Bollen and Lennox 1991). There are costs to this alternative approach, in that it requires the researcher to make explicit assumptions about the relationship of the unobserved Z to the various indicators, including distributional assumptions. Theory can rarely supply the justification for such assumptions and they are not easily tested by statistical means. The task of evaluating such models is rendered more difficult by the fact that the indicators are discrete-valued and the latent construct is inherently unobservable. The potential benefit of the approach, however, is that in some cases the researcher can use both the X and P variables to model the latent construct. Although the issues are complicated, the latent variables approach may prove to be a useful complement to the proxy variables approach that we have pursued.

THE SLIS AND CONSUMPTION

We have referred at several points to our view that consumption expenditure per adult comes closest, among the available measures, to the concept of income in representing a household's command over resources. We now make clear what aspects of theory and measurement support this view. In poor countries, no single empirical measure can be expected to display all of the facets of the concept of income. Our judgment is that consumption per adult is the best measure among those collected in cross-sectional surveys. Nevertheless, we should not proceed without comment on its potential weaknesses and the empirical alternatives. The main lines of the argument are given below; see also Appendix C.

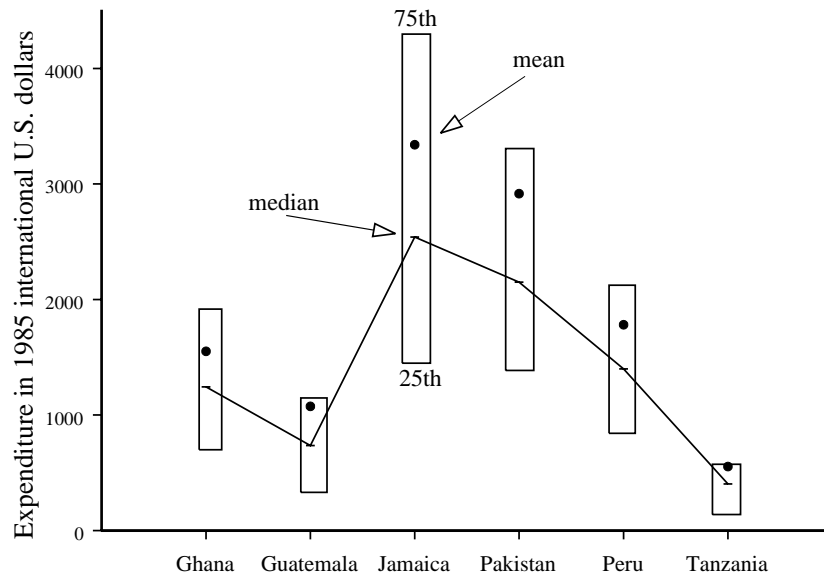
Why not use incomes as such? The reason is that in developing countries, households often draw their incomes from multiple sources that can change from year to year and even season to season. To properly measure income for a single year requires attention to the details of primary and secondary employment and the nature of payment for each adult household member. Transfers and income derived from other sources also need to be measured, as do the costs (in family farms or businesses) of generating income. The transitory nature of some employment, coupled with the uncertainty of net economic return, makes it implausible to think of any one year's income as being representative of the incomes earned over the longer time span in which demographic decisions are made.

Most households will have access to means of transferring resources across time periods, and will save and borrow through a variety of mechanisms that shield their consumption levels from the variabilities of income. When it is possible to borrow and save without incurring heavy transactions costs, then one can view income as having both transitory and permanent components, with consumption being closely related to the permanent component. The notion of frictionless borrowing and saving is, of course, something of a fiction in poor countries. Even where such possibilities are limited, however, consumption should be a “smoothed” version of a highly variable income stream and should better represent the conceptual ideal than would income as such. One could edge even closer to the conceptual ideal by gathering longitudinal data on incomes—but this is not feasible in a cross-sectional survey.

Properly designed surveys will measure the implicit value of consumption activities that do not pass through the market. For instance, the LSMS program has a well-developed protocol for assigning monetary value to farm goods that are produced and consumed by farm households, and these values are added to the total of consumption expenditures for the household. Also, the “services” implicitly provided by consumer durables are calculated and added to the total, using the dates of purchase of the durables, estimates of current value, and assumptions about depreciation. Hence, many of the consumer durables listed in the standard-of-living indices above (see Table 1) are already included in total consumption expenditures.

Why should total consumption expenditures, calculated in this way, then be divided by the number of adults in the household? The intention is to capture the command over resources wielded by the adults who make demographic decisions. Although a simple average for consumption is conceptually inferior to a measure of each adult’s resources, consumption data are not easily linked to particular household members. In demographic applications, the use of total household size as the denominator, rather than the number of adults, must be avoided. Such per capita measures would logically confound an explanatory variable (consumption per capita) with a component of the dependent variable (fertility, mortality). As Casterline (1988) and Casterline et al. (1989) have shown in a penetrating analysis of Egyptian mortality, the per capita specifications can produce seriously misleading results.

Figure 3 Expenditure per adult by country: 1985 international U.S. dollars



Empirical Associations

We now come to the central question—Do the SLI measures described above serve as good proxies for consumption expenditure per adult? Figure 3 shows box plots of consumption per adult in the six study countries.⁶ As can be seen, the sample countries vary considerably in the level and dispersion of consumption per adult. The line in the middle of each box represents the median value; above it is the mean. The lower and upper borders of each box indicate the 25th and 75th percentiles of the distributions. Tanzania is the poorest country in this sample and, indeed, is among the poorest countries in the world. Ghana, by comparison at least, is much better off, and in our sample Jamaica exhibits the highest consumption levels. There is considerable overlap across countries in these distributions, with the poorer Jamaicans being roughly on par with the richer Ghanaians.

Figures 4 and 5 depict the distribution of consumption per adult by country, according to the number of items in the SLI. Box-and-whisker plots depict the central

Figure 4 Consumption expenditures per adult and the standard of living index: Ghana, Guatemala, Pakistan, and Tanzania

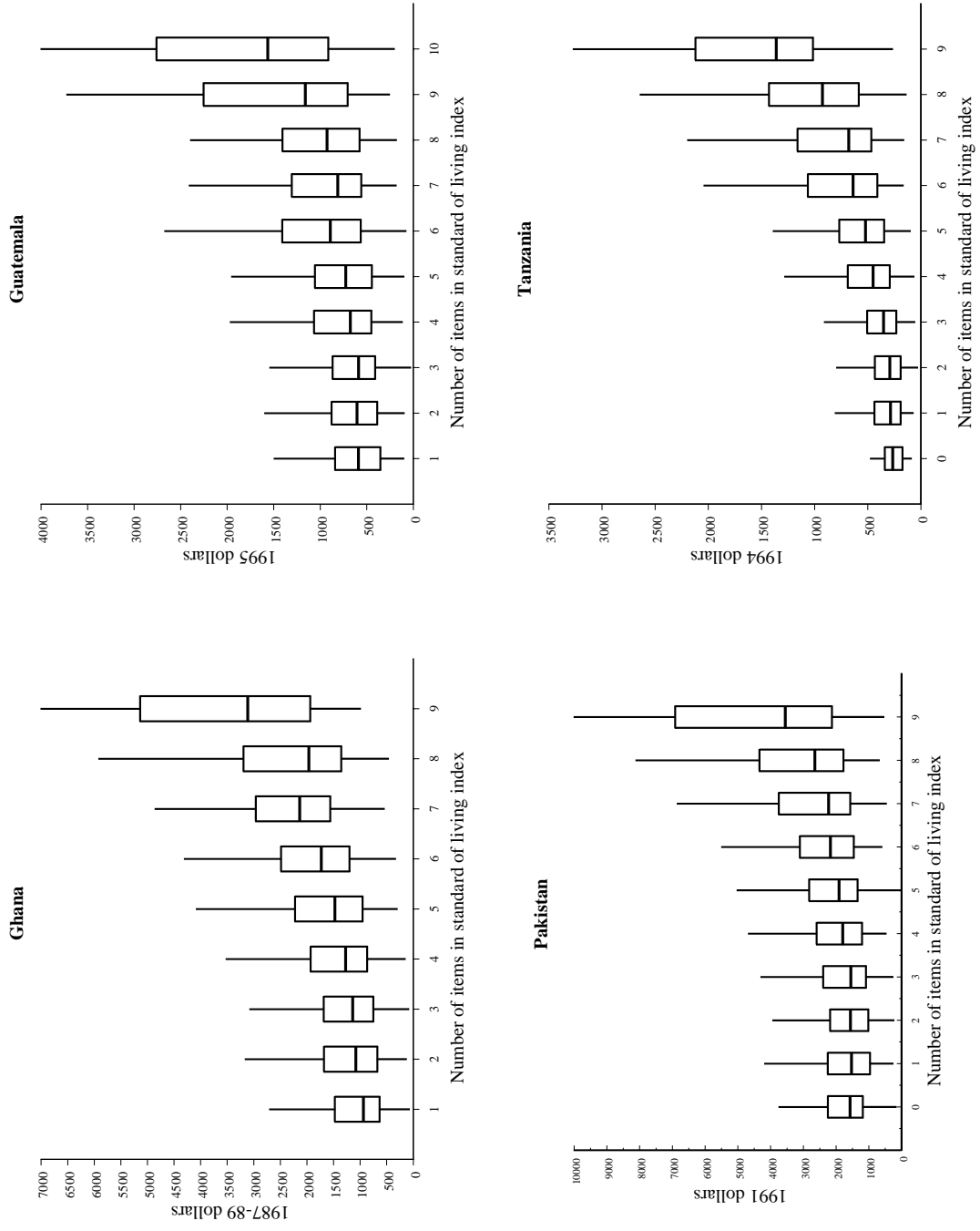
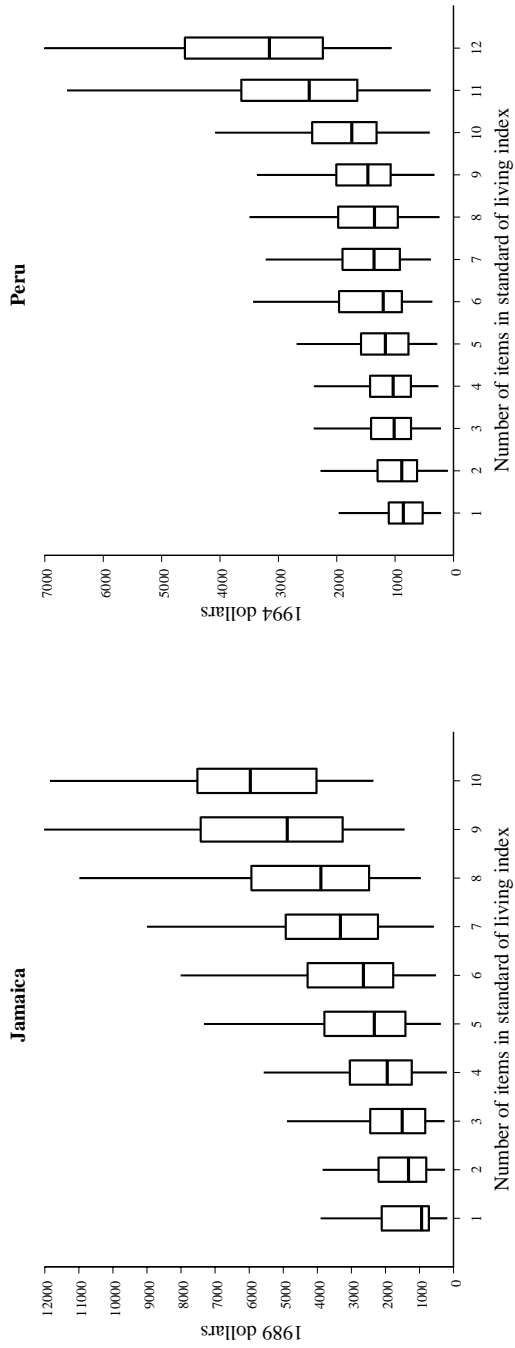


Figure 5 Consumption expenditures per adult and the standard of living index: Jamaica and Peru



tendencies and ranges. Although some irregularities are apparent in these figures, they show that the SLIs do contain information about the level of consumption per adult. The median values of consumption tend to increase with the number of items in the index, as do the 25th and 75th percentiles, apart from a few exceptions. Similar figures (not shown) are produced if we plot separately the relationships for urban and rural areas within each country. As a general rule, the urban consumption levels are higher than the rural at each value of the SLI; but within both urban and rural areas, the index is positively associated with consumption.

