African Economic Conference 2009

Fostering Development in an Era of Financial and Economic Crises

11 – 13 November 2009 • United Nations Conference Centre • Addis Ababa, Ethiopia

Demand for health in Ethiopia: exploratory analysis from welfare monitoring surveys

Alemayehu Geda
Abebe Shimeles

African Development Bank Group

Economic Commission for Africa

Demand for health in Ethiopia: exploratory analysis from welfare monitoring surveys[1]

Alemayehu Geda
Department of Economics
Addis Ababa University
ag@ethionet.et

Abebe Shimeles
(Corresponding author)
African Development Bank
a.shimeles@afdb.org

Paper presented for African Economic Conference
Addis Ababa
November 11-13, 2009

[1] Authors thank the AERC for financial support. Earlier version of this paper was presented in Accra AERC workshop, August 2008. We thank participants for their insightful comments.

Abstract:

This paper investigates demand for health in Ethiopia using a large welfare monitoring survey collected in recent period using alternative indicators of health status such as self-reported illness episodes, number of days lost due to illness and stunting. We found strong evidence that health status varies with socio-economic characteristics of an individual. Consistent with the large empirical evidence, our findings suggest that the level of schooling achieved by the individual in rural areas progressively affects health status and the result is robust to different estimation approaches. In addition, access to health services, affordability and attitudes towards health facility, as well as employment status determine infirmity experienced by individuals.

1. Introduction

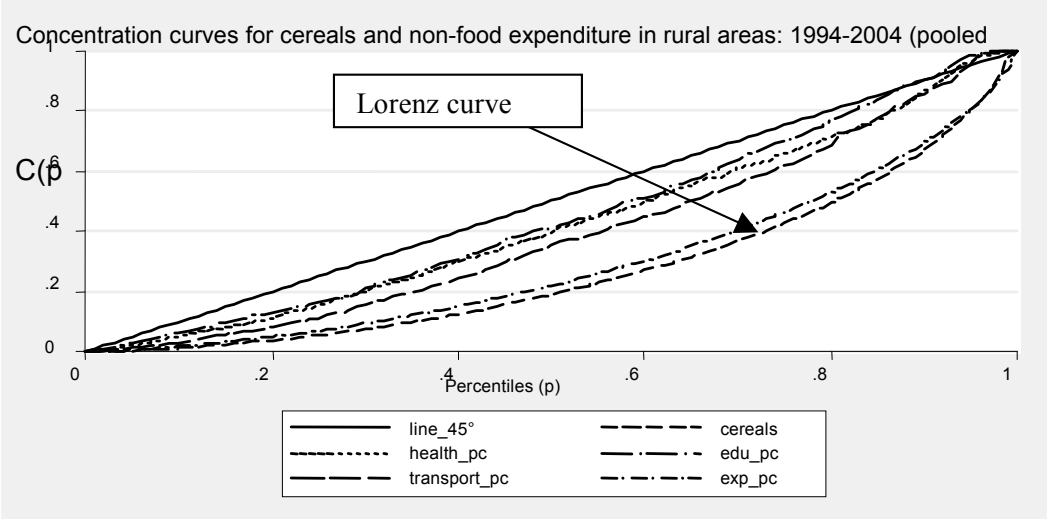
Ethiopia's disease burden is one of the highest in Africa. Life expectancy at birth is close to 40 years and improved little over the years. Close to 50% of children are stunted and killer diseases such as Tuberculosis are on the rise. The country spends about 5% of its GDP on health related services but with negligible impact due to very low GDP. In absolute terms, percapita health consumption expenditure is about 5 USD in current prices, with more than 50% spent by the government (see Table 1). Household expenditure on health related services is quite low relative to other countries. It is estimated that households in Ethiopia spend approximately 2.3% of their total budget on health related services, which compared to 4% for Africa is significantly small (Figure 1). Such low demand for health obviously undermined health services in the country. Hospital beds per 100,000 people are less than 1 and have not improved much over time. Infant mortality rate though improved recently still is one of the highest (79 per 1000 live births) in the world.

Table 1: selected health indicators in Ethiopia

Indicators	2000	2005
ARI treatment (% of children under 5 taken to a health provider)	16	18.7
Births attended by skilled health staff (% of total)	5.6	5.7
Health expenditure per capita (current US$)	5.1	5.6
Health expenditure, private (% of GDP)	2.5056	2.5705
Health expenditure, public (% of GDP)	2.8944	2.7295
Health expenditure, total (% of GDP)	5.4	5.3
Immunization, DPT (% of children ages 12-23 months)	56	69
Immunization, measles (% of children ages 12-23 months)	52	59
Incidence of tuberculosis (per 100,000 people)	307.3787	343.9043
Life expectancy at birth, total (years)	42.2935	42.65231
Malnutrition prevalence, height for age (% of children under 5)	51.5	46.5
Malnutrition prevalence, weight for age (% of children under 5)	47.2	38.4
Mortality rate, infant (per 1,000 live births)	92.4	79.7
Mortality rate, under-5 (per 1,000)	150.6	127
Out-of-pocket health expenditure (% of private expenditure on health)	79.1	78.3
Tuberculosis cases detected under DOTS (%)	32.79118	32.69776
Tuberculosis treatment success rate (% of registered cases)	80.12272	79.34106

Source: WDI 2008

Figure 1: concentration curves for health, education expenditures in rural areas



Source: Shimeles (2007)

Poverty explains much of the health hazard in Ethiopia. According to recent figures, the percentage of households who could not meet the daily calorie intake necessary for normal bodily activities are about 65% suggesting the widespread malnourishment and deterioration of human condition that may have an adverse impact on labor productivity. Thus, improving health conditions evidently becomes an important policy concern. In this paper we investigate determinants of health outcome based on a relatively large data set collected recently (2005) by the Central Statistical Authority of Ethiopia which covered 21,000 households with close to 100,000 individual histories. We investigate health demand based on the theoretical framework that links health outcomes as partially determined by choices individuals make over a life cycle (Grossman, 1972; 2000; Ajake and Mwabu, 2007a). The main thrust of the health demand theory is that individuals consider health status as a means for better life (investment) as well as an end with positive utility gained from good health (consumption). As a result, health is partly

produced by the individual by accessing health care services and conforming to behaviors that derail health, such as avoiding substance abuse, exercising, resting and other mechanisms of protecting the body from fatigue, exhaustion and communicable diseases such as HIV/AIDs. Some of these measures cost money and have prices in a market that functions smoothly. Thus, we expect that factors such as access to health care services and their prices, and income affect the production as well as consumption of health goods such as wellness.

