CSAE Working Paper WPS/1995-02

The Impact of Public Services on Health Care and Morbidity: Evidence from Kenya

Simon Appleton

January 1995

Centre for the Study of African Economies
University of Oxford
St Cross Building
Manor Road
Oxford, OX1 3UL
England

Direct dial Telephone Number: (01865) 281444

The paper reports findings made by Appleton et al. (1991) in a World Bank research project 'Women and Public Services in Africa', for which funding is gratefully acknowledged. However, it represents the views of the author alone and should not be attributed to the World Bank. The author is employed by the Centre for the Study of African Economies, a Designated Research Centre of the Economic and Social Research Centre. The author would like to thank Steve Bond for advice on this topic and Giuseppe Mazzarino for writing the NAG program for numerical evaluation of double integrals. The work extends and refines the analysis in Bevan, Collier and Gunning (1989); the authors of that book provided much useful discussion as well as access to their data. Helpful comments were received from reviewers of the World Bank project and participants at seminars in Oxford and Washington DC.

Health care and status are jointly modelled using household data from Kenya. Both maternal primary education and distance to health facilities affect take-up of child health care, but the former is more powerful. Corrections for selectivity when modelling health demand of the sick are insignificant. Parental education, distance to health facilities and piped water all increase reporting of illness symptoms. The first two results are interpreted as reporting biases but the effect of piped water remains disturbing. The impact of health care on the duration of illness is estimated controlling for endogeneity and found to be favourable but insignificant.

1. Introduction

This paper uses household survey data from rural Kenya to address three main questions: what determines who falls ill? what determines who seeks treatment when ill? and what determines the duration of illness amongst those that fall ill? Of particular interest are those determinants which are amenable to policy: notably three government services — health, education and water supply.

In the case of health care, the analysis permits inferences about the relative effectiveness of different government services (subject to the usual caveats surrounding causal inferences from cross-sectional analysis). For example, one concern is to encourage parents to take sick children for treatment. Which is likely to more effective in enhancing take-up: putting a clinic in every village or giving mothers primary education? A tentative answer is provided, based on patterns of health care use amongst households with different levels of maternal education and access to health facilities.

Health status is measured by survey respondent reports of various symptoms of illness. Such data has not been extensively used in the literature[1]. Consequently, the issue is not so much the *relative* effects of different government services, but whether *any* significant effects can be identified. One particular concern is whether there are systematic reporting biases which undermine the usefulness of such data for assessing the impacts of public services. Perverse estimates of the effects of education and health facilities on the incidence of illness are consistent with such biases, although some disturbing results on the safety of piped water are less easy to dismiss. An analysis of the effectiveness of health care in reducing the duration of illness encompasses previous work on the same data-set by Bevan, Collier and Gunning (1989) but obtains substantially different results.

Health care and health status are modelled jointly in the paper. In response to Behrman and Phananiramai (1991), the analysis of who seeks treatment tests whether the use of samples of ill people only requires controls for selectivity. More crucially, investigating the effect of health care on the duration of illness poses an econometric problem because one might expect those who seek treatment to have more serious illnesses. Following Heckman (1978), attempts are made to control for this by using the results of the models of who seeks treatment.

The rest of this paper is organised as follows. Section 2 explains the measurement of the dependent variables and the structure of the econometric model relating them. Section 3 reports the inferences concerning the use of health facilities. Section 4 turns to the determinants of reported morbidity itself, starting first with the determinants of the incidence of illness and then with those of the duration of illness. Section 5 summarises and concludes.

2. Econometric Specification

A widely used economic model of health care and health outcomes in developing countries follows the theory of household production (see Pitt, 1993). Each household is viewed as allocating the time of its members in order to maximise a utility function with members' health states as arguments alongside their consumption of goods and leisure. They face constraints in the form of a time constraint and a 'full income' constraint. Technological functions exist for the production of goods within the household and the health of household members. From the solution to this problem, one can obtain reduced forms expressing household members' health care and health outcomes as functions of common exogenous variables. The model used here broadly follows this reduced form approach with the important exception that it attempts to estimate the direct effect of using health facilities upon the duration of illness. Nonetheless, the equation for the duration of illness is a 'quasi-reduced form'

[1] In a survey of the literature, Behrman and Phananiramai (1991) found only two published micro-econometric studies of health status in developing countries using self-reported measures of morbidity.

rather than an attempt to estimate a health production function proper, since information is lacking on other inputs to the health production function such as nutrient consumption.

The vectors of exogenous determinants of health care and health outcomes, X_{ji}, used here include three policy instruments: education, distance to health facilities and piped water supply. Other regressors include age, sex, relation to household head, household type (polygamous or female headed), variables for household size and demographic composition together with regional dummy variables[2].

For each individual i there are three endogenous variables: I_i, a zero-one dummy variable for whether she has suffered an incidence of illness in the last three months, T_i, a zero-one dummy variable for whether she sought treatment for their last illness and D_i, the duration of her most recent illness. Distinguishing between the incidence and the duration of illness is important because they may have different determinants. For example, use of curative health facilities may reduce the duration but not necessarily the incidence of illness. The empirical measurement of these variables is discussed briefly before the formal model is set out.

A person was defined as seeking treatment if she sought qualified medical advice[3]. Modelling health care as in this binary fashion is a strong simplification and further dimensions could be added (see Dor and van der Gaag, 1987; Gertler et al, 1987, and Gertler and van der Gaag, 1990). However, focusing on the act of seeking medical advice isolates a decision which is likely to be a matter of individual or household choice. Other measures of health care — number of visits, place of treatment, type of health practitioner, total expenditures — may be determined more by health practitioners than patients. Furthermore, whether qualified medical help was sought appears the single most relevant measure of treatment when analyzing the effectiveness of curative health services.

Health outcomes were measured by morbidity as reported by household respondents rather than being based upon qualified medical diagnosis. Consequently, some variation in the health measures used here may reflect differences in sensitivity to illness, whether real or illusory, rather than differences in illness itself[4]. Interestingly, reviews of studies of the quality of such data by Idler (1991) and Idler and Kasl (1991) have made favourable evaluations. Idler concluded that early studies of the reliability of self-reported measures of health found that there was generally a good correspondence with clinical data. She emphasised that where the two differed, this did not imply the former to be inferior. Idler and Kasl reported that recent studies generally have found self-reported data to be a better predictor of mortality than supposedly more objective measures. However, none of the studies reviewed were of developing countries and it may be that self-reported data is less reliable when

[2] Given a large number of potential determinants of the demand for health care and health outcomes, a general to specific methodology was applied, beginning with estimates of models with very large numbers of variables and gradually rejecting those of low significance (see Hendry, 1987). As a rough rule of thumb, insignificant variables were usually not retained in the final models. Gender differences in the effects of explanatory variables were systematically explored using gender interaction terms; only those which were significant were retained in the final forms.