These central tendencies are encouraging, but to judge the quality of the SLI proxies, we need to know the partial R^2 values, which measure the strength of association between consumption per adult and the SLI proxies with the joint effects of other covariates removed. The empirical variability of consumption Z net of other X covariates, a factor that we denote by $\tilde{\sigma}_{Z|X}^2$, is also an important consideration. So, too, is the sample size.

Table 3 presents the key items of information. To provide context, the top panel of the table gives the ordinary R^2 values taken from a regression of consumption expenditures per adult on the various SLI measures. We show the R^2 values for the full samples and separately for the urban and rural subsamples of each country. These “raw” R^2 values are alarmingly low. They indicate that although the proxy indices may contain some information about consumption they do not contain very much information. Of the three indices, the one with the greatest explanatory power is SLI-3, which is the dummy variable specification with dummies for each of the specific index items.

The picture is not improved when we consider the theoretically appropriate partial R^2 values shown in the second panel of the table. (The other covariates employed include the woman’s education, her age and powers of age, and a set of dummies for regions. We discuss these covariates in the fourth section.) Here, the R^2 values drop as low as .015 for Pakistan, and, even at their best, attain a level of only .150 for Jamaica. If taken by themselves, these partial R^2 s would indicate that proxy-based tests for the relevance of consumption might have distressingly low power.

Fortunately, the power of the χ^2 test is not decided by the partial R^2 values alone. It also depends on the variability of consumption expenditures per adult.

Table 3 Regressions of expenditure per adult on standard of living indices

	Ghana	Guatemala	Jamaica	Pakistan	Peru	Tanzania
	R^2 values					
SLI-1	.104	.077	.143	.030	.154	.155
Rural, urban	.014, .082		.106, .094	.025, .036	.132, .108	.017, .114
SLI-2	.129	.103	.154	.033	.220	.185
Rural, urban	.020, .123		.135, .101	.033, .038	.163, .180	.031, .140
SLI-3	.156	.178	.224	.077	.226	.215
Rural, urban	.041, .136		.177, .192	.047, .141	.192, .184	.079, .170
	Partial R^2 values net of other covariates					
SLI-1	.030	.047	.078	.015	.063	.072
SLI-2	.064	.064	.090	.017	.124	.109
SLI-3	.072	.136	.150	.053	.130	.137
	$\hat{\sigma}_{Z X}^2$ values					
	131.4	162.2	811.8	1,323.7	191.7	38.8

This is the factor denoted by $\tilde{\sigma}_{Z|X}^2$, shown in the bottom panel of Table 3. We see that consumption is highly variable net of other covariates, and such variation will enhance the test’s ability to detect departures from the null hypothesis.⁷ Test power is further enhanced by large sample sizes. To roughly gauge the power of the χ^2 test, one can multiply the partial R^2 value by both $\tilde{\sigma}_{Z|X}^2$ and the sample size. If the product exceeds 20 (see Appendix B), then one can reasonably expect the χ^2 test statistic to have acceptable power.⁸ Even in Tanzania, which displays the lowest value for $\tilde{\sigma}_{Z|X}^2$ in Table 3, this conclusion will hold for samples of the size normally used by demographers.

In summary, the message delivered by Table 3 is ultimately reassuring. Although the SLI proxies are extremely weak when judged by their partial R^2 values, demographic sample sizes are usually large enough, and consumption per adult would seem to be variable enough, for tests based on the SLI proxies to be worth considering. The χ^2 test can be expected to reject the null hypothesis—that consumption does not matter—with reasonably high probability when that null is false. To be sure, the test is not as powerful as a test based on consumption itself, this being the consequence of the low partial R^2 values, and certainly one would not want to further consider the SLI specifications with the lower partial R^2 s. In what follows, then, we retain the SLI-3 measure, which is the best-performing of a poorly performing group, and assess its role in models of fertility, child mortality, and children’s schooling.

APPLICATIONS TO FERTILITY, CHILD MORTALITY, AND CHILDREN’S SCHOOLING

In this section, we compare estimates based on consumption expenditures per adult to alternatives based on SLI-3. For the fertility models, we specify a base set of explanatory variables, which include woman’s education, her age, age squared and age cubed, urban residence, and as detailed a set of regional indices as are available in the LSMS/EGSF data. The approach is generally similar for child mortality and schooling, although here the base set includes measures of the child’s birth order, birth year, and the mother’s age at birth. We then add to this base set the consumption per adult variable and test its significance. In the alternative spec-

ification, we add the SLI-3 dummies and test for their joint statistical significance using the χ^2 test.

In preliminary analyses (not shown here), we found that the effects of per adult expenditure were systematically different in rural and urban areas.⁹ The expenditure variables are therefore entered directly and as an interaction with an urban dummy variable. To accomplish this in the SLI specifications, we have interacted the total number of SLI items (SLI-1) with an urban dummy.

The cumulative fertility models are estimated using the ordered-probit technique (Greene 1997) with children ever born as the dependent variable. We include all women of reproductive age (ages 15–49) in the analysis. As for child mortality, the surveys in three countries—Ghana, Guatemala, and Pakistan—include birth history data for the woman that allow mortality risks to be modeled on a child-by-child basis. We use Cox proportional hazards models for these analyses. The same three countries provide child-by-child information on the years of schooling completed by the woman’s children, which we model using the ordered-probit method.

Table 4 presents the sample means for selected covariates. Here and in what follows, we limit attention to measures of consumption expenditures and women’s education, as the contrast between their effects is perhaps of greatest theoretical interest. Note that the structure of education varies considerably across the study countries. In each case, we have taken the omitted category to be the lowest category, but in Ghana, for example, this is no schooling whereas in Jamaica, it is less than secondary schooling.

Fertility

In Table 5, the ordered-probit estimates of fertility are summarized, with the main results being as follows. In rural areas, expenditure per adult is either insignificant or is positively associated with fertility. By contrast, in urban areas—here one must add the two expenditure coefficients to see the effect—higher expenditures generally either reduce fertility or have no net effect, with Peru being the exception. One can see the significance of the urban factor in the p values of the χ^2 test on urban expenditures. Turning to the women’s education results, we

Table 4 Selected sample means

	Ghana	Guatemala	Jamaica	Pakistan	Peru	Tanzania
	Samples of women					
Children ever born	3.24	2.71	2.06	3.16	2.22	2.87
Woman's age	28.46	25.5	27.13	27.81	29.04	27.45
Urban	.350		.484	.509	.679	.684
Schooling						
Primary	.132	.587		.104		
Middle	.206			.058		.110
Secondary	.167	.091	.687	.163	.438	
Post-secondary			.143		.194	
	Samples of children					
Child mortality (_{5/0})	.170	.087		.164		
Mean years of schooling	3.57	1.37		2.64		

find that, apart from the case of Ghana, the coefficients on education are uniformly negative, significant, and large in magnitude.

The coefficients presented in Table 5 are not readily interpretable in terms of demographic magnitudes. To assess their substantive implications, we have calculated predicted values of children ever born, with the woman's age set at 40 years, for different levels of expenditure per adult in both rural and urban areas. The predictions are evaluated across the range of expenditures from the 10th to the 90th percentile. We perform similar calculations for each distinct level of women's education.

In examining these predicted levels of lifetime fertility (not shown), we find that although expenditures per adult generally have a statistically significant influence on fertility, their demographic effect is comparatively small. The implied differences in fertility between the 10th and 90th percentiles of expenditure never exceed 0.4 children in either the negative or the positive direction. Perhaps effects of such magnitude should not be dismissed, but they would not appear to be of decisive importance.

When applied to women's education, the same kind of analysis reveals a strikingly different picture. Apart from Ghana, in which the woman's education coefficients are insignificant, inspection of predicted fertility shows that the education effects are clearly of substantive importance. For example, in Jamaica the predicted mean fertility of a woman with post-secondary schooling is 2.82 children, whereas the prediction for women with less than secondary schooling (the omitted category) is 4.28 children. Even larger differentials are evident in Peru, where women without secondary schooling are predicted to have 5.12 children, those with secondary schooling some 3.74 children, and for women with post-secondary schooling, 2.29 children. Important education effects also characterize Guatemala, Pakistan, and Tanzania.¹⁰ These education differentials are all the more striking as the inclusion of consumption per adult controls for the association between education and the standard of living.

