SSOAR Open Access Repository
www.ssoar.info

Gender patterns in household health expenditure allocation: a study of South Africa

Kingdon, Geeta; Irving, Margaret

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Empfohlene Zitierung / Suggested Citation:
Kingdon, G., & Irving, M. (2008). *Gender patterns in household health expenditure allocation: a study of South Africa.*. London. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-69452

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by/4.0/deed.de

Terms of use:
This document is made available under a CC BY Licence (Attribution). For more Information see:
https://creativecommons.org/licenses/by/4.0

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Gender patterns in household health expenditure allocation: A study of South Africa

by

Margaret Irving

Geeta Gandhi Kingdon*

October 2008

Abstract

This paper explores the extent and nature of gender differences, by age, in household health expenditure allocation. Using South African data, we adopt a hurdle methodology, constructing a sequence of decision stages (reporting sickness, consulting medical practitioner, incurring positive medical expenditure, and the conditional amount of expenditure) in order to examine all these possible channels of gender differentiation. Our results provide evidence of significant pro-*female* bias among prime age persons (ages 16-40) after controlling for gender differences in the opportunity cost of time spent on seeking medical attention. We infer that expenditure on female health is viewed as an important investment in household welfare in light of women's contribution to household production, particularly over child bearing/rearing ages. This provides an alternative narrative to the 'investment motive' hypothesis traditionally employed to explain differential allocation of resources to males and females within the household. We also compare the relative explanatory power of household and individual level equations in revealing intra-household gender bias. Our findings suggest that the dimensions of gender differentiation are revealed more clearly in individual level regressions.

JEL classification: I31, I38

Keywords: Health expenditure, gender bias, hurdle models, South Africa

Author contact details:

*Corresponding author: Professor Geeta G. Kingdon, Institute of Education, University of London, 20 Bedford Way, London WC1H 0AL. Tel: 020 3073 8316; Fax: 020 3073 8316; Email: g.kingdon@ioe.ac.uk

1. Introduction

The `investment motive' hypothesis maintains that resources are allocated to household members according to their expected returns in the labor market. This has been used as a prominent explanation for gender differences in household allocation decisions. For instance, Pitt, Rosenzweig & Hassan (1990) argue that the relatively higher levels of calories allocated to males than females within the household reflect men's participation in energy-intensive labor market activities. Rosenzweig and Schultz (1982) find that children who are projected to be more economically productive adults receive a larger share of family resources. Rose (2000) finds that the birth of a son in poor households in rural India leads to a substitution of the mother's time from other productive activities into childcare. The fact that this does not hold for the birth of a daughter suggests a greater perceived `investment value' in male children, due to factors such as dowry payments for girls and higher labor market returns for men. However, there are several weaknesses in a simplistic interpretation of the investment motive hypothesis. This paper examines the validity of the investment motive hypothesis in the context of household health expenditures, and explores alternative explanations for observed gender patterns in within-household health expenditure allocations.

While the investment motive constitutes a plausible explanation for gender differences in health expenditures within households, there are several other candidate explanations that have arguably received less attention in the literature. Intra-household gender differences in health expenditures could be driven by gender differences in the value of male and female health in household production (in particular, child bearing and rearing), the opportunity cost of foregone wages when consulting a doctor, and possible inter-temporal substitution both over an individual's life cycle and between generations.

In this paper, we examine the age structure of the gender gap in household health expenditure allocation using survey data from South Africa, to obtain a more nuanced understanding of household decision-making processes. Specifically, we ask whether the investment value hypothesis applies at certain ages but not at others, given the differing roles and perceived contributions of men and women at different ages. It is possible that women of childbearing and maternal ages receive higher health expenditures relative to their (higher-earning) male counterparts if women's non-market work is valued and if the opportunity cost of their time (for consulting a medical practitioner) is lower than that of men. Such favourable treatment of women vis a vis men may occur even while there is discrimination in curative expenditure against (young, unmarried) girls compared to boys, which may occur due to the family's dynamic planning problem and the relatively negligible contribution of girls to parental household welfare post-marriage in patrilineal societies.[1] More generally, we test whether sickness

[1] Where women migrate to their husband's household following marriage.

reporting, medical consultation rates and health expenditure among prime-age adults differ from those at younger and older ages. Finally, we examine whether the methodology using household-level data (the 'Engel Curve' approach) and that using individual-level data, are equally effective in detecting gender differentiated treatment within the household.