We use three indicators of wellness to capture health status. These are episodes of self-reported illness that had occurred in the last two months of the survey, number of days spent in bed or inactive due to illness and stunting. These indicators are direct measures of infirmity with a potential to affect both labor productivity as well as quality of life for the individual. We proxy health inputs and their relative prices with such variables as physical access to health facilities, affordability and individual preference for health care (beliefs on modern medical care). Our findings suggest that socio-economic characteristics of the individual are strongly correlated with health outcomes. Of particular importance is the role that education plays in determining health outcomes in Ethiopia. Generally, the higher the level of schooling achieved by the individual, the lower the probability of falling sick or staying bed or remaining stunted. Policy measures designed to improve education leads also to better utilization of health services and preventive measures to avoid communicable diseases. Access to health facilities is the other most important factor that influence health outcomes in Ethiopia as well as demographic factors such as age, family size, and spatial differences. The rest of the paper is organized as follows: the next section sketches the methodological framework, section 3 describes the data, section 4 discusses the results and section 5 concludes the discussion.

2. Methodological framework and empirical strategy

2.1. Sketch of the health demand theory

The seminal paper by Grossman (1972) and his subsequent works laid the foundation for the literature that emerged in the last three decades on the demand for health services and formed the basis for the large empirical work that followed (see for excellent survey Strauss and Thomas (1998); Ajakiye and Mwabu, 2007a). The main thrust of the theory behind health demand in the literature is that generally individuals make informed choices over the life cycle to improve their health status by investing in health care systems, conforming to health-enhancing behaviors and practices given their initial health endowment inherited genetically from their parents and environment. These decisions take the form of investment in health following the literature on human capital, particularly education, as a means to better livelihood and consumption where good health features as an end in itself. Following the set up and notations in Grossman (2000) we sketch below the formal model to motivate our empirical model. Suppose we define an intertemporal utility function of a typical consumer over stock of health (H_t), and consumption of other good (representative commodity)- Z_t so that:

$$U = U(H_t, Z_t) \quad t = 0,1,2,.....,n \tag{1}$$

Equation (1) postulates that individual consumer gains positive utility (by the quasi-concavity assumption) out of the consumption of health flows in each period. In this set up, the stock of health at t=0 is given or to be precise inherited from and nourished by parents (and society). For subsequent periods however health stock is endogenous to the choices the individual makes subject to life-time budget constraints. The first constraint is that the individual is responsible for the production of health in the life time by making appropriate investment (I_t). Thus net investment in health equals gross investment (I_t) less depreciation (δ_t) and is given by:

$$H_{t+1} - H_t = I_t - \delta_t H_t \tag{2}$$

Investment in health is a function of a vector of health inputs (M_t), such as those bought from health care systems and deliberate avoidance of health hazards (substance abuse, etc) and other appropriate precautions individuals make to gain better health. It also

depends on time allocation decisions on health-enhancing activities (TH_t) and stock of human capital, which is time-invariant (E) and exclusive of health capital (or conditional on specific health capital). Similarly the production of Z_t is a function of a vector of inputs (X_t), the same time variables and stock of knowledge. The production function for health investment and thus other good then can be specified as follows:

$$I_t = I_t(M_t, TH_t; E) \tag{3}$$

$$Z_t = Z_t(X_t, T_t; E) \tag{4}$$

The budget constraint for the consumer must satisfy equality between the discounted life-time expenditure on the purchase of inputs (M_t and X_t), and life-time income or earnings. This may be captured with two equations, one for the expenditure-income flows and the other for time constraint.

$$\sum_{t=1}^{n} \frac{P_t M_t + Q_t X_t}{(1+r)^t} = \sum_{t=1}^{n} \frac{W_t TW_t}{(1+r)^t} + A_0 \tag{5}$$

Where P_t, Q_t represent prices of respective inputs, W_t is hourly wage rate, TW_t is number of hours of work, and the last term on the right hand side is initial assets, r is rate of discount rate usually proxied by market interest rate. Since time is also scare resource to the consumer, exhaustive allocation of it can be specified in the form of time allocated for work (TW_t), inputs for the production of health and the other good (TH_t, T_t) and time lost due to illness or other infirmity (TL_t). For a fixed total time available each period, the time constraint can be specified simply as:

$$\Omega = TW_t + TH_t + TL_t \tag{6}$$

Equations 1-6 provide the building block of the health demand function that has been a subject of intense theoretical investigation and empirical application. The consumer's

objective is now to maximize (1) subject to the constraints specified in equations 3-6. The equilibrium conditions that emerge from the model offer insights into the optimal health capital to be attained from investment in health. We skip key complications that arise in formulating closed form solutions to the optimizing framework and the characterization of equilibrium when optimal life time is endogenously determined by the model. We only consider special cases that have been widely applied in the literature. One of these cases is the assumption where the marginal utility of healthy time is zero to the consumer or healthy time does not enter as an argument directly into the utility function. In this case, health becomes a purely investment good so that the consumer's interest is in ensuring the marginal return to investment in health is equal to the opportunity cost of capital. Grossman (2000) derives a reduced form equation that can be estimated from routine household survey data:

$$\ln H = \alpha \ln M + \rho_H E - \delta_t - \ln \partial_0 \qquad (7)$$

Equation (7) is a linear health production function where health status or outcome is a function of health inputs, human capital, mainly education and depreciation of health over time, usually associated with age. The key predictions of the model are education increases the efficiency of health production so that they are positively correlated. Health inputs lead to higher production of health status for the individual. Other specifications also include income and wealth as further important variables in generating more health for the consumer.

2.2. Empirical strategy

The acknowledged problems of estimating equations such as (7) or any of its variants are the following. First, in real life from which data is drawn to test the hypothesis laid out in this model the key variables are interrelated in a complex chain so that it is difficult to establish causality say from education to good health and vice versa (Grossman, 2004). Secondly, unobserved factors that routinely enter the error term are in most cases highly correlated with the regressors of the health demand function. As a result, estimation

methods such as the OLS are biased and inconsistent (Ajakiye and Mwabu, 2007a). Third, most health outcome indicators are measured with error, particularly in the context of developing countries further generating sources of bias in the role of key determinants. It is not uncommon to find in household surveys from developing countries that episodes of illness are strongly and positively correlated with per capita income mainly due to systematic biases in reporting illness episodes between poor and non-poor households. Poor households are accustomed to "live" with minor ailments, sometimes without even recognizing it. In a number of applications this may then lead to counter intuitive relationship between income and health outcomes (Mwabu, 2009).