Table 6 presents estimates of the fertility models in which the SLI proxies are employed in place of consumption per adult. Because the coefficients on the individual index items confound two associations—that between the index item and consumption per adult, and between consumption per adult and fertility—it is

Table 5 Models of children ever born: Selected ordered-probit coefficients

	Ghana	Guatemala	Jamaica	Pakistan	Peru	Tanzania
Number of women	4,291	2,816	2,153	5,193	4,558	6,742
Expenditure per adult coefficients						
Expenditure per adult (00's)	.040·10 ⁻¹	.063·10 ⁻¹	-.002·10 ⁻¹	.014·10 ⁻¹	.055·10 ⁻¹	.033
(Z statistic)	(1.85)	(3.94)	(0.18)	(2.94)	(1.80)	(3.96)
Expenditure-urban	-.087·10 ⁻¹		-.033·10 ⁻¹	-.023·10 ⁻¹	-.026·10 ⁻¹	-.035
	(3.11)		(2.08)	(2.84)	(0.79)	(4.10)
Urban	-.242		-.018	.122	-.211	-.225
	(4.18)		(0.25)	(3.23)	(3.32)	(5.17)
χ^2 tests on expenditure coefficients: <i>p</i> values						
Both equal zero	.006	.000	.008	.005	.020	.000
Urban equals zero	.010		.002	.167	.027	.316
Women's education coefficients						
Primary school	.097	-.322		-.252		-.139
	(1.95)	(7.13)		(4.88)		(3.96)
Middle school	.007			-.359		
	(0.16)			(4.71)		
Secondary school	-.061	-1.018	-.137	-.702	-.619	-.850
	(1.31)	(11.80)	(2.07)	(11.56)	(14.27)	(14.07)
Post-secondary school			-.659		-1.418	
			(7.40)		(24.26)	
χ^2 test on education coefficients: <i>p</i> values						
All equal zero	.070	.000	.000	.000	.000	.000

Table 6 Comparing fertility models using proxies and expenditures

	Ghana	Guatemala	Jamaica	Pakistan	Peru	Tanzania
χ^2 test on proxies: <i>p</i> values						
All equal zero	.000	.000	.000	.011	.000	.000
Women's education coefficients based on proxies						
Primary school	.081	-.251		-.285		-.111
Within confidence band	yes	yes		yes		yes
around Table 5 estimate						
Middle school	.010			-.353		
Within confidence band	yes			yes		
Secondary school	-.037	-.751	-.126	-.685	-.493	-.576
Within confidence band	yes	no	yes	yes	no	no
Post-secondary school			-.562		-1.193	
Within confidence band			yes		no	
χ^2 test on education coefficients: <i>p</i> values						
All equal zero	.253	.000	.000	.000	.000	.000

not obvious that the index coefficients merit discussion. (Some of these items show a reasonably consistent association with fertility across countries—for example, water on the premises, electricity, possession of a refrigerator—but others do not.) We have argued that the principal role of these proxies is to enable χ^2 tests of the null hypothesis that consumption is irrelevant. As can be seen in Table 6, the p values on the χ^2 test reveal the index items to be highly significant as a group. Using these proxies, we would decisively reject the null hypothesis that the true consumption coefficient is zero.¹¹

Table 6 also shows that the women’s education coefficients are not much affected, whether in sign, significance, or magnitude, by the use of the proxy SLI variables. For the most part, the estimated education effects either fall within the 95 percent confidence bands surrounding the preferred estimates from Table 5 or lie near these bands. At least for the countries studied here, one would not be misled about the importance of women’s education to fertility if given only the results based on proxies.

We have repeated this analysis using measures of recent fertility in the sample countries for which this is possible. The results for both expenditure and schooling are qualitatively similar to those shown above, although statistical significance is diminished.

Child Mortality

Three of the data sets provide usable information on child mortality. Employing the Cox regression method, we have estimated proportional hazards models of mortality risk in the first five years of life. In addition to the explanatory covariates used in the fertility analyses, we consider the mother’s age at birth (under 20 years or over 40 years), the child’s sex, and dummies for first births and high parity (> 6) births, as well as a year of birth variable to capture trend.

Tables 7 and 8 give the results, with the focus again on women’s education, the consumption per adult variable, and its SLI proxies. Table 7 shows that consumption per adult has a statistically significant negative effect on mortality in urban and rural areas in Ghana, and has a marginally significant effect in Guatemala. (The estimate for urban Pakistan implausibly suggests a positive effect.) As in the fertility analyses, mother’s education is estimated to have little influence in Ghana,

Table 7 Models of child mortality: Selected Cox regression coefficients

	Ghana	Guatemala	Pakistan
Number of children	12,869	7,338	17,355
Expenditure per adult coefficients			
Expenditure per adult (00's)	$-.064 \cdot 10^{-1}$	$-.088 \cdot 10^{-1}$	$.002 \cdot 10^{-1}$
(Z statistic)	(2.22)	(1.93)	(0.38)
Expenditure·urban	$-.042 \cdot 10^{-1}$		$.018 \cdot 10^{-1}$
	(0.86)		(1.69)
Urban	-.136		-.067
	(1.40)		(1.16)
χ^2 tests on expenditure coefficients: <i>p</i> values			
Both equal zero	.003	.054	.113
Urban equals zero	.008		.040
Women's education coefficients			
Primary school	.053	-.264	-.226
	(0.78)	(2.95)	(2.77)
Middle school	.046		-.513
	(0.76)		(3.68)
Secondary school	.000	-.962	-.463
	(0.00)	(2.79)	(3.92)
χ^2 tests on education coefficients: <i>p</i> values			
All equal zero	.786	.001	.000
Access to safe water coefficients			
Access to clean water	.128	.252	.050
	(1.57)	(2.21)	(0.58)
Water on premises	-.370	-.382	-.102
	(2.91)	(3.24)	(1.17)
Time to water < 30 min	-.050	.495	.163
	(0.80)	(2.40)	(1.64)
Toilet facility	-.023	-.064	.058
	(0.38)	(0.35)	(1.01)
Flush toilet	-.230	-.174	-.359
	(1.30)	(0.61)	(6.27)
χ^2 tests on water coefficients: <i>p</i> values			
All equal zero	.018	.006	.000

but it is significantly associated with reduced mortality risks in both Guatemala and Pakistan. Where magnitudes are concerned, the implications are similar to those of the fertility analyses: The consumption variable alters the predicted value of ${}_5q_0$ by no more than 1–2 percentage points over the 10th to 90th percentiles of consumption, whereas changes in women’s education have an effect of some 6 percentage points in Guatemala and Pakistan.

In assessing the role of proxy variables in these mortality models, we have removed the water and toilet measures from the proxies category and assigned them a potentially causal role. The proxy-based tests for the relevance of consumption are therefore based only on the remaining SLI variables. These χ^2 tests are shown in Table 8. For Guatemala and Pakistan, the test rejects the hypothesis that consumption per adult is irrelevant, although it fails to reject this hypothesis in Ghana.¹² Once again, the estimates of women’s education effects in the proxy variables model fall within the 95 percent confidence bands surrounding the preferred Table 7 estimates. The same is true of the estimates of the water and toilet coefficients. As in the fertility analyses, here the proxy variable estimates seem to provide reliable guidance to the preferred estimates based on consumption.

Children’s Schooling

Among the determinants of children’s years of schooling, both consumption expenditure and women’s education make highly significant contributions. As can be seen in Table 9, in the otherwise diverse settings of urban Ghana, rural Guatemala, and urban and rural Pakistan, the consumption coefficient is positive and statistically significant; the coefficient for rural Ghana, although insignificant by the conventional criterion, is also positive. Likewise, the coefficients associated with women’s education are positive and highly significant, with Ghana again a partial exception. To assess the magnitude of these effects, we have calculated predicted years of schooling for a child of age 18, a prediction that should approximate completed schooling, and have then examined how such predictions vary with consumption percentiles and the level of the woman’s education. As in the cases of fertility and mortality, we find larger effects associated with the woman’s education than with consumption expenditures. The predicted values for children’s schooling, evaluated at the 10th and 90th percentiles of consumption, differ by no

Table 8 Comparing mortality models using proxies and expenditures

	Ghana	Guatemala	Pakistan
χ^2 test on proxies: <i>p</i> values			
All equal zero	.140	.090	.000
Women's education coefficients based on proxies			
Primary school	.053	-.217	-.190
Within confidence band around Table 7 estimate	yes	yes	yes
Middle school	.046		-.419
Within confidence band	yes		yes
Secondary school	.000	-.836	-.289
Within confidence band	yes	yes	yes
χ^2 test on education coefficients: <i>p</i> values			
All equal zero	.853	.010	.001
Water coefficients based on proxies			
Clean water	.147	.203	.035
Within confidence band	yes	yes	yes
Water on premises	-.316	-.337	-.079
Within confidence band	yes	yes	yes
Time to water	-.070	.488	.152
Within confidence band	yes	yes	yes
Toilet	-.027	-.057	.038
Within confidence band	yes	yes	yes
Flush toilet	-.202	-.148	-.296
Within confidence band	yes	yes	yes
χ^2 test on water coefficients: <i>p</i> values			
All equal zero	.120	.026	.000

Table 9 Models of children's schooling: Selected ordered-probit coefficients

	Ghana	Guatemala	Pakistan
Number of children	6,697	2,795	12,498
Expenditure per adult coefficients			
Expenditure per adult (00's)	$.029 \cdot 10^{-1}$	$.077 \cdot 10^{-1}$	$.010 \cdot 10^{-1}$
(Z statistic)	(1.64)	(5.47)	(4.05)
Expenditure-urban	$.016 \cdot 10^{-1}$		$.047 \cdot 10^{-1}$
	(0.71)		(8.54)
Urban	.242		.127
	(5.14)		(4.30)
χ^2 tests on expenditure coefficients: <i>p</i> values			
Both equal zero	.002	.000	.000
Urban equals zero	.002		.000
Women's education coefficients			
Primary school	.002	.592	2.478
	(0.04)	(11.38)	(68.18)
Middle school	.089		3.930
	(2.42)		(78.21)
Secondary school	.135	1.356	5.217
	(3.54)	(10.23)	(88.78)
χ^2 tests on education coefficients: <i>p</i> values			
All equal zero	.001	.000	.000

more than 0.4 years of schooling. For women's education, however, a similar comparison suggests differences ranging from a low of 0.4 years in Ghana to nearly 10 years in Pakistan. The differences are not just important at the extremes: They are also substantial at intermediate levels of women's education.

A comparison of schooling estimates based on the SLI proxies for consumption is shown in Table 10. The null hypothesis that consumption is irrelevant is decisively rejected by the proxy-based χ^2 test, a result that agrees with what was shown in Table 9. The estimates of the women's education coefficients reveal that, in Ghana, the proxy-based estimates lie within the confidence bands surrounding

Table 10 Comparing schooling models using proxies and expenditures

	Ghana	Guatemala	Pakistan
χ^2 test on proxies: <i>p</i> values			
All equal zero	.000	.000	.000
Women's education coefficients based on proxies			
Primary school	.003	.412	2.288
Within confidence band around Table 9 estimate	yes	no	no
Middle school	.091		3.677
Within confidence band	yes		no
Secondary school	.102	.762	4.921
Within confidence band	yes	no	no
χ^2 test on education coefficients: <i>p</i> values			
All equal zero	.011	.000	.000

the preferred estimates. In Guatemala and Pakistan, however, these estimates are smaller than the preferred Table 9 estimates and lie below the confidence bands. This might be taken as evidence of severe large-sample bias in the proxy estimators, but the qualitative conclusions one would draw from Table 10 are much the same as those from Table 9. In either case, a compelling argument is made for the importance of the woman's education to the educational attainment of her children.

CONCLUSIONS

On the whole, the results of this research should offer some encouragement to demographers, who have had little alternative but to rely on proxy variables for their measures of household living standards. We find that the SLI proxies are very weak predictors of consumption per adult, having partial R^2 values that are extremely low. But when enlisted mainly to test whether consumption is relevant to behavior, the SLI proxies are rescued by two other factors. First, there is considerable variability in consumption expenditures per adult, so that even weak proxies for consumption are able to detect departures from the null hypothesis. Second, demographers are fortunate in having access to relatively large samples, and sam-

ple size further enhances the power of the proxy-based tests. With the aid of these two factors, even proxies as weak as these indices can provide useful information. They furnish the basis for tests of the relevance of consumption that should have acceptable power in most demographic applications.