A paper close in spirit to ours is on China by Gao and Yao (2006). The authors find that while it is supported in the young age groups, the 'market value' hypothesis is rejected in the prime age adult groups since, in this age range, men receive lower health expenditure despite having higher earnings than women. Our paper differs importantly from Gao and Yao (2006) in that we use a hurdle model method approach, as described below. We argue that this approach is more helpful than the unconditional equations approach adopted in Gao and Yao because it is more effective in revealing the channels through which gender differentiation takes place in healthcare behavior. We construct a sequence of healthcare decision stages and employ a hurdle model methodology for estimation in order to allow for the possibility that bias may act in opposing directions at different stages. We also consider the extent to which gender differences may be usefully examined using household level data by comparing the results from an Engel curve household analysis to those from regressions run at an individual level. Our findings reinforce the work of Kingdon (2005) who finds that aggregation of data at the household level makes it more difficult to capture the full extent of gender bias. Finally, unlike Gao and Yao, we also employ household fixed effects analysis.

2. Gender and health: some evidence and our hypotheses

A priori, we anticipate that bias against girls in the South African context should be less evident than in, for instance, South Asia[2]. However, studies of the gender dimensions of medical care in the South African context are limited. Case and Deaton (2002) include a brief consideration of health expenditure as part of a broader study of consumption patterns and gender in South Africa. They use data from the 1995 Household Income and Expenditure Survey, exclusively examining black and 'coloured' households and adopting the parsimonious specification outlined below:

$$x_h = \sum_{j=1}^{J} \varphi_j n_{jh} + \varepsilon_h \qquad (1)$$

[2] One explanation for women's higher status in Africa than South Asia is Africa's far higher land to labor ratio which implies that women are more valued there as they provide an extra pair of (productive) hands. In South Asia, women may be seen as an extra mouth to feed due to the smallness of plots and surplus labor. Wood (2002) shows that the land to labor ratio in Africa is 14 times that in South Asia (where there is only 0.5 square km of land per 100 adults).

where x_h is medical budget share for household h, and n_{jh} are the numbers of people in household *h* in each of *J* age-gender groups.[3] Estimating φ in this model, Case and Deaton fail to find any evidence of systematic gender differences in medical expenditure, though there is some evidence that middle-aged men (ages 36 - 55) are favored over women in the same age category. Given the higher earning power of men, they invoke an economic argument, citing men's higher ability to pay rather than arbitrary discrimination as the cause of this differential.

Indeed - there is evidence from other spheres to support the notion of minimal gender disparity within South African households. Drawing on the same dataset as that we employ in this study, Ray (2000) consistently fails to find evidence of gender bias in household expenditure across a broad range of goods. More specific examples may also be cited. Parents may have a preference for male children on the grounds that sons are more likely to sustain familial and kinship ties following marriage, thus augmenting household income and providing old age parental support. Gangadhara and Maitra (2003) document the widespread literature on this phenomenon in Asian countries, but find limited evidence of son preference in the South African context, apart from in the Indian community where they argue that preservation of cultural norms such as the payment of dowry[4] means daughters are more likely to be regarded as a `burden'. While interesting, Indian households form less than 3% of our sample.

We may also consider education: 1996 enrollment rates for children aged 5 - 15 years were marginally higher for girls, while educational attainment for those under age 25 similarly reflects a small pro-female advantage (Africa, Budlender & Mpetsheni, 2001). Of course this should not be taken as a definitive picture. Wittenberg (2005), for example, finds that adolescent girls bear a disproportionate burden of domestic responsibilities as compared with their male counterparts.

Given our interest is in curative health care, we observe four sequential mechanisms through which gender differences may occur in medical expenditure: (1) women may report sickness with differential frequency than men; (2) conditional on reporting sick, women may be differentially likely to consult a healthcare practitioner than men; (3) conditional on reporting consultation, women and men may differ in their likelihood of incurring positive medical expenditure[5]; and (4) conditional on positive expenditure, average medical expenditure may be different for men and women. We highlight three possible explanations of any observed gender differentials in

[3] The authors distinguish ten key categories, namely males and females in each of five age cohorts: 0 - 5, 6 - 15, 16 - 35, 36 - 55 and older than 55.

[4] In traditional black households the system of *lobola* or 'bride price' reverses the nature of this payment i.e. the payment goes from the groom's to the bride's parents.

[5] While this stage may appear unnecessary, consultation does not imply positive expenditure for a substantial percentage of our sample - 19% of those who report seeking healthcare advice also report zero medical expenditure, most likely due to the use of state-subsidised care.

morbidity, consultation practices and medical expenditure:

1. *cultural or social factors lead household resource allocation processes to favor male over female members, or vice versa*