The common practice to address these problems is to search for instruments that are correlated closely with the endogenous variable(s) but are uncorrelated with the error terms in a systems' equation setting that also addresses the potential problem of simultaneity bias. In case where longitudinal data is available, it is also possible to decompose the error term into components that are easy to control using several techniques to reduce possible sources of endogenity. Since we use cross-sectional data in this study, the problem of endogenous regressors becomes naturally very problematic as the residual includes almost everything that is unexplained. To avoid some of the pitfalls indicated, first we decided to use a health outcome variable which is relatively easy to recognize by respondents and less likely to be contaminated by measurement error. Among the potential indicators of health outcome, episode of illness/infirmity in the last two months was reported by all respondents. The response is a dummy variable where individuals would indicate whether or not they were sick over the last two months and the number of days they had missed from performing their main duties. From a sample of close to 1000,000 individuals, nearly 22% reported to have fallen sick in the last two months and of these 70% was due to serious outbreaks such as malaria, diarrhea, tuberculosis and other intestinal infections. More than 50% of the sick had been bed-ridden for more than one week which is a substantial time loss for those in employment or attending school. Thus, these two variables may capture health outcomes well with a potential to be influenced by policy measures that are exogenous to the individual. Thus we specify a generic latent linear probability model given by:

$$H^* = X'\beta + u \quad \text{So that} \tag{8}$$

$$\begin{aligned} H &= 1 \quad if \quad H^* > 0 \\ &= 0 \quad otherwise \end{aligned} \tag{8'}$$

Where the probability of falling sick is conditioned on a vector of explanatory variables given by X. Some of the key variables entering equation (8) are human capital (E)-proxied by highest grade attained, access to health care (M) , wage (W)and age which is correlated with the rate of depreciation of health capital. Assuming the disturbance terms follow a normal distribution, our model takes a probit model that can be estimated by maximum likelihood method. We make attempt to control for the endogenity of some of the regressors using a set of instruments. Particular concern is the education variable which is clearly correlated with a host of unobserved individual attributes such as ability, school quality, etc. It is also plausible that health of the individual could very well influence the level of education attained. Future or current income as well could also be affected by current health shocks potentially creating a problem of reverse causality if wage or current income is used in the regression.

To obtain consistent estimates of health demand model, particularly with respect to education and income we follow the difficult but commonly implemented method of instrumental variable approach and simultaneous equation model using two-step procedures. Evidently, if one could find a variable that could influence health outcome only through education, then, it is straight forward to obtain consistent estimate on the effect or impact of education on health[2]. From the theoretical model, the level of education of the individual is assumed to enhance the efficiency of producing health, assuming other factors to be constant. The idea is that the higher the level of education the better the individual takes care of his/her health, etc. In the literature, distance to the nearest school is often used as a possible instrument for the level of education attained in which case a negative association between years of education and health outcome could only be explained through a causality running from education to health. Some of the

[2] Classic example in the empirical literature is the use of tax on tobacco as instrument for the link between smoking and lung cancer (see example Imbens, 2009)

criticisms on this approach are that one distance to school may not be entirely exogenous as it is possible that individuals may choose to live in areas where schools are available nearby. Second, distance to school and health outcomes may be jointly determined by other factors such as government decision to set up schools and health care services simultaneously in the same locality may be a cause for the positive association. More than the intuition however it is the weak explanatory power of distance to nearest school that is often the source of problem to serve as good instrument[3].

As a possible way of establishing some robustness, we also used structural feature of the demand model where education is assumed to be determined by individual and community characteristics, including health outcome and unobserved other factors. Thus, we followed a two-step procedure where in the first step we estimated a model of educational attainment as functions of a number of individual and community characteristics where in the second stage we used the predicted values in the health demand model.

3. Data and descriptive statistics

The data used in this study comes from the 2004/2005 nationally representative welfare monitoring survey that covered close to 21000 households with 100,000 individual histories. This is one of the most comprehensive surveys conducted in recent years in Ethiopia on welfare and related issues. The data is rich with detailed questions on household living standard, livelihood, access to basic services and other factors. It covers nearly all parts of the country except some pastoralist areas in Afar region. Table 2 summarizes the key variables used in this study.

Table 2: Summary statistics of key variables

[3] It is also important to mention some of the caveats discussed recently by Deaton (2009) on conditions to be met for an instrument to be exogenous. See also Heckman and Urzua (2009)

Variable	No of observations	Mean	SD
Basic demographics			
Household size	57880	5.56	2.61
Small agricultural holder	57880	0.40	0.50
Male	57880	0.48	0.50
Age	57870	32.73	15.64
Single	57863	0.37	0.48
Married	57863	0.50	0.50
Divorced	57863	0.05	0.22
Separated	57863	0.02	0.12
Widowed	57863	0.07	0.26
Attended formal education (proportion yes)	57869	0.54	0.50
Reasons for not attending school			
Not attending because of work	25994	0.07	0.25
Family too poor to send to school	25994	0.47	0.50
Shortage of money	25994	0.06	0.23
No school in my area	25994	0.16	0.36
Marriage	25994	0.10	0.30
Illness	25994	0.01	0.11
Disability	25994	0.00	0.07
Learning has no benefit	25994	0.10	0.30
too young to learn	25994	0.00	0.04
Too old for learning	25994	0.02	0.14
Other	25994	0.02	0.13
Highest grade attained	32397	11.31	16.00
Employment status			
Employer	33120	0.03	0.17
Self employed in formal sector	33120	0.30	0.46
Self employed in the informal sector	33120	0.17	0.38
Formal sector employee	33120	0.07	0.26
Informal sector employee	33120	0.03	0.17
Public sector employee	33120	0.04	0.19
Government employee	33120	0.08	0.28
NGO employee	33120	0.01	0.10
Domestic worker	33120	0.04	0.19
Unpaid family worker	33120	0.22	0.41
Other	33120	0.01	0.09
Ever sick last two months, if yes type of sickness:	57855	0.23	0.42
Malaria	6235	0.52	0.50
Diarrhea	6235	0.11	0.31
Injury	6235	0.04	0.20
Dental problem	6235	0.06	0.24
Eye problem	6235	0.12	0.32
Skin disease	6235	0.04	0.19
Intestinal infection	6235	0.06	0.23
Tuberclousis	6235	0.06	0.23
Per capita daily calorie	57880	2304.30	876.03
Per capita consumption expenditure	57880	2293.66	4617.51
Proportion stunted	10732	0.44	0.49
Reason for not seeking medical help			