In illustrating the theory through models of lifetime fertility, child mortality, and children's schooling, we have found evidence that consumption per adult has a statistically significant effect on demographic behavior. Tests for the relevance of consumption generally reject the null hypothesis of no effect, whether based on proxies for consumption or on the variable itself. For the reasons outlined previously, the proxy variables coefficients cannot reveal the magnitude of such effects, but the consumption coefficients suggest that the demographic impacts are rather small. If the effects are as small in general as our estimates indicate, the null hypothesis of no effect will often fail to be rejected, especially when that hypothesis is tested with weak proxies.

In contrast to the consumption results, in all countries save Ghana, women's education seems to exert a substantial influence on fertility, child mortality risks, and children's schooling. Because the effects of consumption per adult are controlled in these models, the strength of the education effects is remarkable. The case of Ghana, evidently unusual, bears further inspection. It may be that the quality of schooling in Ghana has been low, with the result that women's education is not associated as tightly with demographic behavior as it is in other countries (see Glewwe 1999). One would be hard-pressed to say, however, that school quality is uniformly lower in Ghana than in other poor countries, such as Tanzania and Pakistan.

As was discussed, another potential benefit from using proxy variables is to reduce the inconsistency in the $\hat{\beta}$ estimators attached to women's education and other covariates. In our limited sample of countries, we find little to suggest that the effects of women's education will be badly mis-estimated if proxy variables are used in place of the preferred consumption measure.

To be sure, this conclusion and the conclusion about the relative magnitudes of the education and consumption effects are based on a small sample of country experience and simple behavioral models of fertility, child mortality, and children's schooling. We hesitate to suggest that the results are more generally applicable,

and perhaps a wider sample or more detailed empirical investigation would overturn some of our conclusions. In particular, we would welcome a comparison of competing perspectives on the measurement of living standards. Although we favor viewing SLIs as proxies for consumption per adult, there is room in the debate for alternative views. A rigorous comparison to latent variables models could prove especially instructive.

Our main substantive result will likely withstand such scrutiny. Our estimates provide striking evidence of the varied roles that women's education plays in demographic behavior, roles that appear to be separable from education's direct links to living standards. If we have controlled adequately for living standards by using consumption per adult, the estimates suggest a decisive contribution of education in affecting cognitive abilities, attention to and receipt of information, social confidence, and women's decisionmaking autonomy. Our models do not reveal which of these factors makes the difference—perhaps all do—but they underscore their centrality to demographic behavior.

**APPENDIX A: INVENTORY OF RECENT DEMOGRAPHIC STUDIES
USING LIVING STANDARDS MEASURES**

Table A-1 Research on developing countries using measures of living standards and published in *Demography* or *Population Studies*, 1990–96

Author(s)	Title	Data source	Living standards measure
Eric Jensen	“The fertility impact of alternative family planning distribution channels in Indonesia,” <i>Demography</i> (1996)	Indonesia 1991 DHS	Index of durables, quality of housing
Narayan Sastry	“Community characteristics, individual and household attributes, and child survival in Brazil,” <i>Demography</i> (1996)	Pesquisa Nacional sobre Saude Materno-Infantil e Planejamento Familiar—Brasil. 1987 DHS	Total household income
Anne Pebley, Noreen Goldman, Germán Rodríguez	“Prenatal and delivery care and childhood immunization in Guatemala: Do family and community matter?” <i>Demography</i> (1996)	National Survey of Maternal and Child Health in Guatemala in 1987	Occupation, woman’s and spouse’s education, whether woman has ever worked, presence of modern toilet facility in household
Pradip Muhuri	“Estimating seasonality effects in child mortality in Matlab, Bangladesh,” <i>Demography</i> (1996)	Matlab Demographic Surveillance System (DSS) and 1982 Matlab Census	Receives by way of remittances or owns a radio, bicycle, watch, quilt, or lamp; dwelling space per capita; mother has no schooling; child born in Matlab treatment area

(continued)

(Table A-1 continued)

Author(s)	Title	Data source	Living standards measure
Ilene Speizer	"A marriage trichotomy and its applications," <i>Demography</i> (1995)	Cameroon 1991 DHS	Type of flooring; ownership of durables: radio, bicycle, television, electricity, motorcycle, car, stove, refrigerator
Ulla Larsen	"Differentials in infertility in Cameroon and Nigeria," <i>Population Studies</i> (1995)	World Fertility Surveys and DHS for Cameroon (1971, 1991) and Nigeria (1981–82, 1990)	Polygynous union (many wives indicating higher economic status), woman's and spouse's education, occupation
Akinrinola Bankole	"Desired fertility and fertility behaviour among the Yoruba of Nigeria: A study of couple preferences and subsequent fertility," <i>Population Studies</i> (1995)	Surveys by Department of Demography and Social Statistics, Obafemi Awolowo University, Nigeria 1984, 1986	Spouse's occupation: agricultural, nonagricultural, not working
Kofi Benefo	"The determinants of the duration of postpartum sexual abstinence in West Africa: A multilevel analysis," <i>Demography</i> (1995)	World Fertility Survey for Côte d'Ivoire 1978, Ghana 1979, Cameroon 1978	Mother's education, nonfamily work, urban residence, spouse's nonmanual occupation
Kenneth Bollen, David Guilkey, Thomas Mroz	"Binary outcomes and endogenous explanatory variables: Tests and solutions with an application to the demand for contraceptive use in Tunisia," <i>Demography</i> (1995)	Tunisia 1988 DHS	Sum of household assets

(continued)

(Table A-1 continued)

Author(s)	Title	Data source	Living standards measure
Renata Forste	“The effects of breastfeeding and birth spacing on infant and child mortality in Bolivia,” <i>Population Studies</i> (1994)	Bolivia 1989 DHS	Partner’s education, mother’s employment
Kerry Richter, Chai Podhisita, Aphichat Chamratrithirong, Kusol Soonthornhdada	“The impact of child care on fertility in urban Thailand,” <i>Demography</i> (1994)	Child Care, Women’s Status, and Fertility in Urban Thailand, 1990–91	Spouse’s occupation
Deborah Balk	“Individual and community aspects of women’s status and fertility in rural Bangladesh,” <i>Population Studies</i> (1994)	Survey collected by the MCH-FP Extension Project of the International Centre for Diarrhoeal Disease Research, Bangladesh, 1988	Respondent’s secular education, dwelling size, landholdings, landlessness, husband a daily laborer
Guang Guo, Laurence Grunmer-Strawn	“Child mortality among twins in less developed countries,” <i>Population Studies</i> (1993)	Demographic and Health Surveys for 26 developing countries	Urban/rural residence, mother’s and spouse’s education, husband’s occupation (agricultural, manual), owns car or TV

(continued)

(Table A-1 continued)

Author(s)	Title	Data source	Living standards measure
Cynthia Lloyd, Anastasia Gage-Brandon	“High fertility and schooling in Ghana: Sex differences in parental contributions and educational outcomes,” <i>Population Studies</i> (1994)	Ghana 1987–88 L.SMS	Consumption per household head
Andrew Foster	“Household partition in rural Bangladesh,” <i>Population Studies</i> (1993)	Matlab Demographic Surveillance System; census data 1974–82	Number of items owned: radios, watches, quilts, lamps; number of cows owned; whether household receives remittances
Linda Adair, Barry Popkin, David Guilkey	“The duration of breast-feeding: How is it affected by biological, sociodemographic, health sector, and food industry factors?” <i>Demography</i> (1993)	Cebu Longitudinal Health and Nutrition Survey 1983–84	Ownership of refrigerator, gas or kerosene stove; piped water; mother worked in modern wage sector during pregnancy; spouse present in household
John Stewart, Barry Popkin, David Guilkey, John Akin, Linda Adair, Wilhelm Flieger	“Influences on the extent of breast-feeding: A prospective study in the Philippines,” <i>Demography</i> (1991)	Authors’ survey of metropolitan Cebu, Philippines	Family’s real income in top quartile

(continued)

(Table A-1 continued)

Author(s)	Title	Data source	Living standards measure
Deborah DeGraff	“Increasing contraceptive use in Bangladesh: The role of demand and supply factors,” <i>Demography</i> (1991)	1982 Matlab Socioeconomic Survey; 1984 In-depth Contraceptive Knowledge, Attitudes, and Practice Survey	Owens land, owns boat, spouse’s occupation, composition of dwelling walls, access to clean drinking water
Wamucii Njogu	“Trends and determinants of contraceptive use in Kenya,” <i>Demography</i> (1991)	Kenya Fertility Survey 1977–78 (WFS)	Region of residence
John Knodel, Malinee Wongsith	“Family size and children’s education in Thailand: Evidence from a national sample,” <i>Demography</i> (1991)	Thailand 1987 DHS	Wealth index derived from: types of vehicles owned, type of flooring, type of toilet facility
David Hamill, Amy Tsui, Shyam Thapa	“Determinants of contraceptive switching behavior in rural Sri Lanka,” <i>Demography</i> (1990)	Rural Family Planning Survey 1985–86	Type of housing and household water source
Martin Brockerhoff	“Rural-to-urban migration and child survival in Senegal,” <i>Demography</i> (1990)	Senegal 1986 DHS	Region; mother’s marital status, education, work status, occupation; spouse’s occupation; toilet facilities, source of drinking water

(continued)

(Table A-1 continued)

Author(s)	Title	Data source	Living standards measure
Michael Koenig, James Phillips, Oona Campbell, Stan D'Souza	"Birth intervals and childhood mortality in rural Bangladesh," <i>Demography</i> (1990)	Matlab Demographic Surveillance System	Household area, mother's education
Abdur Razzaque, Nurul Alam, Lokky Wai, Andrew Foster	"Sustained effects of the 1974-5 famine on infant and child mortality in a rural area of Bangladesh," <i>Population Studies</i> (1990)	Matlab Demographic Surveillance System	Additive index of articles owned or received as remittances: quilts, hurricane lamps, radios, watches, cash remittances

SOURCE: See references for citations

APPENDIX B: DERIVATION OF STATISTICAL RESULTS

The Linear Case

As in the text, we consider the linear regression model

$$Y = X\beta + Z\delta + \epsilon \quad (\text{B-1})$$

in which the dependent variable Y is of dimension $n \times 1$, the X matrix of explanatory variables is $n \times k$, and the $n \times 1$ vector Z represents the preferred measure of consumption expenditures. (The results discussed here can be generalized to the case of multiple Z variables.) The $n \times 1$ disturbance vector ϵ is assumed to have mean zero and covariance matrix $\sigma_\epsilon^2 I$. We make a standard assumption about the relationship between ϵ , X , and Z , namely, that ϵ is weakly exogenous to both X and Z . By “weakly exogenous,” we mean that $\text{plim } n^{-1} X' \epsilon = 0$ and $\text{plim } n^{-1} Z' \epsilon = 0$, where the term “plim” should be read to mean “probability limit.” If the consumption measure Z were available, these and mild additional assumptions would guarantee the consistency of $\hat{\beta}$ and $\hat{\delta}$, the ordinary least-squares estimators.