This explanation encompasses gender bias driven by household `investment motive' as well as a number of alternatives. In a review of the literature, Case & Paxson (2004) highlight that women tend to report illness with higher frequency and make greater use of healthcare facilities than men, even though they suffer lower mortality risk at every age. One explanation for this apparent paradox is that women are objectively healthier than men, but suffer from a relative bias in their perceived health status due either to their being less stoical than men, or to a greater awareness of their own state of health. If it is the case that women are more predisposed to care for their health, then for the same given severity of illness, they will be more likely than men to report ill for extended periods in order to facilitate recovery, as well as be more likely to consult a healthcare practitioner. Similarly, the higher likelihood of women to engage in preventative behavior (such as taking vitamins or washing hands) translates into improved health status (Waldron 1984). Lifestyle factors may play a role, with higher male mortality and morbidity often attributed to their higher rates of smoking and drinking, and greater exposure to occupational hazards. Evidence also suggests that men have a higher propensity to engage in risky behavior (*ibid*). Finally, the measurement of morbidity may be underestimated for men relative to women due to more frequent proxy reporting on behalf of men in the collection of household survey data.

2. *there may be biological differences in terms of susceptibility to illness or in the nature of illness to which men and women are particularly vulnerable, with these differential susceptibilities varying by age*

An alternative posited by Case and Paxson (2004) to explain the gender puzzle outlined above, is that there are gender differences in the distribution of chronic conditions. Women may be more likely to suffer from minor ailments such as migraines that result in poor self-assessed health but are relatively unlikely to contribute to mortality risk. In the South African context, Puoane *et al* (2002) find levels of obesity among women to be almost double those of men (56.6% of women and 29.2% of men were recorded as `overweight' or `obese'), exposing them to a range of health risks, including cardiovascular disease, musculoskeletal problems, and respiratory ailments. Furthermore, the demanding reproductive functioning of women is cited as a reason for higher illness reporting and frequency of medical consultations among women in Western countries (Waldron 1983). However Strauss *et al* (1993) find somewhat contradictory evidence on this issue: women who have given birth do appear to have relatively more difficulty with vigorous activities, but having children has no impact on the probability of suffering more severe health conditions, and also appears to have a positive correlation with higher self-reporting of overall good health. They conclude that the role of fertility in explaining gender differences in health status is most likely small.

3. *potentially the most compelling explanation in the South African context may be that economic imperatives are the source of behavioral differences.*

In particular, expenditures are affected both by an individual's needs and her ability to pay in order to meet them. We note that healthcare has an opportunity cost in terms of time and direct cost in terms of money. Thus, if women are less likely to form part of the labor force (as is the case in South Africa, see Table 1), on average they face a lower opportunity cost in reporting sick and seeking medical consultation.

[Table 1 about here]

Additionally, since women's wages are on average lower than men's (Table 1), they are less able to pay for medical care conditional on reporting sick, especially if incomes are not pooled – or only partially pooled – within the household, as is often the case in Africa (Haddad, Hoddinott and Alderman, 1997). On these grounds we would anticipate that medical expenditure would be lower both for women, and in households where women comprise a larger proportion of the employed members. Furthermore, employed women are disproportionately represented in domestic work and in the informal sector. Minimal health and safety regulations exist in these non-unionized sectors of the economy, and employer subsidies for health insurance are rare.

As outlined in our motivation, we also note that there may be some implicit recognition of women's non-market contribution to household welfare. Particularly over the child-bearing and maternal periods, women play an important role in fostering the well-being of the next generation. Thus while women may add less to the household monetary resource base than men, expenditure on women's health is arguably an important household investment.

3. Overview of data

We use the South African Integrated Household Survey (SAIHS) of 1993, the first survey to cover the entire South African population, sampling 43,984 individuals in 8,854 households. The survey included both a comprehensive household questionnaire and a community questionnaire, the latter largely relating to the availability of facilities and infrastructure.

We split the sample into five age cohorts, designed to represent the young (age 0 - 5); an intermediate juvenile group (age 6 - 15); a prime age working group (age 16 – 40) which also incorporates women of child-bearing age; a middle-aged working group (age 41 - 64); and the elderly (those 65 and older). Summary statistics are presented in Table 2.

[Table 2 about here]

Across the pooled (all ages) sample (lower panel of Table 2), women are 1.82 points (or 27%) more likely to report sick than men and, conditional on seeking medical consultation, are also statistically significantly more likely to incur positive medical expenditure. Consultation rates and conditional medical expenditures are roughly similar across genders, when we take all ages pooled. Striking results emerge on more detailed examination by age group (upper panel of Table 2). Incidence of reporting illness is U-shaped in age for both men and women: it is high at young age, falls sharply for the intermediate juvenile group and then increases with age. However, the upward sloping part increases with age more rapidly for women than for men. Thus, the gender gap in the chances of reporting sickness grows with age, and becomes statistically significant from age group 16-40 onwards. However, its very large size in the 65+ age group is partly the result of demography and biology. It is well known that women have higher life expectancy than men, so it is not surprising that women in this age group are, on average, 0.8 years older than men, a statistically significant difference. There is no significant gender difference in the chances of consulting a medical practitioner, conditional on reporting sick. However, in the 16-40 and 41-64 age groups women are significantly more likely to incur positive medical expenditure conditional on reporting sick and consulting a medical practitioner. Finally, there is little gender difference in medical expenditure conditional on incurring positive expenditure, except weakly in the 41-64 age group, where men's medical expenditure is higher than women's at the 10% level of significance (both in levels and as a percentage of household per capita expenditure).