Variable	No of observations	Mean	SD
No need to consult	45561	0.86	0.35
Financial problem	45561	0.12	0.32
Expensive service	45561	0.005	0.07
Service too far	45561	0.01	0.11
Not confident with service quality	45561	0.007	0.04
Do not believe in medical remedies	45561	0.01	0.08
No qualified health worker	45561	0.002	0.04
Poor equipment	45561	0.001	0.04

Some of the demographic indicators are typical. Large family size (5.6), young population (mean 32 years) and largely uneducated labor force. The highest grade completed in the country is grade 11 with 46% having never been to any formal school in their life. Poverty is the main reason for failing to go to school. Close to 60% of individuals attributed lack of money as the main reason for not going to school. About 7% also said that they had to work instead of going to school probably to support family businesses. Access explains only 16% of the reason for failing to go to school. Family formation explains a significant proportion (10%) of avoiding school probably more for women particularly in rural areas. Other reasons include disability and sickness which affect school attendance. Negative attitude is also one of the important variable explaining aversions towards schooling (10%). This may be explained in many ways. In urban areas, the growing youth unemployment, particularly among the educated has led many to believe that it is really a waste of time and resources to go to school. Growing poverty and unemployment might have discouraged many from investing in building human capital. Other possible reasons include cultural values, religious beliefs, which may prejudices people from going to school.

Ethiopian labor market profile is typical of a poor agricultural economy. Self employment and unpaid family labor account 62% of employment in the country. Formal wage employment accounts only for a fraction of total employment thus livelihood for many is a risky affair which is vulnerable to several shocks. Because of low labor productivity in self-employment, it is no wonder that a major concern by households in Ethiopia is food shock. Nearly 16% of households responded as having faced serious food shortage during the survey period and for the majority food aid (40%) and distress

sell of cattle (34%) are important remedial mechanisms suggesting the weakness of local institutions in addressing major shocks as basic as food shortage. It is very clear from Table 2 that a majority of the Ethiopian population are malnourished, with 2300 mean calorie intake, which is very close to the 2100 calorie used to draw the food poverty line. In fact, some studies put the calorie requirement to 2400 calories in which case the absolute food poor could reach as much as 65% of the population. It is no wonder that 44% of respondents were reported to have been stunted, a clear evidence of major health hazard in the country.

Sickness is a frequent presence among the Ethiopian population. Close to 23% of individuals in the survey reported as having been ill last two months. The main causes are easily preventable diseases such as malaria, diarrhea and intestinal problems. Among the sick, few seek health care, a majority (59%) due to financial problems. Distance to health center is also important to a certain extent accounting for 7% of the reasons for not consulting health professional. From this description, it is clear that there is a strong correlation between poverty, bad health and demand for health care.

4. Discussion of results on determinants (correlates) of demand for health in Ethiopia

We estimated equation (8) for three alternative indicators of health status of the individual. The first variable we chose to use as an indicator of health status in Ethiopia is self-reported episodes of illness that had occurred in the last two months of the survey. Evidently, this variable reveals a lot about the current health status of an individual though it should not be regarded more than merely a snapshot of the bigger picture. But, in the context of Ethiopia, this variable has a potential to capture health burden borne by individuals and could be a good indicator of overall wellbeing and vulnerability to epidemics such as malaria, tuberculosis, and HIV/AIDs. Illness episodes could be chronic, which is a better measure of life-cycle effects of health demand, or transitory that could prevail for any number of reasons without necessarily varying with individual socio-economic characteristics. Thus, it may not be straightforward whether or not one

should expect significant effect of individual characteristics. As a supplement to episodes of illness, we also selected the number of days lost in the last two months due to illness as one of our indicators of health status. This indicator tends to be sharper in discriminating the most chronic from the transitory illness episodes and might certainly complement the first indicator. Finally, we used stunting, one of the commonest indicators of health status as one of our dependent variables to examine the relative roles of variables laid out in the literature. As determinants of health outcome, we followed the specifications in equation (8) where the key variables of interest are human capital, proxied by the highest grade attained by the individual; health care inputs, proxied by affordability, distance or access to nearest health care facility, and overall attitude towards the health care facilities; and finally income which is proxied by the sector of employment of each individual. We have also used total real consumption expenditure as alternative proxy for wages or income and the results remained unchanged. We restricted our sample to the working age population (between 15 and 65) in the case of illness episodes and number of days in bed to reduce the problem of missing values such as employment status, etc. and also to conform with the theory that infirmity or sickness has an opportunity cost in terms of working hour lost or school attendance that one way or another affects income. For stunting, we used the whole sample since children were also included in the observation. Also, the number of observations falls drastically when we restrict our sample to those in the working age group only.

Table 3 summarizes the key findings from the estimation of the probit model using a simultaneous equation set up where predicted educational attainment is used from the first stage regressions including health outcome itself[4]. We reported results separately for rural and urban areas. The results are quite interesting. Demographic factors seem to be correlated with health outcomes. In both urban and rural areas, male individuals have lower probability of falling sick than females, older people tend to get sick and families with large size have lower probabilities of falling sick. Some of these results could be capturing unequal intra-household distribution of calorie consumption (gender),

[4] Because education is a continous variable, we used two-step estimation procedure where Smith and Blundell (1984) derivation is used to test for the weak exogenity of the education variable.

depreciation of health due to passage of time (age) and survival at early age (large families) or other life-cycle effects.