In the absence of Z , we entertain the possibility of using, in its place, a proxy variable P or a set of $l > 1$ such proxies. (We maintain the assumption that, like X and Z , the proxies P are weakly exogenous to the disturbance ϵ .) By inserting P in place of Z , we would estimate the misspecified equation

$$Y = X\beta + Pd + v, \quad (\text{B-2})$$

with d being $l \times 1$, where the composite disturbance $v = \epsilon - Pd + Z\delta$. What are the large-sample consequences of such a specification error?

Estimates of β

Using the FWL theorem (Davidson and MacKinnon 1993) or partitioned inversion, we obtain the following expression for $\hat{\beta}$,

$$\hat{\beta} = \beta + (X' M_P X)^{-1} X' M_P Z \delta + (X' M_P X)^{-1} X' M_P \epsilon, \quad (\text{B-3})$$

with $M_P = I - P(P'P)^{-1}P'$. As P and X are weakly exogenous to ϵ , it follows that

$$\hat{\beta} \xrightarrow{P} \beta + \text{plim}(n^{-1} X' M_P X)^{-1} \text{plim}(n^{-1} X' M_P Z) \delta, \quad (\text{B-4})$$

where the notation “ \xrightarrow{P} ” means “converges in probability,” this being a synonym for the term “plim.” The important point is that $\hat{\beta}$ is inconsistent for β . That is, the use of proxies has damaging consequences for the other coefficients of the model.

Estimates of δ and σ_ϵ^2

Applying the same techniques to the estimator of the proxy variables coefficients \hat{d} , we obtain

$$\hat{d} = (P' M_X P)^{-1} P' M_X Z \delta + (P' M_X P)^{-1} P' M_X \epsilon, \quad (\text{B-5})$$

where we have used the fact that $M_X X = 0$. In the limit,

$$\hat{d} \xrightarrow{P} \text{plim} (n^{-1} P' M_X P)^{-1} \text{plim} (n^{-1} P' M_X Z) \delta. \quad (\text{B-6})$$

Note that \hat{d} and δ may be of different dimensions, this occurring when a set of $l > 1$ proxies is employed for the single unobserved variable Z . In such a case, no direct comparison of \hat{d} and δ would be sensible. Even if \hat{d} is of dimension one, however, equation (B-6) shows that \hat{d} is generally inconsistent for δ .

It can also be shown that the usual estimator of the disturbance term variance, s^2 , converges to a quantity that exceeds the true variance σ_ϵ^2 . That is,

$$s^2 \xrightarrow{P} \text{plim} n^{-1} \delta' Z' M Z \delta + \sigma_\epsilon^2, \quad (\text{B-7})$$

in which the matrix M is analogous to M_P and M_X above but includes both X and P .

Reducing the Inconsistency of $\hat{\beta}$

By adapting the approach of Wickens (1972), we can assess whether using the proxies P will tend to reduce the inconsistency of the $\hat{\beta}$ coefficients on X , drawing the estimates closer to the true value of β than they would be if P were simply omitted. The direction of effect depends on the nature of the auxiliary structural model that links Z to P . Consider the case of a single P variable. If one can write the auxiliary model as $P = Z\theta + w_1$, with w_1 being uncorrelated with Z , then using the proxy will reduce inconsistency in $\hat{\beta}$. The key assumption here is that w_1 is uncorrelated with Z . A safer and more general model would be $P = X\gamma + Z\theta + w_2$. In this more general case, it is not obvious that including the proxy will reduce the inconsistency of $\hat{\beta}$. The γ coefficients of the auxiliary structural model will now cloud the comparison, leaving the net effect ambiguous.

Testing Hypotheses About Consumption

The large-sample biases noted above will invalidate tests of hypotheses, at least in the general case. One special case, however, can be rigorously assessed. Under the null hypothesis that the consumption coefficient $\delta = 0$, the estimator \hat{d} reduces to

$$\hat{d} = (P' M_X P)^{-1} P' M_X \epsilon,$$

and it is then easy to establish that $\hat{d} \xrightarrow{p} 0$. On making the additional central limit assumption that $n^{-1/2} P' M_X \epsilon \xrightarrow{d} \mathcal{N}(0, \sigma_\epsilon^2 Q)$, where the notation “ \xrightarrow{d} ” means “convergence in distribution,” we have

$$n^{1/2} \hat{d} \xrightarrow{d} \mathcal{N}(0, \sigma_\epsilon^2 Q^{-1})$$

with $Q = \text{plim } n^{-1} P' M_X P$. This last result provides the ingredients for a test statistic T ,

$$T = \frac{\hat{d}' P' M_X P \hat{d}}{s^2} \xrightarrow{d} \chi_l^2 \quad (\text{B-8})$$

in which l , the number of degrees of freedom, is the number of proxy variables employed in the test. We use here the fact that if a vector W is multivariate normal with mean 0 and positive definite covariance matrix $\sigma^2 V$, then $W' V^{-1} W / \sigma^2$ is central χ^2 with degrees of freedom equal to the dimension of W . We also use the fact that, under the null hypothesis, s^2 converges to σ_ϵ^2 . The test described in equation (B-8) is no more than a simple χ^2 test applied to \hat{d} , and most statistical packages will implement this procedure.

Test Power

Can we rely on such a test to reject the null hypothesis $\delta = 0$ when that null hypothesis is false? This is the central question, for if rejection could not be reasonably well assured, the test would have little diagnostic value. The power of the test—that is, the probability of rejecting the null when the null is false—depends on the relationship between the proxy variables P and the unobserved Z measure. MacKinnon (1992) and Davidson and MacKinnon (1993) present analyses of the power function. The discussion below is adapted from MacKinnon (1992).

Suppose that the regression disturbance ϵ is normally distributed.¹³ Then, given that

$$\hat{d} = (P' M_X P)^{-1} P' M_X Z \delta + (P' M_X P)^{-1} P' M_X \epsilon,$$

the estimator \hat{d} is normally distributed (conditional on X , Z , and P) with a mean of $(P'M_X P)^{-1}P'M_X Z\delta$ and a variance of $\sigma_\epsilon^2(P'M_X P)^{-1}$.

Returning to the test statistic of equation (B-8), imagine that the term in the denominator were σ_ϵ^2 rather than s^2 . With $\delta \neq 0$, the quadratic form defining T would now be distributed as noncentral χ_l^2 with noncentrality parameter λ , where

$$\lambda = \delta'Z'M_X P(P'M_X P)^{-1}P'M_X Z\delta/\sigma_\epsilon^2.$$

This result is due to the fact that if a random vector W is multivariate normal with mean μ and covariance matrix σ^2V , the quadratic form $W'V^{-1}W/\sigma^2$ is distributed as noncentral χ^2 with noncentrality parameter $\lambda = \mu'V^{-1}\mu/\sigma^2$.

Of course, the test statistic T is actually defined in terms of s^2 rather than the true σ_ϵ^2 . Moreover, when $\delta \neq 0$, the estimator s^2 does not converge to σ_ϵ^2 , as was noted above in equation (B-7). Hence, for given δ , the test statistic T will converge to a random variable that is proportional to a noncentral χ_l^2 variate. This detail does not affect the essence of the argument that follows.

With other things equal, the greater the noncentrality parameter λ , the more powerful the test. The relationship between λ and test power is depicted in Figure B-1, which shows the probability of rejection of the null hypothesis for a conventional 5 percent test. Note that as λ increases, so does the power of the test. Note, too, that with other things equal, test power declines as a function of l , the number of proxy variables.

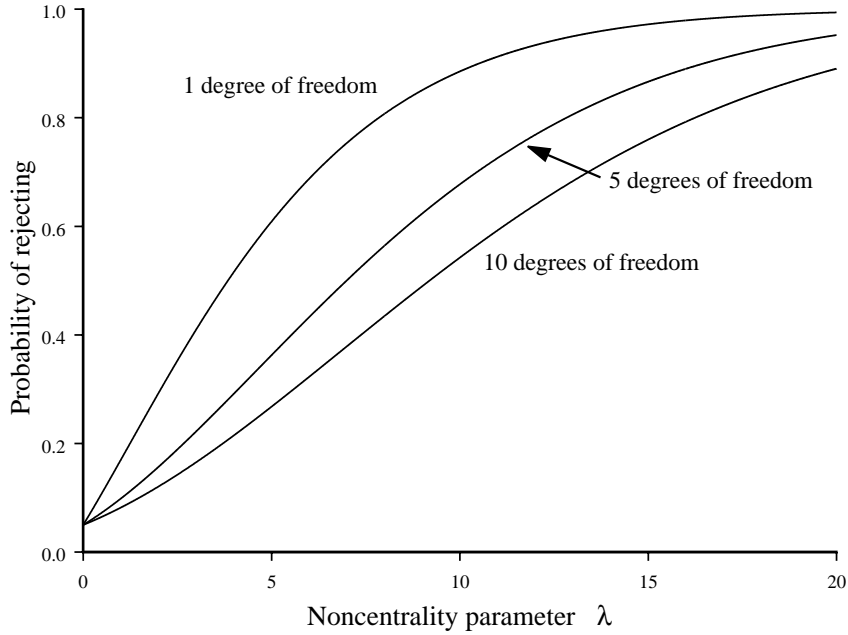
Quality of Proxy Variables

We have arrived at a means of judging the quality of the proxy variables. We will show that the higher the partial R^2 between Z and the set of proxies P , the greater the noncentrality parameter λ and, by extension, the greater the power of the test. Although the partial R^2 is of principal interest in what follows, it is only one of four factors that together determine λ .

To see the role of the proxies in their proper context, imagine that the preferred consumption measure Z were actually available. In this instance, λ would be the product of the following three factors:

$$\lambda = \left(\frac{\delta}{\sigma_\epsilon}\right)^2 \cdot \bar{\sigma}_{Z|X}^2 \cdot n. \quad (\text{B-9})$$

Figure B-1 Association of noncentrality parameter λ to test power



Here, δ is the true value of the consumption parameter and it appears in the expression scaled by σ_ϵ , the standard deviation of the regression disturbance term. We see that the greater the value of the δ parameter relative to the “noisiness” of the regression disturbance, the greater the probability of rejecting the null hypothesis. The second factor appearing in the expression, $\tilde{\sigma}_{Z|X}^2$, is not a model parameter as such, but rather is an empirical indicator of the variability of Z net of X :

$$\tilde{\sigma}_{Z|X}^2 = \frac{1}{n} Z' M_X Z.$$

The net variability of Z is relevant because it affects the ability of the test statistic to detect departures from the null hypothesis $\delta = 0$. The greater such independent variation in Z , the easier it becomes to assemble evidence against the null. The third factor in the expression is the sample size n . In short, if the consumption measure Z were actually available, the power of a test for $\delta = 0$ would depend on

the size of δ in relation to σ_ϵ , the variability of the Z net of other covariates X , and the sample size.