[3] The survey recorded the first and next actions taken by an individual upon their most recent illness in the last three months. A person was defined as seeking treatment if on the first or second act she went to a dispensary, health centre or hospital. People who reported doing nothing, going to bed, buying traditional or other medicine were classified as not seeking treatment.

[4] A related problem is recall error. The survey inquired about morbidity over a three month period, which Deolalikar (1991) argues is too long. Comparing Indonesian data using both one week and three month recall periods, he found the annual prevalence of illness implied using the latter kind of information to be only 15% of that calculated using the former. Against this, Pitt and Rosenzweig (1985) express scepticism about data with a one week recall period because it may lead people to include rather trivial health complaints and because it will only record ill-health infrequently. Interestingly, however, Deolalikar found the results of regression analysis of morbidity data with both one week and three month recall periods led to similar results.

respondents lack basic education. Reporting biases are to some extent controlled for by the fact that the surveys record only illnesses which are associated with certain symptoms[5].

Reviewing the relative advantages of alternative measures of health status, Behrman and Deolalikar (1988) suggest that self-reported data on symptoms of illness such as that used here are relatively promising. The two main alternative measures of health status used in the literature are child mortality and anthropometric outcomes. It is arguable that what these alternatives report is less informative about health than reliable data on morbidity would be. Child mortality is relatively rare and presumably reflects only fairly extreme instances of ill-health. By contrast, substantial differences in weight and height between children do not necessarily indicate variations in child welfare. Both measures may also be subject to serious reporting errors and biases. In the case of mortality, this is due to its sensitivity; in the case of anthropometrics, knowledge of the child's age is essential but often hard to gauge in developing countries. Anthropometric data is also problematic because of the difficulty of choosing an appropriate standardisation for age: developed country standards may not be appropriate (see Thomas, 1991). Interestingly, the choice of health measure may matter more than is commonly thought. Evidence from Uganda shows very low (and sometimes negative) correlations between the three main alternatives (see Mackinnon, 1994).

The duration of illness variable is the number of days a person was ill during the most recent illness, in logarithmic form[6].

The structure of the model is as follows[7]. Both binary dependent variables, I_i and T_i, are modelled using a latent variable formulation, with the latent variables, I_i^* and T_i^*, being linear functions of a vector of the exogenous variables and a normally distributed error term. Thus:

(1)
$$I_i \begin{cases} = 1 \text{ if } I_i^* > 0 \\ = 0 \text{ else} \end{cases}$$

where: $$I_i^* = \beta_1' X_{1i} + u_{1i}$$

If $I_i = 1$, the following is also observed:

(2)
$$T_i \begin{cases} = 1 \text{ if } T_{i*} > 0 \\ = 0 \text{ else} \end{cases}$$

where: $$T_{i*} = \beta_2' X_{2i} + u_{2i}$$

[5] The symptom categories were designed to identify cases of ill health which were sufficiently common, obvious and distinct to be capable of accurate self-diagnosis. Specifically, the following symptoms were inquired about: fever; diarrhoea or vomiting; fever and diarrhoea or vomiting; cough; and cough with blood in sputum. The categories were at the suggestion of the late Sir Michael Wood, then Director General of the African Medical and Research Foundation. The major diseases corresponding to these symptoms are malaria, diarrhoea diseases and chest infections.

[6] The data will be right censored for those individuals ill at the time of the survey, but as these individuals cannot be identified, this source of error cannot be controlled for.

[7] In fact, separate models were estimated for adults and children (those under 16 years of age). This was because of the very different capabilities of adults and children and because of the different decision processes involved. Decisions about children's health tend to be taken by other household members and hence the characteristics of these others, for example the education of the child's parents, may be of particular importance.

and:

$$D_i = \beta_3' X_{3i} + \alpha T_i + u_{3i}$$

Given the fact that many important explanatory variables are likely to be unobserved, there may be correlation in the stochastic determinants, u_{ji}, of the three endogenous variables. Control for this is carried out by assuming the stochastic terms follow a trivariate normal distribution:

$$\begin{vmatrix} u_{1i} \\ u_{2i} \\ u_{3i} \end{vmatrix} \sim N \begin{vmatrix} 0 & 1 & \sigma_{12} & \sigma_{13} \\ 0, & \sigma_{12} & 1 & \sigma_{23} \\ 0 & \sigma_{13} & \sigma_{23} & \sigma_3 \end{vmatrix}$$

A two-stage estimation strategy is followed. In the first stage, the two dichotomous equations are jointly estimated using maximum likelihood methods. This is stage is essentially a bivariate probit with sample selection (see Van de Ven and Van Praag, 1981)[8]. The log-likelihood function, L, is:

$$L = \Sigma_{i=1,n} \{ I_i.T_i.\Phi_2(\beta_1' X_{1i}, \beta_2' X_{2i}, \sigma_{12}) \quad (4)$$
$$+ I_i.(1-T_i).\Phi_2(\beta_1' X_{1i}, -\beta_2' X_{2i}, \sigma_{12}) + (1-I_i).\Phi(-\beta_1' X_{1i})\}$$

where $\Phi_2(.)$ and $\Phi(.)$ are the cumulative distribution functions for the bivariate and univariate normals respectively.

The second stage of the model estimates equation (3) over the sample of people reported ill. Estimation of this equation is complicated by two problems: the endogeneity of T_i and the fact the sample is selected according to I_i. Following Heckman (1978, 1979) both problems — when occurring in isolation — can be seen as misspecifications of the regression equation, with an additional term required for the expected value of u_{3i} conditional upon T_i or I_i as appropriate. Given the normality assumptions, consistent estimates can be obtained by augmenting the regression equation with the conditional expectations of u_{1i} or u_{2i} derived from prior estimates. By the same reasoning, the simultaneous existence of treatment effects and sample selectivity will bias ordinary least squares estimates due to the omission of a term for the expected value of u_{3i} conditional upon both T_i and I_i.