Table 3: Marginal effects of Probit estimates of the determinants of episodes of illness in Ethiopia (schooling approximated by auxiliary first stage regression)

Variables	Rural areas		Urban areas	
(Dependent is dummy if individual was sick last 2 months)	coef	pval	coef	pval
Demographic & consumption characteristics of the individual				
Sex of individual is male	-0.007087***	[5.78e-05]	-0.019668***	[3.35e-05]
Age in years	0.000552***	[0]	0.001565***	[0]
Household size	-0.000971***	[0.00689]	-0.004040***	[0.000332]
Log per capita consumption expenditure	0.001258	[0.350]	0.014252***	[0.00879]
Interaction b/n consumption & schooling	0.000506**	[0.0168]	-0.00115	[0.435]
Schooling (Residuals from first-stage regression)	-0.003963**	[0.0146]	0.013345	[0.212]
Marital Status (Single is reference group)				
Individual is currently married	-4.9E-05	[0.982]	0.019171***	[0.00357]
Individual is divorced	0.002855	[0.479]	0.010631	[0.409]
Individual is separated	0.006307	[0.307]	0.012892	[0.531]
Individual is widowed	0.004701	[0.267]	0.007676	[0.526]
Reasons for not seeking modern health care services (reference is no need to consult)				
Financial incapability	0.422289***	[0]	0.584110***	[0]
Expensive service	0.337564***	[7.43e-10]	0.539795***	[0]
Service too far	0.386589***	[0.000462]	0.623292***	[0]
Not confident with the quality	0.450467***	[1.55e-07]	0.611455***	[0]
Do not believe in medical treatment	0.610149***	[0]	0.751449***	[0]
Lack of qualified health personnel	0.541179***	[2.06e-09]	0.688086***	[0]
Poor service/equipment	0.409899***	[3.75e-05]	0.740174***	[0]
Regional dummies (Tigrai region is reference group)				
Afar	-0.00446	[0.405]	-0.00981	[0.495]
Amhara	0.010377*	[0.0545]	0.017965*	[0.0732]
Oromia	0.002903	[0.511]	-0.00108	[0.905]
Somale	0.017698**	[0.0175]	0.012952	[0.315]
Benishangul_gumu	0.033648***	[0.00219]	0.040639***	[0.00497]

SNNP	0.015580**	[0.0196]	0.023619**	[0.0189]
Harari	0.087670***	[1.26e-06]	0.087745***	[0.000144]
Addis Ababa	-0.00453	[0.226]	-0.00064	[0.966]
Dire dawa	0.036410***	[0.00144]	0.040668**	[0.0390]
Likelihood ratio	-3360.86	-4841.91		
Smth-Blundell test of weak exogenity (p-values)	0.946		0.2164	
Observations	26676		18765	

The effect that household size may have on overall incidence of infirmity from Table (3) sounds counterintuitive since it implies that the larger the size of the family, the lower the disease burden. Some of the possibilities for this may be that there are omitted variables correlated with household size that may have positive influence in reducing disease burden. The other possibility may be that there is a threshold effect whereby disease burden declines as the number of people in the household increase and tends to rise afterwards as the household grows to be crowded. We found such non-linearity in both specifications in Table 3 (not reported), so that after a certain threshold large family size leads to high incidence of illness. Similarly, men tend to have lower reported illness compared to women which is consistent with what is commonly observed about health hazards women face particularly in poor countries.

The effect of per capita consumption expenditure in both rural as well as urban areas is reflective of the measurement error problem raised in the preceding section (Mwabu, 2009) where richer households tend to be sensitive in reporting health status as compared to poor households. Despite the fact that we used interaction terms to capture the measurement error problems it has a negative and significant effect on health outcome in urban areas.

Variables that were used to capture access to health care system as self-reported by respondents tend to be strongly correlated with episodes of illness in both rural and urban areas[5]. It is important to note that the self-reported visit or failure to visit health facilities

[5] These variables are used only for decomposition purposes as they have strong correlation with the dependent variable by construction. The results remain unchanged even when we used distance to health care centers as proxy for health inputs (see Annex Table 1)

were not conditional on being sick in the last two months. The exact sequence of questions asked were as follows: first respondents were asked "did you visit any of health facilities in the last two months?" and immediately they were asked "reasons for not consulting health service". Most responded by saying they did not need to go to health facilities apparently because they were not sick. The other reasons were as reported in Table 3 and imply that some or most of them had suffered some form of illness during the period. We may safely assume however that with or without illness, for reasons reported in Table 3 most people who suffered from some form of illness did not get treatment. In a dynamic setting it thus implies that lack of access to health care exposes individuals to frequent attack of illness, such as for instance malaria, diarrhea, or intestinal infections, which are the most important factors behind self-reported episodes of illness. It is interesting to note that among the reasons outlined for not accessing health facilities, individual aversion towards modern medical facilities tend to have a much stronger impact on the probability of falling sick. To examine directly the effect that access may have on illness, we used distance from the nearest health post, clinic and hospital for the individual as an alternative for the self-reported reasons for not accessing health facilities. Despite the acknowledged weakness of such proxies, we could see that in either rural or urban areas, accesses to health care services are important variables in affecting health outcomes (see Annex Table 1).

The other approach we used to address the endogenity of schooling is instrumental variable method where distance to the nearest primary and secondary school was used as instrument for the level of education attained by the individual. The results show that education plays a very important role in improving health status. Its effect in urban areas is insignificant. This is more or less consistent with intuition. In rural areas, preventive measures, such as personal hygiene and understanding of communicable diseases could help a lot in preventing health hazards. In urban areas, though education could be important in making health production efficient, the weakness of distance as an instrument may have led to insignificant correlation between health and schooling which is also what was found in the two-step procedure.

What is remarkable however about the instrumental variable approach is that per capita consumption expenditure is no more counter-intuitive in both rural and urban areas. Actually, in rural areas, the higher per capita consumption expenditure, the better health outcomes turn out to be. The interaction term has successfully captured the measurement error arising from systematic reporting biases between poor and non-poor households or individual. The other interesting aspect of the IV approach is that none of the variables that capture entry into modern health care system tend to be significantly important for health outcomes in rural areas.

The discussion above can be made slightly sharper if we consider the number of days lost due to illness as our indicator of health status or wellbeing. This specification allows us to explore chronic illness (the more the number of days a person is bedridden, the greater the importance of access to health care facilities) better than the simple dummy variable of illness episodes.