When the test is based on proxies P rather than the true Z , a fourth factor comes into play: the partial R^2 between Z and P , which measures their correlation after Z and P are purged of association with X . With P used in place of Z , the noncentrality parameter λ becomes

$$\lambda = \left(\frac{\delta}{\sigma_\epsilon} \right)^2 \cdot \tilde{\sigma}_{Z|X}^2 \cdot n \cdot R^2. \quad (\text{B-10})$$

The new factor is the uncentered R^2 from an artificial or diagnostic regression:

$$M_X Z = M_X P c + \text{residuals}. \quad (\text{B-11})$$

The uncentered R^2 from this regression is

$$R^2 = \frac{Z' M_X P (P' M_X P)^{-1} P' M_X Z}{Z' M_X Z}. \quad (\text{B-12})$$

Note that in regression (B-11), both Z and P are premultiplied by the matrix M_X . The transformation is equivalent to regressing each of the variables in P and Z on the common set of X variables and saving the residuals, which then become the explanatory and dependent variables for regression (B-11). The R^2 from the regression thus summarizes the ability of the proxies P to explain the empirical variation in Z after the effects of X have been taken into account. The higher the R^2 from this regression, the larger λ and the more powerful the χ^2 test, other things being equal.

Applications to Qualitative Variables Models

Although one expects the spirit of these results to carry over to nonlinear and qualitative variables models, the details are not easily reproduced in such settings. The problem in using proxy variables P for Z is that of estimating a misspecified maximum-likelihood model. We adapt below the approach of Kiefer and Skoog (1984) for probit and ordered-probit models.¹⁴

In these models, the notional dependent variable is Y^* , a latent variable, and the structural equation of interest is expressed as $Y^* = X\beta + Z\delta + \epsilon$. In the simple probit case, we observe $Y = 1$ if $Y^* > 0$, and observe $Y = 0$ otherwise. In the

ordered-probit case, we observe $Y = 1$ if $Y^* \leq \alpha_1$, $Y = 2$ if $\alpha_1 < Y^* \leq \alpha_2$, and so on. In each case, ϵ is assumed to be distributed as standard normal.

Let the relationship between income Z and its proxies P be represented in the simple form

$$Z = P\theta + u$$

and insert this into the latent variables structural equation, giving

$$\begin{aligned} Y^* &= X\beta + P\theta\delta + u\delta + \epsilon \\ &= X\beta + P\gamma + u\delta + \epsilon \end{aligned} \tag{B-13}$$

with $\gamma \equiv \theta\delta$. Imagine that the “measurement error” u is somehow made available to the researcher and can be employed as an ordinary covariate. Imagine further that the researcher uses u without recognizing the relationship between γ and δ and treats these as wholly distinct parameters. With u observed, the parameters β , γ , and δ of equation (B-13) can be consistently estimated by the usual probit methods. This artificial case supplies us with a benchmark for the next step of the analysis.

Rather than estimate equation (B-13) with u in hand, we are required to estimate equation (B-14),

$$Y^* = X\beta + P\gamma + \text{residuals} \tag{B-14}$$

in which the residuals are $u\delta + \epsilon$. Comparing equations (B-13) and (B-14), we see that the estimation problem has several aspects. First, the consumption coefficient δ is hidden in the l composite parameters $\gamma = \theta\delta$ that are attached to the proxy variables. It cannot be separately identified. Second, the measurement error u , a component of the disturbance term in equation (B-14), will be correlated with P and perhaps with X as well. If $\delta \neq 0$, this correlation invalidates ordinary probit and ordered-probit estimation. Third, u is unlikely to be normally distributed, causing the disturbance distribution to be misspecified and again rendering invalid the usual probit estimators.

Using the approach of Kiefer and Skoog (1984), the net effects of these problems can be seen in a local analysis of the inconsistency of $\hat{\beta}$ and $\hat{\gamma}$. The nature of the inconsistency can be clarified if we begin with the artificial benchmark equation (B-13), in which u is included as if it were a covariate and all parameters

can be consistently estimated. Equation (B-13) has parameters β , γ , and δ , and associated with them is an information matrix \mathcal{J} .

If we omit u and estimate equation (B-14) rather than (B-13), the inconsistency in $\hat{\beta}$ and $\hat{\gamma}$ can be determined by a Taylor expansion of the scores of the true model, that is, the scores associated with equation (B-13). Kiefer and Skoog (1984) show that to a local approximation (i.e., for small δ),

$$\text{plim} \begin{pmatrix} \hat{\beta} - \beta \\ \hat{\gamma} - \gamma \end{pmatrix} = \mathcal{J}_{11}^{-1} \mathcal{J}_{1\delta} \delta$$

where

$$\mathcal{J}_{11} = \begin{pmatrix} \mathcal{J}_{\beta\beta} & \mathcal{J}_{\beta\gamma} \\ \mathcal{J}_{\gamma\beta} & \mathcal{J}_{\gamma\gamma} \end{pmatrix}$$

and

$$\mathcal{J}_{1\delta} = \begin{pmatrix} \mathcal{J}_{\beta\delta} \\ \mathcal{J}_{\gamma\delta} \end{pmatrix}.$$

On inspection, the large-sample bias is seen to be similar in character to omitted-variables bias in the case of linear regression. As in the linear case, the bias from omitting a variable depends on the cross-products of the omitted with the included variables; but here the relevant terms are the off-diagonal elements of the information matrix \mathcal{J} of the true model, that is, the model of benchmark equation (B-13).

We will not further explore here the large-sample biases and related test statistics, because an analysis of the issues becomes extremely complex. The reader is referred to the instructive dissection of the probit model by Kiefer and Skoog (1984) and to the wider-ranging findings of Yatchew and Griliches (1984). Stefanski and Carroll (1985) and Carroll et al. (1995) have investigated the logit case.

APPENDIX C: WHY CONSUMPTION PER ADULT?

Surveys that emphasize the assessment of living standards often collect a great variety of measures of income. To decide on the appropriate measure is no simple matter (Ravallion 1996a, 1996b). The decision involves consideration of theoretical concepts in light of the local economic structure, as well as judgments about the probable extent and nature of measurement error.

In most LSMS surveys, an effort is made to gather information on total household incomes, and the data are then adjusted to account for the income derived from major economic activities that are not marketed. For instance, a detailed accounting is made of the food grown and consumed by farm households. Such food has an implicit income value, in the sense that it could have been taken to market. Calculating the value of marketed and own consumption entails listing the quantities of production in detail and attaching to each the appropriate market price. The prices themselves are often gathered in a separate community-level or market-level data collection exercise. In the LSMS household questionnaire, disaggregated data are also collected on additional sources of nonfarm income, including business income, labor earnings, and income from assets. Transfer income derived from remittances is also measured. The data are then further adjusted by netting out the various costs that households must bear to generate their incomes, such as the costs of inputs in the case of farm or business households. Incomes calculated in this way are not infrequently found to be negative, a situation that indicates a transitory shortfall of the household's revenues in relation to its costs.

In poor rural economies, it is especially difficult to accurately measure all components of household net income. In recognition of the difficulties, LSMS surveys complement the data on income with measures of household consumption expenditures. Hentschel and Lanjouw (1996: 1) describe one aspect of the rationale:

While poor households are probably purchasing and consuming only a narrow range of goods and services, their total income may derive from a myriad of different activities with strong seasonal variation and with associated costs that are not always easily assigned. Getting an accurate net income figure for such households can be frustratingly difficult.

As with income, the consumption expenditures must be adjusted to incorporate items that are not priced in the market, such as the value of housing services derived from owner-occupied housing. Consumer durables are likewise converted to flows of services by making use of detailed data on the vintages of the items owned coupled with assumptions regarding their rates of depreciation.¹⁵

Although the errors are arguably smaller and less systematic than those associated with household income, consumption expenditure data are not themselves error-free (see Bouis and Haddad 1992; Bouis 1994; Grosh et al. 1995; Scott and Amenuvegbe 1990). However, the rationale favoring the use of consumption expenditures as against income per se is not limited to the issue of measurement error. Consumption is likelier than income to represent the household's medium-term, and possibly its longer-term, command over resources. In theory, if a household can both borrow and save, its yearly consumption is a more reliable index of the standard of living than is one year's flow of income. Consumption can be regarded as a "smoothed" representation of a series of possibly noisy and highly variable annual incomes; in the extreme case of perfect capital markets, one can speak of consumption as reflecting "permanent income." Of course, this rationale is not wholly persuasive for poor economies in which households have limited abilities to save and borrow. Nevertheless, when coupled to considerations of relative measurement error, the weight of the argument suggests that consumption expenditures are generally to be preferred.

Adjusting for Household Size and Composition

The LSMS surveys estimate consumption expenditures at the level of the household. Because households can differ greatly in their size and composition, some further adjustment would seem to be in order to estimate the command over resources typically enjoyed by household members. Yet, it is far from obvious how the adjustment should be made—that is, whether consumption should be expressed as per capita, per adult, per adult equivalent, or in some other terms—and there is little consensus in the literature (see Hentschel and Lanjouw 1996).

When the data are to be employed in analyses of fertility and mortality, however, certain specifications must be ruled out. It is clearly inappropriate to specify a model of children ever born in terms of household consumption per capita, as

the denominator of the consumption variable reflects fertility. For the same reason, mortality analyses should not be specified in terms of consumption per capita or per adult equivalent. The risks of doing so are vividly illustrated by the research of Casterline (1988) and Casterline et al. (1989) for Egypt, which shows how the effect of income on child mortality can be distorted by a naive use of per capita measures.

On balance, we conclude that an appropriate measure of living standards is given by consumption expenditures per adult. In some circumstances one might attempt further refinements, such as accounting for economies of scale in households with different numbers of adults, but we have not pursued these here.

Variation in Prices

Because the consumption per adult measures depend on prices, it is vital to take into account any areal or regional variation in prices. Otherwise, differences in consumption levels might not reflect living standards as such, but rather differences in the costs of living. This is clearly an important issue in comparing rural and urban living standards, but it also deserves consideration in comparisons within these sectors.

Among the countries studied in our analysis, regional price indices have been devised for Ghana, Jamaica, and Peru (World Bank 1993; 1996a, 1996b), and we have adjusted the Pakistan data for regional variation in food prices (World Bank 1995a). Estimated price variation is considerable in Peru, somewhat less so in Pakistan, small in Jamaica, and evidently trivial in Ghana. It is not clear that regional price indices exist for Tanzania for the period in 1993–94 during which its survey was fielded. For Tanzania, therefore, the analyses on which we report should be treated with caution. The Guatemala consumption data, which are limited to four rural departments, are not believed to be much affected by cross-department price differences.

Exogeneity

In the text, we made the assumption that the true income measure Z and the proxy variables P were statistically exogenous to ϵ , that is, uncorrelated with the disturbance term of the equation. This is a strong and potentially objectionable assumption.

The difficulty is that in many developing countries, older children make direct or indirect contributions to household income, whether through working on the family farm or business, or by sending remittances as young adults. In analyses of fertility, therefore, one might anticipate a statistical correlation between a variable such as consumption expenditures per adult and the disturbance term of the fertility equation. A similar argument can be made with respect to mortality and children's schooling.