Assuming a trivariate normal distribution, consistent estimates can be obtained by entering linear functions of estimates of the expectations of u_{1i} and of u_{2i} conditional upon both T_i and I_i. This estimation method is discussed in the context of multiple criteria for sample selectivity by Maddala (1983, pp. 278–283) and the extension to the case of a single criterion for sample selection coupled with an endogenous variable is straightforward. To provide consistent estimates, equation (3) should be augmented with an estimate of the following term:

$$E(u_{3i} | I_i = 1, T_i) = \sigma_{13}.M_{13i} + \sigma_{23}.M_{23i} \quad (5)$$

where:

$$M_{j3i} = (1-\sigma_{12}.\sigma_{12})^{-1}.(E_{ji} - \sigma_{12}.E_{ki}) \qquad j,k = 1,2$$
$$W_{1i} = -\beta_1' X_{1i}$$
$$W_{2i} = -\beta_2' X_{2i} \qquad \text{if } T_i = 1$$
$$= -\infty \qquad \text{else}$$

[8] The model can be estimated using the econometric software LIMDEP (Greene, 1991).

$$E_{ji} = \Phi_{2i}^{-1} \int_{W_{1i}}^{\infty} \int_{W_{2i}}^{W_{3i}} u_{ji}.\phi_2(U_{1i},u_{2i},\sigma_{12})\ du_{1i}.du_{2i}$$

$$\begin{aligned} W_{3i} &= \infty && \text{if } T_i = 1 \\ &= -\beta_2'X_{2i} && \text{else} \\ \Phi_{2i} &= \Phi_2(\beta_1'X_{1i},\beta_2'X_{2i},\sigma_{12}) && \text{if } T_i = 1 \\ &= \Phi_2(\beta_1'X_{1i},-\beta_2'X_{2i},\sigma_{12}) && \text{else} \end{aligned}$$

where $\phi_2(.)$ is the probability density function of the bivariate normal distribution.

Hence, using the estimates of β_1, β_2 and σ_{12} obtained from the first stage, one can obtain estimates of M_{13i} and M_{23i}[9]. Ordinary least squares regression using equation (3) augmented by these two selectivity terms will then yield consistent estimates of β_3. The coefficients upon M_{13i} and M_{23i} will provide consistent estimates of σ_{13} and σ_{23}[10].

Few other studies of developing countries have attempted to assess the effect of health care upon health status, perhaps because the use of health facilities is potentially endogenous to health status. The only example appears to be the analysis contained in Bevan, Collier and Gunning (1989) — henceforth BCG — using the same Kenyan survey data as that analyzed here (as well as data for Tanzania). However, BCG's econometric specification differs from that set out below both in its structure and in the specific variables used. They adopt a three stage approach. Firstly, they regress the duration of the most recent illness on various explanatory variables, including the number of instances of illness in the last three months. They then take the predicted duration of the most recent illness and use it as a determinant of whether people sought treatment for their most recent illness. Finally, they take the predicted probability of seeking treatment as a regressor in a model of the number of days in the last three months that a person was too ill to carry out their normal activities. BCG do not properly consider the simultaneity involved in the three stages of their approach. Issues of identification not discussed. For example, they include variables for the number of instances of various symptoms in tne last three months as determinants of the duration of illness but not of seeking treatment. However, why the variables should affect the former but not the latter is not discussed. Nor is it ever justified why these variables for the number of illnesses can be used as exogenous determinants (via the predicted probability of seeking treatment) of the total number of days of work lost from illness. However, as discussed below in Section 4.2, these differences of econometric methodology probably have less effect on the substantive conclusions drawn than the selection of explanatory variables. This paper employed a richer set of explanatory variables in initial general specifications. Consequently, it produces final forms which encompass and to some degree overturn those reported in BCG.

3. The Demand for Health Care

Tables 2.1A and 2.2B report estimates of parameters β_2 in equation (4) which determine whether sick people sought treatment for their most recent illness. Definitions of the mnemonics used for explanatory variables are given in Table 1. The unobservables influencing the demand for health care are allowed to be correlated with those for whether an individual is ill. The estimated covariance of

[9] In the econometric results, estimates of M_{13i} are denoted LAMBDAT and M_{23i} are termed LAMBDAI. Evaluation of double integrals of the sort in equation (5) was done numerically using NAG routine D01GBF.

[10] The standard errors from such a two-stage procedure will be incorrect.

these unobservables, Rho(1,2), is positive but with insignificant t-ratios[11]. Likelihood ratio tests for the restriction to a zero covariance give chi-square values of 1.5 for adults and 1.86 for children. Hence the restrictions are not rejected at 5% significance. This is important given Behrman and Phananiramai's (1991) assertion that the widespread practice of analyzing health demand on sub-samples of the sick leads to selectivity biases. For the data used here, at least, there is no evidence of significant sample selectivity. Similarly, models of health demand which impose the restriction of zero covariance do not lead to substantially different inferences (see Appleton, 1991). This is encouraging given the unattractive nature of Behrman and Phananiramai's proposed solution to the sample selectivity problem, namely to model health demand on a sample of the sick and the non-sick. Implemented as in Behrman and Phananiramai (1991), the analysis regards the non-sick as identical to the sick who do not seek treatment. In particular, a binary choice model is estimated across the whole sample, with the non-sick and untreated sick being both assigned zero health demand[12]. Consequently, such analysis does not permit the investigator to determine whether a given explanatory variable raises the demand for health care by increasing usage amongst the sick (presumably a good thing) or by increasing the numbers of the sick (quite definitely a bad thing). Which interpretation is correct seems essential from a welfare or policy point of view. (Behrman and Phananiramai justify their approach in terms of its usefulness in forecasting health care demand.)

Table 1:
Definitions of Variable Names

AGE	age (years)
BOREHOLE	1 if household drinking water source is a borehole, 0 else
DAM	1 if household drinking water source is a small dam, 0 else
FEMALE	1 if individual is female, 0 else
FEMHH	1 if household is female headed, 0 else
HDIST	distance to nearest dispensary or hospital (miles)
HSIZE	total number of household members
KIAMBU	1 if household is in Kiambu District, 0 else
KIRINYAG	1 if household is in Kirinyaga, 0 else
KISII	1 if household in Kisii District, 0 else
KISUMU	1 if household in Kisumu District, 0 else
LAMBDAT	correction for endogeneity of treatment (an estimate of M_{23i} in eq.5)
LAMBDAI	correction for selectivity of illness (an estimate of M_{13i} in eq.5)
LANDPC	land in use for agriculture by household (acres per capita)
LATRINE	1 if household uses pit latrine, 0 else
LIVPC	total value of household livestock per capita (shillings)
LPRIM	grades of primary schooling [a]

cont ...

[11] The positive covariance estimates suggests that those who for unobserved reasons are more likely to fall ill are also more likely to seek treatment if ill. This is somewhat surprising, but may partly reflect reporting biases affecting the health status variables, discussed further below.

[12] Pitt (1993) proposes modelling health demand using a multinominal logit with three categories: not sick; sick but not treated; and both sick and treated. This is similar in spirit to the approach adopted here. However, the bivariate probit with sample selectivity used here could be viewed as less restrictive than the multinominal logit in that it does not impose the 'independence of irrevelant alternatives assumption' (see Greene, 1993).

Table 1 cont ...