Table 4: Determinants of illness episodes in Ethiopia (Instrumental variable estimates)

Variables	Rural areas		Urban areas	
	Coef	pval	Coef	pval
(Dependent is dummy if individual was sick last 2 months)				
Demographic & consumption characteristics of the individual				
Sex of the individual is male	-0.01324	[0.834]	-0.14608***	[8.70e-07]
Age	-0.00085	[0.881]	0.009861***	[0.000648]
Log of real per capita consumption	-1.38039***	[0.000285]	0.459002	[0.640]
Interaction b/n consumption & schooling	0.181215***	[4.16e-05]	-0.218577	[0.743]
	-1.387102***	[3.09e-05]	1.581092	[0.747]
Marital Status (Single is reference group)				
Currently married	-0.17351***	[0.000249]	0.203743	[0.202]
Divorced	0.022773	[0.655]	0.175999*	[0.0613]
Separated	-0.03081	[0.751]	0.152294	[0.228]
Widowed	-0.02889	[0.676]	0.144499**	[0.0466]
Reasons for not seeking modern health care services (reference is no need to consult)				
Financial incapability	0.983619	[0.221]	1.915606**	[0.0426]
Expensive service	0.861924	[0.193]	1.548056*	[0.0506]
Service too far	0.991999	[0.170]	1.787670**	[0.0283]
Not confident with the quality	0.697385	[0.413]	1.698447	[0.122]
Do not believe in medical treatment	1.263226	[0.171]	2.182349**	[0.0401]
Lack of qualified health personnel	0.761222	[0.431]	1.894229*	[0.0690]
Poor service/equipment	1.268343*	[0.0661]	2.187779***	[0.00215]
Regional dummies (Tigrai region is reference group)				
Afar	-0.0957	[0.269]	-0.022581	[0.888]
Amhara	-0.00817	[0.943]	0.151173*	[0.0849]
Oromia	-0.04053	[0.531]	-0.002528	[0.962]
Somale	-0.09292	[0.573]	0.133045	[0.213]
Benishangul_gumu	0.079735	[0.714]	0.262444***	[0.00400]
SNNP	-0.01071	[0.940]	0.191656***	[0.00599]
Harari	0.364558	[0.301]	0.362883*	[0.0989]
Addis Ababa	-0.102565*	[0.0619]	-0.007738	[0.949]
Dire dawa	0.197116	[0.335]	0.268838**	[0.0287]
Afar	-1.387102***	[3.09e-05]	1.581092	[0.747]
Observations	26676		18765	
Loglikelihood ratio		-29189.5	-3375.6292	
Wald test of exogenity		0.1405	0.7533	

We used distance to nearest primary and secondary school as our instrument to estimate the model reported in Table (5). For the whole sample, we found that one year more schooling could reduce by about 49% the number of days that an individual could spend in bed due to illness. The effect of education on health remained largely positive and significant in rural areas, but again with no statistically significant impact in urban areas, which is what also we found in the preceding tables. In this set up, we used employment

status to control for possible confounding between schooling, employment and number of days absent from work due to illness. We notice also that per capita consumption expenditure now has a distinct and positive impact on the number of days an individual spent in bed due to illness. For other demographic characteristics the result remained similar and even much sharper.

Table 5: Determinants of number of days lost due to illness (Instrumental variable estimates)

	All sample		Rural areas		Urban areas	
VARIABLES	coef	pval	coef	pval	coef	pval
Demographic & consumption characteristics of the individual						
highest grade completed	-0.494024***	[0.000120]	-0.198399***	[0.00910]	-0.50362	[0.562]
Age in years	0.001349**	[0.0255]	0.000043	[0.940]	0.003931***	[1.74e-08]
Log of per capita consumption expenditure	-0.363251***	[0.000128]	-0.200207***	[0.00710]	-0.0986	[0.577]
Household size	-0.007821***	[7.65e-07]	-0.003252***	[0.00856]	-0.006307**	[0.0248]
Interaction b/n consumption & schooling	0.064038***	[0.000123]	0.025535***	[0.00944]	0.067955	[0.566]
Reasons for not seeking modern health care services (reference is no need to consult)						
Financial incapability	0.894874***	[0]	0.643952***	[0]	1.017386***	[0]
Expensive service	0.588729***	[1.27e-08]	0.21211	[0.103]	0.703771***	[4.17e-09]
Service too far	0.839850***	[0]	0.167977	[0.482]	0.873037***	[0]
Not confident with the quality	0.596232***	[0.00251]	0.290487	[0.163]	1.001823***	[0.000112]
Do not believe in medical treatment	1.093444***	[0]	0.359942***	[0.00843]	1.336537***	[0]
Lack of qualified health personnel	0.956323***	[4.44e-09]	0.723374***	[0.00571]	1.100077***	[1.35e-07]
Poor service/equipment	0.721866***	[4.20e-05]	0.380515	[0.100]	0.884975***	[0.000148]
Urban dummy	0.016431	[0.435]				
Marital Status (Single is reference group)						
currently married	-0.035363***	[0.00454]	-0.019297*	[0.0574]	0.000301	[0.992]
Divorced	0.001626	[0.923]	0.011341	[0.452]	0.033746	[0.289]
Separated	-0.00289	[0.930]	0.022034	[0.458]	0.08622	[0.166]
Widowed	-0.00623	[0.779]	0.004968	[0.823]	0.03407	[0.355]
Employment status (self employed (formal sector) is the reference group)						
Self employed in formal sector	0.008122	[0.725]	0.006541	[0.778]	0.012489	[0.670]
Self employed in the informal sector	-0.01816	[0.137]	-0.020262*	[0.0904]	0.025497	[0.374]
Formal sector employee	-0.00812	[0.542]	-0.01743	[0.109]	0.020926	[0.579]
Informal sector employee	-0.00655	[0.724]	-0.02477	[0.134]	-0.01508	[0.736]
Public sector employee	-0.068430***	[0.000140]	-0.033603***	[0.00404]	-0.125206**	[0.0448]
Government employee	-0.045688***	[0.00248]	-0.023734**	[0.0216]	-0.23523	[0.602]
NGO employee	-0.076327**	[0.0308]	-0.046773***	[0.00288]	-0.0717	[0.759]
Domestic worker	0.042255**	[0.0105]	0.004353	[0.749]	0.000619	[0.990]
Unpaid family worker	-0.035405**	[0.0226]	-0.00504	[0.779]	0.032707**	[0.0229]
Other	-0.100772***	[0.00467]	-0.075514***	[0.00266]	-0.03683	[0.647]
Regional dummies (Tigrai region is reference group)						
Afar	-0.01966	[0.331]	0.003481	[0.830]	-0.05456	[0.104]
Amhara	0.003113	[0.852]	0.007916	[0.602]	0.001272	[0.963]
Oromia	0.009972	[0.545]	-0.00176	[0.903]	-0.01776	[0.417]
Somale	0.041703*	[0.0608]	0.012455	[0.549]	0.092177***	[0.00916]