If such a correlation is feared, the usual statistical prescription is to estimate the model by the method of instrumental variables and then to test the null hypothesis of no correlation using Hausman's test (Hausman 1978). We have not pursued this approach here for four reasons. First, we believe that the potential for statistical endogeneity is present to much the same degree in both the consumption expenditure measures and the proxy variable indices. Second, the identification of appropriate instruments is always conceptually difficult, and the recent econometric literature has emphasized the dangers to proper inference of using weak instruments, such as those with low predictive power. Third, unless the SLI-style indices are collapsed into one or two summary measures—a procedure that we would resist on other grounds—many more instruments will be required to estimate models with indices than would be required for equivalent models using consumption expenditures. Fourth, in nonlinear models with measurement error, probit models being one example, standard instrumental variables techniques cannot be applied, as shown by Hausman et al. (1995) among others. Interesting variations on the standard techniques have been devised, and these merit investigation in future research.

NOTES

¹Tanzania's Human Resources and Development Survey, is not, strictly speaking, an LSMS survey, but was conducted using the same protocols. See Grosh and Glewwe (1996) for a summary of the LSMS program and detailed discussion of the generic LSMS questionnaire. Peterson et al. (1997) describe the 1995 Guatemalan survey (EGSF) and Gragnolati (1999) evaluates the consumption module.

²With respect to items having to do with water and toilet facilities, there are logical linkages among some index items (e.g., water available on the premises implies that water is less than 30 minutes away). LSMS surveys generally obtain distance to water sources. We have converted these distance measures to what we think are reasonable time equivalents so as to obtain a variable comparable to that of the DHS.

³We do not consider here the implications of clustered or multilevel samples, heteroskedasticity, and other statistical complications. For the linear model, an extension of the large-sample results to these cases is straightforward. The exogeneity issue is addressed in Appendix C.

⁴By "diverge," we mean that $\hat{\beta}$ is inconsistent for β , the parameter of the causal model specified in equation (1). See Appendix B for definitions of terms and more precise statements of conditions and conclusions. Note that orthogonality of X and Z does not guarantee that the standard errors of $\hat{\beta}$ will be consistently estimated.

⁵We use the conventional 5 percent rejection criterion and assume a value of $\rho = 0.25$ for these simulations. The points shown in the figure are the average values of 5,000 replications. The figure is symmetric about zero, and attention can therefore be restricted to nonnegative δ values.

⁶In generating this figure, we have converted all local currency amounts to their equivalents in international U.S. dollars in the survey year. The conversion involves the use of indices of purchasing power parity (PPP) taken from the most recent Penn World Tables of version Mark 5.6, an update of the data described in Summers and Heston (1991). For Tanzania, no recent estimate of PPP is available, and we have used exchange rates to convert the Tanzanian shilling to dollars, a procedure that underestimates the standard of living. For Peru, the most recent PPP measure available in the Penn World Tables is for 1992, and given the extent of macroeconomic turmoil from 1992 to 1994, the Peruvian estimates may also be artificially low. These conversion difficulties affect Figure 3 and distort

cross-country comparisons, but do no harm to our within-country analyses. The conversions simply scale the expenditure variables.

⁷See Appendix B for a definition of $\hat{\sigma}_{Z|X}^2$. The expenditure data are scaled here and in what follows by 100.

⁸Two issues need further consideration. First, the ratio of the true coefficient on expenditures δ to σ_ϵ will affect the power of the test. When δ is expected to be very small in relation to σ_ϵ , this will imply a need for larger samples and a highly variable consumption measure if test power is to reach acceptable levels. Second, because the χ^2 test statistic is based on s^2 rather than the true σ_ϵ^2 , it will not be distributed as noncentral χ^2 but rather as a random variable that is proportional to such a noncentral χ^2 variate; see Appendix B. The factor of proportionality is less than unity, causing the test statistic to be smaller than the noncentral χ^2 variate and further reducing the likelihood of rejecting the null hypothesis.

⁹The definition of urban differs across countries. For example, in Jamaica the urban areas are Kingston Metropolitan Area and “other towns.” In Peru an urban area is one that has a population of at least 2,000 inhabitants.

¹⁰In light of these differentials, the case of Ghana appears unusual indeed. Further analysis shows that within urban areas in Ghana, secondary education has an important negative influence on fertility, and middle school education has a significant, but not large, negative effect (results not shown).

¹¹The interaction of an urban dummy variable and SLI-1 is included with the other proxies in this test, but excluding it does not affect the conclusion.

¹²In the case of Pakistan, the test based on proxies for consumption decisively rejects the null, whereas the Table 7 analysis using consumption indicates that consumption is insignificant or, at best, marginally significant. According to the statistical theory outlined above, one would expect that proxy-based tests would reject less often than the counterpart test based on the true consumption variable. Of course, because test statistics are random variables, this is not guaranteed to occur in any given case.

¹³If ϵ is not normal, then we can prove what follows asymptotically by using the device known as “Pitman drift.” See Montgomery and Burke (1997) for the details.

¹⁴See Wolfe et al. (1996) for a closely related application and Yatchew and Griliches (1984) for alternative approaches that invoke additional structural assumptions.

¹⁵The EGSF survey for Guatemala differs from the LSMS surveys in that it used a simplified protocol for collecting consumption expenditures, restricting attention mainly to staples, other foods, and other selected items, and including the estimated value of consumption from own production. No income data were gathered. The simplifications adopted by the EGSF were appropriate given the study's focus on four rural departments, but might not be justifiable in more heterogeneous or urban settings. See Gragnolati (1999) for an analysis of expenditure data quality.

REFERENCES

- Adair, Linda S., Barry M. Popkin, and David K. Guilkey. 1993. "The duration of breast-feeding: How is it affected by biological, sociodemographic, health sector, and food industry factors," *Demography* 30(1): 63–80.
- Aigner, Dennis J. 1974. "MSE dominance of least squares with errors-of-observation," *Journal of Econometrics* 2(4): 365–372.
- Balk, Deborah. 1994. "Individual and community aspects of women's status and fertility in rural Bangladesh," *Population Studies* 48(1): 21–45.
- Bankole, Akinrinola. 1995. "Desired fertility and fertility behaviour among the Yoruba of Nigeria: A study of couple preferences and subsequent fertility," *Population Studies* 49(2): 317–328.
- Benefo, Kofi D. 1995. "The determinants of the duration of postpartum sexual abstinence in West Africa: A multilevel analysis," *Demography* 32(2): 139–157.
- Bollen, Kenneth A., David K. Guilkey, and Thomas A. Mroz. 1995. "Binary outcomes and endogenous explanatory variables: Tests and solutions with an application to the demand for contraceptive use in Tunisia," *Demography* 32(1): 111–131.
- Bollen, Kenneth A. and Richard Lennox. 1991. "Conventional wisdom on measurement: A structural equation perspective," *Psychological Bulletin* 110(2): 305–314.
- Bouis, Howarth E. 1994. "The effect of income on demand for food in poor countries: Are our food consumption databases giving us reliable estimates?" *Journal of Development Economics* 44(1): 199–226.
- Bouis, Howarth E. and Lawrence J. Haddad. 1992. "Are estimates of calorie-income elasticities too high? A recalibration of the plausible range," *Journal of Development Economics* 39(2): 333–364.
- Brockerhoff, Martin. 1990. "Rural-to-urban migration and child survival in Senegal," *Demography* 27(4): 601–616.

- Caldwell, John C. 1979. "Education as a factor in mortality decline: An examination of Nigerian data," *Population Studies* 33(3): 395–413.
- Carroll, Raymond J., D. Ruppert, and Leonard A. Stefanski. 1995. *Measurement Error in Nonlinear Models*, London: Chapman and Hall.
- Casterline, John B. 1988. "Household income and survival in Egypt: Sensitivity of estimates to measurement of income," Unpublished paper, Brown University Population Studies and Training Center, Providence, RI.
- Casterline, John B., Elizabeth C. Cooksey, and Abdel Fattah E. Ismail. 1989. "Household income and child survival in Egypt," *Demography* 26(1): 15–35.
- Cleland, John G. and Jerome K. van Ginneken. 1988. "Maternal education and child survival in developing countries: The search for pathways of influence," *Social Science and Medicine* 27(12): 1357–1368.
- Davidson, Russell and James G. MacKinnon. 1993. *Estimation and Inference in Econometrics*, New York: Oxford University Press.
- DeGraff, Deborah S. 1991. "Increasing contraceptive use in Bangladesh: The role of demand and supply factors," *Demography* 28(1): 65–81.
- Elo, Irma T. and Samuel H. Preston. 1996. "Educational differentials in mortality: United States, 1979–85," *Social Science and Medicine* 42(1): 47–57.
- Forste, Renata. 1994. "The effects of breastfeeding and birth spacing on infant and child mortality in Bolivia," *Population Studies* 48(3): 497–511.
- Foster, Andrew D. 1993. "Household partition in rural Bangladesh," *Population Studies* 47(1): 97–114.
- Fuller, Wayne A. 1987. *Measurement Error Models*, New York: John Wiley and Sons.
- Glewwe, Paul. 1999. "School quality, student achievement, and fertility in developing countries," in Caroline H. Bledsoe, John B. Casterline, Jennifer A. Johnson-Kuhn, and John G. Haaga (eds.), *Critical Perspectives on Schooling and Fertility in the Developing World*, pages 105–137, Washington, DC: National Academy Press.

- Gragnotati, Michele. 1999. *Children's Malnutrition in Rural Guatemala: A Multilevel Statistical Analysis*, Ph.D. thesis, Princeton University, Program in Population Studies.
- Greene, William H. 1997. *Econometric Analysis*, Upper Saddle River, NJ: Prentice-Hall.
- Griliches, Zvi. 1974. "Errors in variables and other unobservables," *Econometrica* 42(6): 971–998.
- Grosh, Margaret E. and Paul Glewwe. 1996. "An introduction to the World Bank's Living Standards Measurement Surveys," Living Standards Measurement Survey Working Paper 120, World Bank, Washington, DC.
- Grosh, Margaret E., Qing hua Zhao, and Henri-Pierre Jeancard. 1995. "The sensitivity of consumption aggregates to questionnaire formulation: Some preliminary evidence from the Jamaican and Ghanaian LSMS surveys," Poverty and Human Resources Division, Policy Research Department, The World Bank, Washington DC.
- Guo, Guang and Laurence M. Grummer-Strawn. 1993. "Child mortality among twins in less developed countries," *Population Studies* 47(3): 495–510.
- Hamill, David N., Amy O. Tsui, and Shyam Thapa. 1990. "Determinants of contraceptive switching behavior in rural Sri Lanka," *Demography* 27(4): 559–578.
- Hausman, Jerry A. 1978. "Specification tests in econometrics," *Econometrica* 46: 1251–1271.
- Hausman, Jerry A., Whitney K. Newey, and James L. Powell. 1995. "Nonlinear errors in variables: Estimation of some Engel curves," *Journal of Econometrics* 65(1): 205–233.
- Hentschel, Jesko and Peter Lanjouw. 1996. "Constructing an indicator of consumption for the analysis of poverty: Principles and illustrations with reference to Ecuador," Living Standards Measurement Survey Working Paper 124, The World Bank, Washington, DC.