LSEC	grades of secondary schooling [b]
MALE	1 if individual is male, 0 else
MKTDIST	distance to nearest market, miles
MURANGA	1 if household is in Muranga District, 0 else
NYANDARU	1 if household is in Nyandarua district, 0 else
NYANZA	1 if household is in Nyanza Province, 0 else
NYERI	1 if household is in Nyeri district, 0 else
OTHREL	1 if individual is not a member of the household head's nuclear family, 0 else
PFAD	proportion of household members women
PIPE	1 if household drinking water source is communal piped water, 0 else
PKID	proportion of household members children aged under 15
POLYHEAD	1 if individual is had of a polygamous household, 0 else
POND	1 if household drinking water source is a pond, 0 else
SIAYA	1 if household in Siaya district, 0 else
SNYANZ	1 if household in South Nyanza district, 0 else
SON	1 if individual is son of the household head, 0 else
TAP	1 if household draws its drinking water from a private tap, 0 else
TREAT	1 if individual sought treatment, 0 else
TREEPC	number of trees owned by household per capita
URBDIST	distance to nearest large urban centre (miles Nairobi for Central, Kisumu for Nyanza)
WIFE	1 if individual is wife of household head, 0 else

Squared terms are denoted by the variable name followed by 2 and cubic terms by the variable name followed by 3
Gender interactions are given by prefixing *FZ* before the variable name if interacting the variable with FEMALE and by prefixing with *MZ* if interacting with MALE.
Rho (1,2) is the estimated covariance between the unobservables in the incidence and treatment equations.

Notes: [a] If followed by: *M* refers to individual's mother's schooling; *F* refers to individual's father's schooling; *W* refers to senior female in household; *H* refers to male head of household.

[b] Coded as above.

Table 2.1:
Joint Probits for the Incidence of Adult Illness and its Treatment

Sample size	2383
% ill	21.9
% of ill who sought treatment	70.3
Log-Likelihood	-1414.6

Variable	Coefficient	T-ratio	Mean	Std.Dev.
A. Determinants of who falls ill				
CONSTANT	0.616E-01	(0.16)		
FEMALE	-0.630	(-1.70) *	0.553	0.497
FZAGE	0.180E-01	(1.31)	19.490	20.928
FZAGE2	-.156E-03	(-1.05)	817.629	1224.070
MZAGE	-.190E-01	(-1.31)	16.219	21.578
MZAGE2	0.266E-03	(1.79) *	728.480	1323.667
LPRIM	0.178E-01	(1.39)	3.513	3.301
FZLSEC	-0.182	(-3.38) ***	0.129	0.645
FZHDIST	-.216E-01	(-1.26)	1.695	2.376
MZHDIST	-.495E-01	(-2.24) **	1.350	2.306
BOREHOLE	0.445	(3.59) ***	0.529E-01	0.224
FZDAM	-0.317	(-1.62)	0.302E-01	0.171
MZDAM	-1.304	(-3.16) ***	0.206E-01	0.142
MZLATRIN	-0.357	(-2.47) **	0.386	0.487
MZLIVPC	0.107E-03	(2.16) **	308.529	702.929
CHILD	-0.341	(-2.68) ***	0.328	0.469
OTHREL	-0.431	(-3.31) ***	0.128	0.334
HSIZE	-.416E-01	(-4.37) ***	8.390	4.047
PKID	-0.420	(-2.48) **	0.445	0.214
MKTDIST	0.438E-01	(2.70) ***	2.815	2.117
FZKIRI	-0.226	(-1.49)	0.495E-01	0.217
MZKIRI	-0.603	(-2.82) ***	0.474E-01	0.213
SIAYA	0.358	(3.62) ***	0.119	0.324
NYANZA	0.217	(2.81) ***	0.614	0.487
WIFE	-0.257	(-2.55) **	0.236	0.425
B. Determinants of who seeks treatment				
CONSTANT	0.751E-01	(0.07)		
FEMALE	0.614	(0.58)	0.632	0.483
FZAGE	-.709E-02	(-1.21)	24.521	22.619
MZAGE	0.350E-01	(0.85)	15.044	23.044
MZAGE2	-.421E-03	(-1.05)	756.335	1475.239
LPRIMH	0.395E-01	(1.47)	1.320	2.507
LPRIM	0.407E-01	(1.56)	2.912	3.201
MZHDIST	-0.159	(-3.47) ***	1.000	1.929
FZHDIST	-.645E-01	(-2.03) **	1.893	2.347
LANDPC	-0.102	(-1.42)	0.689	0.909
POND	0.359	(1.63)	0.939E-01	0.292
SON	-0.577	(-1.37)	0.130	0.337
POLYHEAD	0.830	(1.54)	0.421E-01	0.201
FZKISII	-0.477	(-2.49) **	0.119	0.324
MZKISII	-0.870	(-3.27) ***	0.651E-01	0.247
Rho (1.2)	0.312	(1.47)		

Table 2.2:
Joint Probits for the Incidence of Child Illness and its Treatment

Sample size	2539
% ill	19.3
% of ill who sought treatment	70.1
Log-Likelihood	-1351.8

Variable	Coefficient	T-ratio	Mean	Std.Dev.
A. Determinants of who falls ill				
CONSTANT	-0.393	(-1.64)		
FEMALE	-0.310	(-1.93) *	0.475	0.499
MZAGE	-.775E-01	(-2.03) **	4.046	5.008
MZAGE2	0.422E-02	(1.70) *	41.445	64.236
LPRIMH	0.303E-01	(2.80) ***	2.617	3.177
LPRIMW	0.230E-01	(1.77) *	1.813	2.726
MZLSECW	-0.455	(-2.19) **	0.287E-01	0.287
HDIST	-.445E-01	(-2.96) ***	2.988	2.333
BOREHOLE	0.425	(3.16) ***	0.575E-01	0.233
PIPE	0.229	(1.65) *	0.524E-01	0.223
TAP	0.479	(2.57) **	0.339E-01	0.181
LATRINE	-0.279	(-2.95) ***	0.866	0.341
LANDPC	0.274	(3.65) ***	0.493	0.409
TREEPC	-.490E-03	(-2.97) ***	94.905	206.801
OTHREL	-0.105	(-1.15)	0.180	0.385
HSIZE	-.880E-01	(-8.40) ***	9.074	3.539
PYKID	0.660	(2.81) ***	0.220	0.145
MKTDIST	0.573E-01	(3.50) ***	2.868	2.016
FZNYANZA	0.754	(6.63) ***	0.269	0.443
MZNYANZA	0.597	(5.07) ***	0.306	0.461
KISII	-0.292	(-3.00) ***	0.213	0.409
MZKIAMBU	0.562	(3.86) ***	0.744E-01	0.263
FZKISUMU	-0.503	(-2.49) **	0.366E-01	0.188
B. Determinants of who seeks treatment				
CONSTANT	2.436	(4.56) ***		
FEMALE	-1.679	(-4.64) ***	0.460	0.499
LPRIMW	0.508E-01	(1.89) *	2.035	2.883
MZHDIST	-0.107	(-3.28) ***	1.560	2.264
LIVPC	-.147E-03	(-1.48)	622.145	887.236
WTIME	-.116E-01	(-2.23) **	13.957	14.179
PFAD	1.167	(1.92) *	0.231	0.111
PYKID	-1.198	(-2.34) **	0.238	0.160
MZURBDIS	-.873E-02	(-4.53) ***	59.798	74.273
KISII	-1.178	(-6.85) ***	0.210	0.408
MZSIAYA	-0.908	(-2.33) **	0.672E-01	0.251
MZKIAMBU	-1.685	(-4.40) ***	0.876E-01	0.283
Rho(1.2)	0.232	(1.25)		

Education appears to increase the likelihood of an individual seeking treatment. For adults, both primary schooling of the individual and that of their (male) household heads has positive coefficients[13]. Amongst children, the primary schooling of the senior female also has a positive effect. By contrast, variables for the education of male heads were rejected as wholly insignificant. This suggests one mechanism for the widely observed negative correlation between maternal education and child mortality; educated mothers appear to be more willing to send sick children for treatment.