	All sample		Rural areas		Urban areas	
VARIABLES	coef	pval	coef	pval	coef	pval
Benishangul_gumu	0.034336*	[0.0977]	0.044598**	[0.0464]	0.021488	[0.468]
SNNP	0.067194***	[7.94e-05]	0.01579	[0.373]	0.048307	[0.114]
Harari	0.089289***	[0.000765]	0.017319	[0.414]	0.044856	[0.176]
Addis Ababa	0.030225*	[0.0844]	-0.00153	[0.906]	0.034932	[0.287]
Dire dawa	0.042236**	[0.0315]	0.044699**	[0.0295]	0.009465	[0.814]
Constant	2.731549***	[0.000149]	1.569704***	[0.00716]	0.639733	[0.628]
Hansen's J statistics for testing overidentification	0.1569		0.1662		0.0224	
Observations	26007		12712		13295	
R-squared	0.157713		0.150977		0.318472	

Finally we examine stunting as one of the important indicators of health outcomes in Ethiopia which affects close to 44% of individuals- see Table (2). Unlike episodes of illness, information on stunting is available only for 10732 individuals, which is 10% of the whole sample. In addition, when we attempted to use the same regessors as those in the previous models, the number of missing observations increased substantially, but with no significant effect on the underlying relationship between stunting and socio-economic characteristics of the individual. Table (5) reports the determinants of stunting in Ethiopia. Evidently, the simple probit estimate is biased and inconsistent due to the presence of endogenous regressors in our model as can be seen by the rejection of the hypothesis of endogenity by the Wald-test. Thus, instrumental variable method of estimation is indicated to get consistent estimates. The challenge however is to find valid instruments that are correlated with the educational achievement variable, but, uncorrelated with the residual terms.

Table 6: Determinants of stunting in Ethiopia

	OLS estimates		IV estimates	
	dy/dx	Z-value	dy/dx	z-value
Never been in formal school	0.114007	6.38	0.474724	7.34
Per capita calorie consumption	-1.1E-05	-1.55	-3.1E-05	-1.75
region2*	-0.01811	-0.61	-0.04584	-0.6
region3*	0.131862	6.52	0.329709	6.42
region4*	-0.03558	-1.88	-0.09295	-1.92
region5*	-0.02886	-1.15	-0.07552	-1.17
region6*	-0.0197	-0.74	-0.05013	-0.74
region7*	-0.00295	-0.15	-0.0102	-0.2
region8*	-0.08538	-2.84	-0.22408	-2.79
region9*	-0.10818	-4.89	-0.26113	-4.36

region10*	-0.13498	-4.64	-0.35731	-4.36

*Marginal effect is for discrete variable from 0 to 1.
Instruments include age, household size, residence in rural areas, access to schooling (affordability, distance, attitude, and other factors such as marriage, etc.)

We followed the human capital literature that routinely uses supply side determinants to address the problem of endogenous schooling variables (Card, 2001). That is, to find variables that capture the availability of schooling opportunities to the individual, which include distance to the nearest school (proxy for access), parental background (affordability and also a taste for education) and other factors such as unemployment rate in areas of residence, etc. In our case, we used variables which could be correlated with schooling but are exogenous to the individual, such as age, regions of residence, and availability of schooling in the area of residence. Other quasi-endogenous variables include size of the household, attitude or taste for education perhaps borne out of cultural and societal beliefs, employment status, etc. The use of these variables as instruments for schooling pushed upwards the coefficient for schooling in determining stunting. This is again consistent with findings in the large empirical work where OLS tends to underestimate returns to schooling in the wage equation. The same finding is also reported for health demand functions (Grossman, 2004). Actually our result for the education variable even become higher and remained robust when we used purely exogenous instruments for education, such as age of the individual and distance in kilometers of primary and secondary school from the residence. Thus, it may be safe to assume that health outcome, as measured by stunting, is highly correlated with the educational achievement of the individual. The issue now becomes that of causality, which according to Grossman (2004) such a finding can be interpreted in three possible ways. One is that education increases or improves health outcomes, or secondly it is good health that promotes education (reverse causation) or third, there may be no causation at all, but, both variables may be affected by other factors such as parental background, etc. Evidently the issue has attracted a large empirical work in recent years, with the weight of evidence pointing towards the causality from schooling to health. In our specification, the tendency for schooling achievement to influence health outcomes sounds a reasonable direction of causation. For instance, the coefficient on the IV-based estimate for the schooling variable remains robust for a specification where we used distance from health

centers (health post, clinics and hospitals), distance from schools (primary and secondary), age of the individual and regions of residence as explanatory variables and instruments for schooling variable.

Stunting is also determined by the daily per capita kilo calorie consumed by the individual, an indicator that may be rather more relevant for determining wasting than stunting. Perhaps this could be because snapshot calorie consumption consists of a substantial long-term component so that perhaps little has changed for many over time.

So far our discussion focused on health outcomes conceived by the individual purely as an investment good, a means for achieving perhaps better income, and improved livelihood. However, health is also consumption good where consumers derive higher utility from its increased availability. Thus, a direct demand for health consumption can provide useful policy insight, through the empirical application is complicated by lack of good data in developing countries, particularly data on consumption of a wide range of health services and their relative prices (see Ajakiye and Mwabu, 2007b). Here we briefly motivate an approach where one can recover income and price responses from a cross-sectional data when price data on health services are not available and illustrate using Ethiopian data.

A convenient utility function that allows recovery of price responses from income elastcities is the Linear Expenditure System which is extensively used particularly in calibrating consumption behavior in CGE frameworks, which is specified as follows:

$$p_{it}x_{iht} = p_{it}\gamma_i + \beta_i(y_{ht} - \sum_{k=1} p_k\gamma_k) \qquad (9)$$

Where p_{it} is price of commodity i prevailing at period t, x_{it} is quantity of i demanded by household h at period t, y_{ht} is total income of household h at period t and γ_i and β_i are parameters to be estimated, representing respectively the "subsistence" consumption of commodity i, and β_i is the marginal budget share. The structure of the LES is motivated

by the assumption that regardless of income levels, each household allocates its income first on subsistence goods and the remaining is driven by consumption preference. Estimation of (9) is complicated by the non-linear term linking marginal budget share with the "supernumerary" income or consumption expenditure so that a numerical approximation is used in the context of non-linear system of equations. Despite some limitations, the LES provides a simple framework to capture the welfare implications of changes in relative prices. Estimation of (9) from one cross-section data can be made using additional information on consumption decision, such as savings (e.g Howe, 1975; Lluch, 1974). The price responses are linked with the income responses through what is the known as the inverse Frisch parameter or share of the subsistence consumption to supernumerary income, or marginal utility of income:

$$E_{ii} = \phi E_i - E_i V_i (1 + \phi E_i) \qquad (10)$$

If income elasticity on health expenditure is about 0.7 (see Figure 1 for Ethiopia) which indicates that health is a necessity, with a Frisch parameter around 0.6027 (see Shimeles, 2007). From (10) own price elasticity for health is services would be around 0.38549. The fact that health services are both income and price inelastic in the Ethiopian case may not be surprising given the limited choices available and its importance to survival and quality of life to consumers.