- Jensen, Eric R. 1996. "The fertility impact of alternative family planning distribution channels in Indonesia," *Demography* 33(2): 153–165.
- Kiefer, Nicholas and Gary Skoog. 1984. "Local asymptotic specification error analysis," *Econometrica* 52(4): 873–885.
- Knodel, John and Malinee Wongsith. 1991. "Family size and children's education in Thailand: Evidence from a national sample," *Demography* 28(1): 119–131.
- Koenig, Michael A., James F. Phillips, Oona M. Campbell, and Stan D'Souza. 1990. "Birth intervals and childhood mortality in rural Bangladesh," *Demography* 27(2): 251–265.
- Larsen, Ulla. 1995. "Differentials in infertility in Cameroon and Nigeria," *Population Studies* 49(2): 329–346.
- LeVine, Robert A., Sarah E. LeVine, Amy Richman, F. Medardo Tapia Uribe, Clara Sunderland Correa, and Patrice M. Miller. 1991. "Women's schooling and child care in the demographic transition: A Mexican case study," *Population and Development Review* 17(3): 459–496.
- Lloyd, Cynthia and Anastasia Gage-Brandon. 1994. "High fertility and schooling in Ghana: Sex differences in parental contributions and educational outcomes," *Population Studies* 48(2): 293–306.
- MacKinnon, James G. 1992. "Model specification tests and artificial regressions," *Journal of Economic Literature* 30(1): 102–146.
- Maddala, G. S. 1977. *Econometrics*, New York: McGraw-Hill.
- Montgomery, Mark R. and Kathleen A. Burke. 1997. "Using mismeasured and proxy variables," Unpublished paper, Policy Research Division, Population Council.
- Muhuri, Pradip K. 1996. "Estimating seasonality effects in child mortality in Matlab, Bangladesh," *Demography* 33(1): 98–110.
- Njogu, Wamucii. 1991. "Trends and determinants of contraceptive use in Kenya," *Demography* 28(1): 83–99.

- Pebley, Anne R., Noreen Goldman, and Germán Rodríguez. 1996. "Prenatal and delivery care and childhood immunization in Guatemala: Do family and community matter," *Demography* 33(2): 231–247.
- Peterson, Christine, Noreen Goldman, and Anne R. Pebley. 1997. *The 1995 Guatemalan Survey of Family Health (EGSF): Overview and Codebook*, no. DRU-1583/3-NICHD in RAND Series, Santa Monica, CA: RAND.
- Preston, Samuel H. and Michael R. Haines. 1991. *Fatal Years: Child Mortality in Late Nineteenth-Century America*, Princeton NJ: Princeton University Press.
- Ravallion, Martin. 1996a. "How well can method substitute for data? Five experiments in poverty analysis," *The World Bank Research Observer* 11(2): 199–221.
- . 1996b. "Issues in measuring and modelling poverty," *The Economic Journal* 106(438): 1328–1343.
- Razzaque, Abdur, Nurul Alam, Lokky Wai, and Andrew Foster. 1990. "Sustained effects of the 1974–5 famine on infant and child mortality in a rural area of Bangladesh," *Population Studies* 44(1): 145–154.
- Richter, Kerry, Chai Podhisita, Aphichat Chamrathirong, and Kusol Soonthorndhada. 1994. "The impact of child care on fertility in urban Thailand," *Demography* 31(4): 651–662.
- Sastry, Narayan. 1996. "Community characteristics, individual and household attributes, and child survival in Brazil," *Demography* 33(2): 211–229.
- Schultz, T. Paul. 1981. *Economics of Population*, Reading, MA: Addison-Wesley.
- Scott, Chris and Ben Amenuvegbe. 1990. "Effect of recall duration on reporting of household expenditures: An experimental study in Ghana," *Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper 6*, The World Bank, Washington DC.
- Speizer, Ilene S. 1995. "A marriage trichotomy and its applications," *Demography* 32(4): 533–542.

- Stefanski, Leonard A. and Raymond J. Carroll. 1985. "Covariate measurement error in logistic regression," *Annals of Statistics* 13(4): 1335–1351.
- Stewart, John F., Barry M. Popkin, David K. Guilkey, John S. Akin, Linda Adair, and Wilhelm Flieger. 1991. "Influences on the extent of breast-feeding: A prospective study in the Philippines," *Demography* 28(2): 181–199.
- Stuebing, Kathleen W. 1997. "Maternal schooling and comprehension of child health information in urban Zambia: Is literacy a missing link in the maternal schooling-child health relationship?" *Health Transition Review* 7(2): 151–171.
- Summers, Robert and Alan Heston. 1991. "The Penn World Table (Mark 5): An expanded set of international comparisons, 1950–1988," *Quarterly Journal of Economics* 106(2): 327–368.
- Wickens, Michael R. 1972. "A note on the use of proxy variables," *Econometrica* 40(4): 759–761.
- Wolfe, Barbara, Robert Haveman, Donna Ginther, and Chong Bum An. 1996. "The 'window problem' in studies of children's attainments: A methodological exploration," *Journal of the American Statistical Association* 91(435): 970–982.
- World Bank. 1993. *Ghana Living Standards Survey (GLSS), 1987–88 and 1988–89: Basic Information*, World Bank, Washington, DC, Poverty and Human Resources Division.
- . 1995a. *Pakistan Integrated Household Survey (PIHS), 1991: Basic Information*, World Bank, Washington, DC, Poverty and Human Resources Division.
- . 1995b. *Tanzania Human Resource Development Survey: Final Report*, World Bank, Washington, DC, Population and Human Resources, Eastern Africa Department.
- . 1996a. *Jamaica Survey of Living Conditions (JSLC), 1988–94: Basic Information*, World Bank, Washington, DC, Poverty and Human Resources Division.

—. 1996b. *Peru Survey of Household Living Conditions (ENNIV), 1994: Basic Information*, World Bank, Washington, DC, Poverty and Human Resources Division.

Yatchew, Adonis and Zvi Griliches. 1984. "Specification error in probit models," *Review of Economics and Statistics* 66: 134–139.

POLICY RESEARCH DIVISION WORKING PAPERS

Recent Back Issues

1997

- 93 James F. Phillips, Fred N. Binka, Martin Adjuik, Alex Nazzar, and Kubaze Frank Adazu, "The determinants of contraceptive innovation: A case-control study of family planning acceptance in a traditional African society."
- *94 John Bongaarts and Sajeda Amin, "Prospects for fertility decline and implications for population growth in South Asia."
- 95 Barbara S. Mensch and Cynthia B. Lloyd, "Gender differences in the schooling experiences of adolescents in low-income countries: The case of Kenya."
- 96 Martin Brouckerhoff and Ellen Brennan, "The poverty of cities in the developing world."
- *97 Carol E. Kaufman, "Reproductive control in South Africa."
- *98 John Bongaarts, "Trends in unwanted childbearing in the developing world."
- 99 Mary Arends-Kuenning, "How do family planning workers' visits affect women's contraceptive behavior in Bangladesh?"
- 100 Mark R. Montgomery and Cynthia B. Lloyd, "Excess fertility, unintended births, and children's schooling."
- 101 Mary Arends-Kuenning, "The equity and efficiency of doorstep delivery of contraceptives in Bangladesh."
- *102 Sajeda Amin, Ian Diamond, Ruchira T. Naved, and Margaret Newby, "Transition to adulthood of female factory workers: Some evidence from Bangladesh."
- *103 Margaret E. Greene and Ann E. Biddlecom, "Absent and problematic men: Demographic accounts of male reproductive roles."
- 104 Michael P. Todaro, "Urbanization, unemployment, and migration in Africa: Theory and policy."
- 105 Geoffrey McNicoll, "Population and poverty: A review and restatement."

* No longer available

- 106 Sajeda Amin and Gilda Sedgh, "Incentive schemes for school attendance in rural Bangladesh."
- 107 Martin Brockerhoff and Paul Hewett, "Ethnicity and child mortality in sub-Saharan Africa."
- 108 Ann E. Biddlecom and Bolaji M. Fapohunda, "Covert contraceptive use: Prevalence, motivations, and consequences."
- 109 John Bongaarts and Griffith Feeney, "On the quantum and tempo of fertility."
- 110 Barbara S. Mensch, Daniel Bagah, Wesley H. Clark, and Fred Binka, "The changing social environment for adolescents in the Kassena-Nankana District of northern Ghana: Implications for reproductive behavior."
- 111 Martin Brockerhoff and Ann Biddlecom, "Migration, sexual behavior, and HIV diffusion in Kenya."
- 112 Zeba A. Sathar and John B. Casterline, "The onset of fertility transition in Pakistan."
- 113 Geoffrey McNicoll, "Government and fertility in transitional and post-transitional societies."
- 114 John Bongaarts, "Fertility and reproductive preferences in post-transitional societies."
- 115 Fiona Steele, Sajeda Amin, and Ruchira T. Naved, "The impact of an integrated micro-credit program on women's empowerment and fertility behavior in rural Bangladesh."
- *116 Cynthia B. Lloyd, Barbara S. Mensch, and Wesley H. Clark, "The effects of primary school quality on the educational participation and attainment of Kenyan girls and boys."
- 117 Sajeda Amin and Cynthia B. Lloyd, "Women's lives and rapid fertility decline: Some lessons from Bangladesh and Egypt."
- 118 James F. Phillips and Mian Bazle Hossain, "The impact of family planning household service delivery on women's status in Bangladesh."
- 119 Mark R. Montgomery and John B. Casterline, "Social networks and the diffusion of fertility control."

- 120 John Bongaarts, "The fertility impact of changes in the timing of childbearing in the developing world."
- 121 James F. Phillips, Wendy L. Greene, and Elizabeth F. Jackson, "Lessons from community-based distribution of family planning in Africa."
- 122 Mark R. Montgomery, "Mortality decline and the demographic response: Toward a new agenda."
- 123 Mark R. Montgomery, Mary Arends-Kuenning, and Cem Mete, "The quantity-quality transition in Asia."
- 124 Barbara S. Mensch, Wesley H. Clark, Cynthia B. Lloyd, and Annabel S. Erulkar, "Premarital sex and school dropout in Kenya: Can schools make a difference?"
- 125 John Bongaarts and Rodolfo A. Bulatao, "Completing the demographic transition."
- 126 Geoffrey McNicoll, "Population weights in the international order."
- 127 Cynthia B. Lloyd, Carol E. Kaufman, and Paul Hewett, "The spread of primary schooling in sub-Saharan Africa: Implications for fertility change."
- 128 John B. Casterline, "The onset and pace of fertility transition: National patterns in the second half of the twentieth century."
- 129 Mark R. Montgomery, Michele Gagnolati, Kathleen Burke, and Edmundo Paredes, "Measuring living standards with proxy variables."