Distance to health facilities reduced the probability of seeking treatment. The impact of distance to health facilities upon usage varied by gender in several samples. Amongst both adults and children, the effect was more negative for males. These findings reinforce the impression from the descriptive statistics that the demand for health services is not lower for females, since one would expect greater 'price' elasticities for females if this was the case[14].

Interpreting the estimated effects of education and access to health facilities as causal is a strong assumption given the cross-sectional nature of the data. Evidence of changes over time or, more convincingly, experimental evidence would be necessary to be fully confident about such an interpretation. However, the results are suggestive and this important caveat will be set aside for the moment. Taken at face value, the effects of maternal primary schooling upon whether children are sent for treatment are quite large (as are those of primary schooling upon adults). Full primary schooling of the senior female in Kenyan households is predicted to raise the probability of a sick child being sent for treatment by 12 percentage points, evaluating at the means of other variables. It is interesting to compare the magnitude of these effects with those of another method of trying to increase take-up, namely increasing the availability of health services. In particular, the coefficients on variable for the distance to the nearest health facility can be used to see what sort of reduction in distance would be needed to generate a comparable rise in the probability of a sick child being sent to treatment to that brought about by complete maternal schooling. Cutting distance to health facilities is predicted to significantly affect boys only. Moreover, a 14 percentage point rise in take up amongst boys (comparable to that generated by complete primary schooling of the senior female) would require a 5 mile reduction in distance to the nearest dispensary or hospital. This required reduction is more than the mean of the distance variables, which is under 3 miles. Consequently, reliance upon an expansion of the number of health facilities — even extending to the provision of a clinic or dispensary in each village — might not be able to achieve a comparable increase in the take up of child health care to that brought about by female primary schooling. Comparable simulations for rural Côte d'Ivoire yield the same conclusion (Appleton, 1991)[15].

Amongst the other explanatory variables which are not directly of policy interest, household demographic characteristics had some significant effects. Higher proportions of young children within the household reduce the probability of children being sent for treatment. Higher proportions of female adults had a positive effect. This may reflect the fact that escorting children to dispensaries is seen as 'women's work', with young children placing competing demands upon female time. None of the

---

[13] Where there is no male head, variables for their education are set to zero. The survey did not identify the mothers of children. Consequently, instead of using the mother's characteristics as explanatory variables we use those of the 'senior female' in the household, where this refers to female heads or eldest wives of male heads. In the majority of cases, these will be the mothers of children in the household. For ease of expression, the discussion of child health sometimes equates senior females with mothers.

[14] See Alderman and Gertler (1989) for a formal presentation of the link between gender bias and the price elasticity of demand. In this sample, 71% of sick boys were sent for treatment compared to 70% of sick girls; 67% of men were sent compared to 72% of women.

[15] Other health policy changes may be more effective than simple duplication of facilities. These include improvements in quality of service, changes in user fees and adjustments in the rationing of services. Estimates of the impact of some of these policy variables on health care in Kenya is provided in Mwabu, Ainsworth and Nyamete (1993). However, the data used here lacked information on these dimensions due to the absence of surveys of local health providers.

measures of physical assets per capita were positive and significant whilst livestock value per capita had perverse effects upon children in Kenya. Relation to the household head was generally unimportant in determining whether treatment was sought.

4. Health Status

4.1. The Incidence of Illness

Tables 2.1B and 2.2B report estimates of parameters β_1 in equation (4) which determine the reported incidence of illness.[16] The results are disturbing: nearly all government interventions appear to increase the reported incidence of illness. Those living further away from health facilities are less likely to report having been ill, *ceteris paribus*. Reported child illness rises with both maternal and paternal primary schooling. Similarly, adult illness rises with personal primary education. Households with piped drinking water report higher illness. There is no consistent pattern in relation to different measures of household assets[17]. Taken at face value, these results would provide rather damning indictment of development! Perhaps more plausibly, they indicate severe limitations as to what data on self-reported morbidity can tell us about policy.

First, consider the effects of education. In fact, these are somewhat mixed. Personal secondary schooling powerfully reduces reported women's illness[18]. Maternal secondary schooling apparently benefiting the health of boys (only). In general, however, the data shows strong negative effects of parental primary schooling on reported child health in particular[19]. Similar findings have been made for other developing countries[20]. These unfavourable effects schooling are implausible, especially given the strong negative correlation between maternal primary schooling and child mortality documented in Kenya and elsewhere (see World Bank, 1989; Cochrane et al. 1980)[21]. These findings can be reconciled if there is a systematic bias in the reporting of illness. Educated parents may be more likely to report as ill a child who has a given objective medical condition. For example, they may have been taught to identify child illnesses. More generally, educated parents may have different standards of

[16] The coefficients can be interpreted similarly to those on independent bivariate probits. In particular, scaling by the probability of falling ill yields the marginal effect of the relevant explanatory variable.

[17] The number of tea and coffee trees per capita had favourable effects upon the incidence of child illness. However, the other two physical asset variables — land and livestock per capita - had mixed effects, being either perverse or insignificant.

[18] Two transmission mechanisms suggest themselves. Firstly, analysis of the same data-set reveals that secondary schooling has a powerful effect on female labour market participation (see Hoddinott and Collier, 1991). This may reduce illness by raising household income, female bargaining power and the costs of illness (on the latter point, see Mwabu and O'Connell, 1989). Secondly, female secondary schooling tends to reduce fertility and this may be beneficial to maternal health (Appleton, 1995).

[19] The effect is visible in simple cross-tabulations: children average 0.42 reported illnesses every three months in households where the senior female is uneducated; where the senior female has full primary schooling, the average is 0.56.

[20] These include Tanzania and the Côte d'Ivoire (Appleton, 1991); Indonesia (Pitt and Rosenzweig, 1986); and Uganda (Mackinnon, 1994); and Ghana (Victor Levy, personal communication, 1990). One counter-example is for Nicaragua, where Wolfe and Behrman (1984) find literate women report less child morbidity.