5. Conclusions

The main objective of this paper is to explore the determinants of selected indicators of health status in Ethiopia using the large welfare monitoring survey conducted in 2005. Some of the key determinants have been motivated and identified from the established results in the theory of health demand which conceives health services as investment as well as consumption good. Three indicators of health status are used in the analysis. The two indicators deal with episodes of illness and the number of days lost due to illness. The other is stunting which usually is measured as inhibited growth usually caused by malnourishment in childhood. The main result that came out of our analysis is that

variables such as schooling, age, household size, marital status, employment and place of residence vary considerably with the health indicator, particularly in rural areas. Apart from socio-economic factors, access to basic health services seem to play a very important role in affecting the three measures of health outcomes. The condition of health status improves with schooling achievement only in rural areas, which could possibly be due to the weakness of the variable we used as instrument for educational attainment. The possibility that schooling determines health is evident from the robustness of the schooling variable in the health demand model when we use variables that potentially influence only schooling not health as instruments. But, there is no denying the fact that health is important for continuing education and other achievements in life, such as higher income. In our data, some individuals failed to go to school because of disability caused by illness and related factors.

Health indicators deteriorate with larger household size, age and residence in rural areas. For stunting, apart from the known individual characteristics, we found per capita calorie consumption to be an important determinant. The importance of calorie consumption for incidence of stunting reflects the possibility that profile of current calorie consumption might have changed little over time.

Preliminary results on price and income responses for health demand suggests that health services are demand inelastic with respect to prices and income that may not be surprising given the limited choices available and the desire to maintain health by consumers.

References

Ajakaiye, O and Mwabu, G (2007a), "The demand for reproductive health services: an application of control function approach", AERC mimeo

Ajakaiye, O and Mwabu, G (2007a), "The demand for reproductive health services: framework for analysis", AERC mimeo

Card, D, (2001), "Estimating the return to schooling: progress on some persistent econometric problems", Econometrica, 69(5):1127-1160

Deaton,A.S. (2009), "Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development", Working paper, National Bureau of Economic Research

Grossman, M, (2004), "The demand for health after three decades", Journal of health economics

Grossman, M., (1999), "The human capital model of the demand for health", NBER Working Paper 7078

Grossman, M, (1972), " On the concept of health capital and the demand for health", Journal of Political Economy 80: S74-S103

Heckman, J and S. Uzrua (2009), "Comparing IV with structural models: what simple IV can and cannot identify" NBER Working Paper 14706.

Howe, H., (1975), "Development of the extended linear expenditure system from simple saving assumptions", European Economic Review, 6: 305-359

Imbens, G. U. (2009), "Better late than nothing: some comments on Deaton, Heckman and Uzrua" NBER Working Paper 14896.

Lluch, C., (1974), "Extended linear expenditure system", European Economic Review, 4, 21-32

Mwabu, G, (2009), "Health economics for low income countries" in Handbook of development economics, ed (H. Chenery & T.N. Srinivasan), ch 53

Shimeles, A., (2007),"Inflation and the poor in Ethiopia: a demand systems approach", World Bank, Washington DC, mimeo.

Annex Table 1: Determinants of illness episode in Ethiopia (simple probit with schooling approximated by auxiliary first stage regression-Marginal effects)

	Rural areas		Urbban ares	
VARIABLES	coef	pval	coef	pval
Demographic characteristics and consumption expenditure				
Sex of the individual is male	-0.035046***	[0]	-0.046129***	[0]
Age	0.003076***	[0]	0.004629***	[0]
Household size	-0.008506***	[0]	-0.008306***	[2.21e-08]
Log of per capita consumption	-0.022085***	[0.00111]	0.021964**	[0.0466]
Interaction b/n consumption & schooling	0.001449*	[0.0599]	-0.00021	[0.904]
Highest grade achieved	-.0131761**	[.00571]	.0005097	[0.968]
Marital status (single is reference group)				
currently married	0.035335***	[1.02e-09]	0.047698***	[5.55e-08]
Divorced	0.048735***	[1.48e-05]	0.073084***	[7.29e-05]
Separated	0.087097***	[2.43e-06]	0.088687***	[0.00501]
Widowed	0.084806***	[0]	0.037606**	[0.0292]
Distance to health center & consumption exp	-0.00025	[0.479]	0.00004	[0.935]
Distance to clinic & consumption exp	0.000003	[0.981]	-9.1E-05	[0.750]
Distance to health post & consumption exp	0.000137	[0.265]	-0.000443**	[0.0265]
Distance from health post	-0.00093	[0.318]	0.003346**	[0.0207]
Distance from nearest clinic	-0.00032	[0.764]	0.000444	[0.828]
Distance from nearest health center	0.000545***	[0.000455]	-0.00027	[0.256]
Regional dummies (Tigrai is the reference region)				
Afar	0.099707***	[5.35e-09]	0.156396***	[0]
Amhara	0.030584***	[0.00597]	0.018935	[0.125]
Oromia	0.046336***	[4.17e-05]	0.037353***	[0.00220]
Somale	0.003447	[0.799]	0.023923	[0.190]
Benishangul_gumu	0.146256***	[0]	0.184595***	[0]
SNNP	0.084061***	[1.00e-09]	0.058199***	[3.73e-06]
Harari	0.095746***	[4.91e-07]	0.047104**	[0.0285]
Addis Ababa	0.011002	[0.277]	-0.108647***	[0]
Dire dawa	0.063019***	[0.000218]	0.030192	[0.177]
Likelihood ratio	-15064.587		-13429.169	
Smith-Blundell test of weak exogenity		0.107		0.9803
Observations	33721		23466	