[21] The persuasiveness of these correlations has been questioned (Behrman, 1991) but remains widely accepted.

what constitutes an illness. Uneducated and objectively less healthy households may regard certain conditions of ill health as normal and not worth reporting as illnesses[22]. Deolalikar (1991) states that maternal education is commonly known to increase the perception and early recognition of disease. If maternal education does have this effect, it may result in higher reported (but not actual) incidence of child illness and lower mortality, in part because of the higher awareness of disease.

A similar argument may explain the negative estimated effect of distance from health facilities. Treatment could conceivably be injurious to health[23]. However, it seems more plausible that the effect is spurious. In particular, vigorous preventive health care is likely to raise awareness of illness and hence may increase its reported occurrence. This argument can be made more general, since it is hard in practice to separate preventive from curative health care. In particular, apparently 'curative' consultations crucially transmit information about the symptoms and causes of disease (see Appleton and Mackinnon, 1993). Furthermore, use of health facilities is likely to increase the probability of recalling an illness. Another possible explanation of the negative relation between health facilities and the incidence of ill-health is reverse causality. Health facilities may be placed in areas of high morbidity (see Rosenzweig and Wolpin, 1986). However, it is unclear to what extent this pattern of placement exists in rural Kenya and casual observation suggests that the reverse may be true, with superior health infrastructure being provided to in more prosperous, healthier areas.

Perhaps the 'perverse' result of policy-related variables which gives most concern is that of piped water. Boreholes, communal pipes and in-house taps all significantly increase the probability of children being reported as having fallen ill in the last three months (compared with the default of getting water from a river or stream). Boreholes have a similar impact on the incidence of adult illness. These effects are also very large, typically more than doubling the risk of illness[24]. This general pattern of results — that more 'developed' sources of water appear to be more dangerous — is surprising. Nonetheless, it is apparent in simple cross tabulations of the data and has also been discovered in some cases for Côte d'Ivoire and Tanzania (see Appleton, 1991). Again, it may be that this again reflects upon the quality of reported morbidity data. However, this seems less plausible than the reporting biases posited in the case of education or health facilities and has not been suggested elsewhere in the literature. Instead, it may be that the results should be taken at face value: piped water may be more insanitary in the countries studied. For example, temporary breakdowns in the flow of water through pipes may sometimes occur, which cause them to become unsafe. Boreholes and communal pipes may be covered in filth and be insanitary; as relatively new innovations, people will not have had the years of experience and tradition associated with the use of natural water sources. It could be that piped water is not in itself a greater health risk but is simply not the health benefit that people perceive it to be. As a consequence, people consume it without taking the precautions which they may do with other natural sources. For example, water from ponds and dams, apparently the most healthy according to the results, is often discoloured and hence unlikely to be consumed directly. By contrast, piped water is usually clear and hence seldom boiled.

Aside from these key policy related variables, a number of other findings emerged. Pit latrines, as opposed to the default of no toilet, appeared to reduce the incidence of illness. Considerable spatial variation in morbidity was discovered, with Nyanza Province appearing to be a markedly less healthy

[22] Similar biases may arise from measures of household economic status, such as assets. This is Deolalikar's (1991) interpretation of his finding of a significant positive effect of household income upon reported child illness.

[23] Little or no effective help may be provided if the quality of treatment is very low, for example due to an absence of essential drugs. In such cases, people's health may actually suffer from seeking treatment as a result travelling long distances and waiting in rooms full of sick people.

[24] Taking the marginal effects evaluated at the mean probability, boreholes increase the probability of being ill in the last three months by 12 percentage points for adults and 13 for children. Amongst children, in-house taps raise the reported incidence of illness by 14 points whilst communal pipes increase it by 7 points. This are very large increases given that the mean probability of illness is around 20%.

environment than Central. Age-illness profiles differed markedly between the sexes. For men, the curve depicting the effect of age on the incidence of illness has a U-shape with the young and the old being more likely to fall ill. For women, there is an inverse U-shape, perhaps reflecting the additional stress on female health caused by child-bearing. Individuals less closely related to the household head are reported as having less illness, *ceteris paribus*. Amongst adults, it is heads and then their wives who are most likely to be reported ill; amongst children, it is those who are offspring of the head. This is contrary to what one might expect if heads gave preference to those closest to them. Perhaps the most plausible interpretation is that variations in reported illness with relation to the head reflect differential reporting bias rather than genuine differences in morbidity itself. In particular, the survey respondents are typically household heads and it would not be surprising if they were most aware of the illnesses of their own offspring over the past three months. Similarly, the negative effect of household size probably reflects a reporting bias, with respondents from larger households being more likely to forget the illness of anyone individual, rather than genuine economies of scale in health.

4.2. The Duration of Illness

Those who seek treatment tend to be ill for longer than those who do not (Table 3 refers).

Table 3:
Duration of Most Recent Illness and Receipt of Treatment

	Men	Women	Boys	Girls
Mean Days Illness (Standard Deviations in Brackets)				
With Treatment	15.1	16.8	11.8	11.6
	(17.1)	(18.9)	(12.7)	(12.2)
Without Treatment	8.8	13.0	8.2	8.7
	(8.2)	(20.6)	(10.5)	(6.8)

One would not wish to infer that the receipt of treatment makes people ill for longer. Instead, it presumably reflects the fact that those with more serious, longer lasting illnesses are also more likely to seek treatment. An attempt to control for this can be made by allowing the unobservables in the equations (2) and (3) to be correlated. Identification of the effect of treatment is achieved by virtue of the normality assumptions and by exclusion of distance to health facilities as a determinant of the duration of illness. The final forms of the models for the duration of illness described in Section 2 are given in Tables 4.1B and 4.2B. For comparability, Tables 4.1A and 4.2A give parallel results but assume that treatment is exogenous. Making this assumption produces regression results similar to the simple descriptive statistics in Table 3: treatment appears to significantly increase the duration of illness. Once controls for endogeneity are made in Tables 4.1B and 4.2B, this result is overturned[25]. For adults, treatment appears to reduce the duration of illness by around 20%; for children the figure

[25] Results (not reported) using the Heckman dummy endogenous variable model but not correcting for selectivity, show that it is the control for endogeneity of treatment rather than for the selectivity of illness which is important (see Appleton, 1991).

is 14%. However, in both cases, the effects are not statistically significant, with t-ratio's less than one. The coefficient on the variable LAMBDAT (the estimated treatment residual) is positive and significant for adults and children. This is consistent with the hypothesis of a positive covariance between the unobservables affecting the probability of seeking treatment and the duration of illness.[26]

Table 4.1:
Determinants of the Duration of Adult Illness

Number of Observations	522
Mean of Dependent Variable	2.265

	A. OLS		B. CORRECTED			
Variable	Coefficient	T-ratio	Coefficient	T-ratio	Mean	Std.Dev.
CONSTANT	1.458	(7.53) ***	1.756	(5.90) ***		
FEMALE	0.375	(2.46) **	0.463	(2.79) ***	0.632	0.483
MZAGE	0.608E-02	(1.97) **	0.704E-02	(2.18) **	15.044	23.044
LPRIMH	-.333E-01	(-2.53) **	-.231E-01	(-1.64) *	1.969	2.861
BOREHOLE	0.349	(2.46) **	0.327	(2.21) **	0.920E-01	0.289
PIPE	0.352	(2.29) **	0.356	(2.32) **	0.613E-01	0.240
PFAD	0.493	(2.51) **	0.413	(1.96) **	0.338	0.195
MZFEMHH	-0.835	(-2.07) **	-0.819	(-2.03) **	0.958E-02	0.975E-01
URBDIST	-.248E-02	(-2.40) **	-.284E-02	(-2.64) ***	106.090	63.132
SNYANZA	0.501	(3.17) ***	0.553	(3.41) ***	0.224	0.417
MZKIRINY	0.817	(2.56) **	0.824	(2.56) **	0.134E-01	0.115
KISII	0.631	(4.96) ***	0.557	(4.21) ***	0.184	0.388
MZKIAMBU	0.475	(2.13) **	0.528	(2.35) **	0.364E-01	0.187
TREAT	0.482	(5.93) ***	-.234E-02	(-0.01)	0.703	0.457
LAMBDAI			0.147E-01	(0.12)	1.221	0.395
LAMBDAT			0.299	(1.89) *	0.638E-02	0.767

Std. Error of Regr.	0.814	0.812
R - squared	0.169	0.176
Adjusted R - Squared	0.148	0.151

[26] LAMBDAI, the correction for the selectivity of illness, is insignificant. The negative coefficient indicates that those who, for unobservable reasons, are more likely to be ill, are likely to recover more quickly from a given illness, *ceteris paribus*. This may reflect the acquisition of greater immunity (for example, to repeated bouts of malaria).

Table 4.2:
Determinants of the Duration of Child Illness

Number of Observations	491
Mean of Dependent Variable	2.03

Variable	A. OLS Coefficient	A. OLS T-ratio	B. CORRECTED Coefficient	B. CORRECTED T-ratio	Mean	Std.Dev.
CONSTANT	2.088	(24.08) ***	2.031	(15.99) ***		
LPRIMH	-.252E-01	(-2.34) **	-.222E-01	(-2.05) **	2.925	3.278
LANDPC	-0.257	(-3.27) ***	-0.206	(-2.57) **	0.562	0.464
TREEPC	-.335E-03	(-1.44)	-.472E-03	(-2.01) **	64.705	150.949
TAP	-0.545	(-2.86) ***	-0.491	(-2.58) **	0.346E-01	0.183
MZKISUMU	0.538	(3.01) ***	0.628	(3.47) ***	0.428E-01	0.203
TREAT	0.254	(3.31) ***	-.358E-01	(-0.26)	0.701	0.458
LAMBDAI			0.183	(2.88) ***	1.232	0.694
LAMBDAT			0.248	(2.79) ***	0.351E-02	0.790
Std. Error of Regr.	0.771	0.764				
R - squared	0.079	0.099				
Adjusted R - Squared	0.068	0.084				

The results reported here differ substantially from those obtained on the same data by Bevan, Collier and Gunning (1989). In particular, BCG find a *significant* favourable effect of treatment on the number on days lost from work due to illness in Kenya (and a significant unfavourable effect in Tanzania). As noted in Section 2, BCG use a different econometric structure to that adopted here but it appears that this is not the cause of the discrepancy in the two sets of results. Instead, BCG's finding seems to be due to the omission of dummy variables for location when modelling the number of days lost due to illness. In particular, the author attempted to replicate BCG's results using their own methods and obtained similar (but not numerically identical) results. However, when a dummy variable for residence in Central region was added to the replicated equation for days lost due to illness, it induced a sign reversal in the effect of the probability of seeking treatment. Ceteris paribus, those in Central Province were more likely to seek treatment and less likely to be ill for a long period. Consequently, omitting the Provincial dummy variable from the illness equation appears to have upwardly biased BCG's estimated coefficient on the probability of seeking treatment.

In contrast to BCG's findings, the rather weak signal-to-noise ratios for treatment reported here are disappointing. However, they appear rather favourable when contrasted with sometimes significantly negative effects of treatment found elsewhere for Tanzania and Côte d'Ivoire (Appleton, 1991). Taken together, these results suggest that the data and methods used here may not be suitable for identifying the impact of health care use on health status. In part, this is due to the formidable econometric problems of separating out the true effects of treatment from the bias induced by those who seek treatment tending to have more serious illnesses. Furthermore, there may be a reporting bias which may offset any objective improvements gained. The receipt of treatment may cause an increase in the reported, but not actual, duration of illness. Receiving treatment may confer an additional importance onto ones' illness, making its duration less likely to be under-reported. Furthermore, a patient may be prescribed a course of drugs and regard herself as ill until the course has finished. A 'reporting bias' interpretation of perverse treatment effects is particularly plausible in those cases where the duration of illness is reported by another party such as the household head. If treatment is sought for an illness, this act may attract the attention of the household respondent to it, making them less likely to downgrade its severity. However, one does not have to appeal to data limitations to explain the insignificant effect of treatment in reducing illness in Kenya. Many common ailments may be either hard to treat or of short duration even in the absence of treatment, which may be received late. Medical services may act not so much to reduce the duration of common illnesses as to alleviate pain and to screen a few potentially very dangerous illnesses. Moreover, the quality of the rural medical facilities may be low.

Aside from treatment effects, a number of other variables of relevance to policy affect the duration of illness. The primary schooling of the male head has a favourable impact on both adults and children. Household holdings of physical assets — land and tree crops — also reduce the duration of illness of children. Unlike the results for the incidence of illness, these effects have the expected sign. This is consistent with Deolalikar's (1991) observation that self-reported data on the duration of illness conditional on its occurrence is less likely to be subject to reporting biases than the data on the incidence of illness itself. However, water supply variables again have surprising results — this time upon adults. Receiving water from a borehole or a communal pipe increases the duration of adult illness by about a third, *ceteris paribus*. This provides some support for the suggestion that the apparently adverse health effects of some man-made water supply sources are genuine.

5. Summary and Conclusions

The above findings include two of particular interest to policy-makers. The first is the powerful apparent effect of maternal primary education in inducing take-up of curative child health services. This may be one of the processes explaining the negative correlation between maternal education and child mortality. Moreover, the results suggest that increased maternal education may be more effective in improving use of child health facilities than providing more accessible health facilities. This is complements simulations presented by Summers (1994) which indicate that female education is a more cost-effective means of reducing child mortality than health expenditure. As such, the results provide some support for the growing international consensus in favour of enhancing female education — most recently expressed at the 1994 World Population Summit. The second potentially important finding is more unexpected: the apparently adverse impact of piped water on health. This may warrant further investigation in the form of specialist research on the long-term health consequences of water supply projects in rural Africa.

The paper also has some methodological implications for researchers in the field. Symptom based subjective reports of illness, such as those used here, may be of limited use in addressing a number of policy issues. In particular, strong adverse effects of parental education on reported child health and of proximity to health facilities on health in general are more likely to reflect reporting biases than genuine phenomena. As such, these kinds of measures of health status appear not to live up to their initial promise as perceived by Behrman and Deolalikar (1988).

The analysis of the impact of treatment on the duration of health facilities had mixed success. Controlling for endogeneity reversed the apparent negative relation but did not yield the significant positive impact reported by Bevan, Collier and Gunning (1989). That earlier optimism was shown to arise through an omitted variables bias.

Finally, controlling for the selectivity of illness was found to be of marginal importance in most models of the demand for health care and the duration of illness. This suggests, contrary to Behrman and Phananiramai (1991), that the large existing literature on health demand using samples only of sick people may still be valid.

References

Alderman, Harold and Paul Gertler (1989) 'Family Resources and Gender Differences in Human Capital' mimeo, International Food Policy Research Institute: Washington DC.

Appleton, S.M. (1991) 'Socio-economic determinants of education, health and fertility in Africa' unpublished PhD. dissertation, University of Oxford: Oxford.

— (1995) 'Socio-economic factors, fertility and its proximate determinants in the Côte d'Ivoire' CSAE Working Paper, WPS 95.3: Oxford.

Appleton, S.M., D.L.Bevan, K.Burger, P.Collier, J.W.Gunning, L.Haddad and J.Hoddinott (1991) 'Public Services and Household Allocation in Africa: Does Gender Matter?' mimeo, CSAE: Oxford.

Appleton, S.M. and J.Mackinnon (1993) 'Health and education in the Least Developed Countries', mimeo, CSAE: Oxford.

Behrman, Jere R. (1991) 'Women's Schooling and Non-Market Productivity: A Survey and a Reappraisal', Women in Development Working Paper, World Bank: Washington DC.

Behrman, Jere R. and A.B. Deolalikar (1988) 'Health and Nutrition' Chapter 14 of H.Chenery and T.N.Srinivasan (eds) *The Handbook of Development Economics* Vol.1, pp. 631–711 North Holland: Amsterdam.

Behrman, Jere R. and Mathana Phananiramai (1991) 'The Demand for Curative Treatment in Thailand: Sample Selection and Interactions' mimeo, University of Pennsylvania: Philadelphia.

Bevan, D.L., P.Collier and J.Gunning (1989) *Peasants and Governments* Clarendon Press: Oxford.

Cochrane, S. H., J. Leslie and D.J.O'Hara (1980) 'The Effects of Education on Health' *World Bank Staff Working Paper* No.405 World Bank: Washington DC.

Deolalikar, Anil B. (1991) 'Intrahousehold Allocation of Health Inputs and Distribution of Health Outcomes among Indonesian Children' in *Robert S. McNamara Fellowships Program: 10th Anniversary Publication* Economic Development Institute, World Bank: Washington DC.

Dor, Avi and Jacques van der Gaag (1987) 'The Demand for Medical Care in Developing Countries: Quantity Rationing in Rural Cote d'Ivoire' *Living Standards Measurements Study Working Paper* No.35 Washington DC, World Bank.

Gertler, Paul and Jacques van der Gaag (1990) *The Willingness to Pay for Medical Care* World Bank: Washington DC.

Gertler, Paul, Luis Locay and Warren Sanderson (1987) 'Are User Fees Regressive? The Welfare Implications of Health Care Financing Proposals in Peru' *Journal of Econometrics* Vol.36 pp. 67–88.

Greene, W.H. (1991) *LIMDEP Version 6.0* Econometric Software, Inc.: New York.

— (1993) *Econometric Analysis* (2nd edition) Macmillan: New York.

Heckman, J. (1979) 'Sample Selection Bias as a Specification Error' *Econometrica* Vol. 56, pp. 91–118.

— (1978) 'Dummy Endogenous Variables in a Simultaneous Equation System' *Econometrica* Vol. 46, pp. 931–59.

Hoddinott, John and Paul Collier (1991) 'Labour' in Appleton et al. (eds.) 1991.

Idler, Ellen L. (1991) 'Self-assessed Health and Mortality: a Review of Studies' *International Review of Health Psychology*.

Idler, Ellen L. and Stanislav Kasl (1991) 'Health Perception and Survival: Do Global Evaluations of Health Status Really Predict Mortality?' *Journal of Gerontology* Vol.46 No.2 S55–65.

Maddala, G. (1983) *Qualitative and Limited Dependent Variables in Economics* New York, Cambridge University Press.

Mackinnon, John (1994) 'Health in Uganda' mimeo, CSAE: Oxford.

Mwabu, Germano, Martha Ainsworth and Andrew Nyamete (1993) 'Quality of medical care and choice of medical treatment in Kenya' *Journal of Human Resources* Vol.28 No.4.

Mwabu, Germano and Stephen O'Connell (1989) 'Health maintenance by workers in rural and urban production structures' mimeo, Dept. of Economics, University of Nairobi: Nairobi.

Pitt, M. (1993) 'Analyzing Human Resources: Health' in *Understanding the Social Effects of Policy Reform* edited by M.Ferroni, L.Demery and C.Grootaert, Washington D.C., the World Bank.

Pitt, M. and M. Rosenzweig (1985) 'Health and nutrient consumption across and within farm households' *Review of Economics and Statistics* Vol.67, pp. 212–223.

— (1986) 'Agricultural Prices, Food Consumption and the Health and Productivity of Farmers' in Inderjit Singh, Lyn Squire and John Strauss (eds.) *Agricultural household models: extensions, applications and policy* Johns Hopkins University Press: Baltimore.

Rosenzweig, Mark R. and K.Wolpin (1986) 'Evaluating the Effects of Optimally Distributed Public Programs' *American Economic Review* Vol.76 No.3, pp. 470–487.

Thomas, D. (1991), 'Like Father, Like Son: Gender Differences in Household Resource Allocations', mimeo, Yale University.

Van de Ven, W. and B. Van Praag (1981) 'The Demand for Deductibles in Private Health Insurance' *Journal of Econometrics* Vol.17, pp.229–252.

Wolfe, Barbara L. and Jere R. Behrman (1984) 'Determinants of Women's Health Status and Health Care Utilisation in a Developing Country: a Latent Variable Approach' *Review of Economics and Statistics* Vol.56 No.4, pp. 696–703.

World Bank (1989) *Kenya: The Role of Women in Economic Development* World Bank: Washington DC.