We turn next to examining the relationship between gender and health expenditure after controlling for observable heterogeneity. Individual level regressors include gender, age, years of education and employment status, the last being included as dummy variables for unemployment, casual employment and regular employment, the base category being those out of the labor force.[6] A number of household variables were also included: household per capita expenditure, household size, dependency ratio (ratio of members aged 0-14 and over 64, to members aged 15 to 64) and gender of the household head. We also include a number of community characteristics in some of our specifications. Summary statistics for included variables are reported in Table 3.

[Table 3 about here]

4. Empirical methodology

In order to investigate our central research question, two models are compared. The first is a regression of unconditional medical expenditure *M* on a set of explanatory variables **x**, with those not reporting sick assigned a

[6] Individuals younger than 16 and older than 64 are categorised as out of the labor force. The `unemployed' dummy includes all those who claim to want work and are not currently employed, regardless of whether they are actively seeking employment, i.e. we use the 'broad' definition of unemployment, which is argued to be the more appropriate definition in South African conditions of high unemployment (Kingdon and Knight, 2006).

value of zero for M. However, as there is essentially censoring of the healthcare expenditure data at zero, the distribution of the dependent variable is not normal, thus the resulting estimates are likely to suffer from bias.[7] The conventional response to censoring is to use a standard Tobit model in estimation. However a significant limitation of this approach is that it presupposes that a single mechanism drives both the choice of positive expenditure, $P(M>0|\mathbf{x})$, and the choice of how much to spend conditional on positive medical expenditure, $E(M|\mathbf{x}, M>0)$. That is, it presupposes that the derivatives $\delta P(M>0|\mathbf{x})/\delta x_j$ and $\delta E(M|\mathbf{x}, M>0)/\delta x_j$ have the same sign, which may be a strong restriction to impose. An alternative is to use a hurdle model, which essentially separates out these two decisions. However, as noted, in the case of healthcare, there are a number of decision stages preceding determination of the level of expenditure. The hurdle approach outlined in Wooldridge (2002: 536 - 7) is therefore extended to allow for a four stage decision model:

1. does an individual report being sick (S=1 or S=0)
2. conditional on having reported sick (S=1) , does the individual seek treatment (D=1 or D=0)?[8]
3. conditional on having sought treatment (D=1) , does the individual report any positive medical expenditure (M=0 or M>0)?
4. conditional on positive expenditure, how much is spent on medical care (E(M)) ?

Conditional on **x,** we assume independence between the decision to report sick, the consultation decision and the positive expenditure decision, and thus write:

$$P(S=0\,|\,\mathbf{x}) = 1-\Phi(\mathbf{x}'\boldsymbol{\gamma}) \tag{2}$$

$$P(D=0\,|\,\mathbf{x}) = 1-\Phi(\mathbf{x}'\boldsymbol{\theta}) \tag{3}$$

$$P(M=0\,|\,\mathbf{x}) = 1-\Phi(\mathbf{x}'\boldsymbol{\eta}) \tag{4}$$

$$\log(M\,|\,\mathbf{x}, S=1, D=1, M>0) \overset{D}{=} \mathrm{N}(\mathbf{x}'\boldsymbol{\beta}, \sigma^2) \tag{5}$$

where Φ represents a standard normal distribution function, Equation (2) indicates the probability of an individual reporting sick, Equation (3) the probability of seeking consultation with a doctor, Equation (4) the probability of positive medical expenditure, and Equation (5) represents conditional medical expenditure, i.e. conditional on reporting sick, seeking treatment and engaging in positive expenditure. The maximum likelihood

[7]This is not to say that unconditional estimates are without value - Dow (1997) observes that the consideration of healthy people in curative care demand is important to the extent that such individuals may adjust inputs into health in the long run in order to affect their probability of illness. To accurately model health behavior, it is thus important to also consider this sample, which may present with somewhat different characteristics to the 'sick' group.

[8] Seeking treatment or 'Consultation' is broadly defined as seeking medical advice from any one or more of a number of sources, namely private doctors, clinics, hospitals, primary healthcare visitors, nurses, traditional healers and pharmacies. We did not include those who reported seeking advice from friends and family (less than 1% of the sick reported such consultations).

estimators (MLE) of $\boldsymbol{\gamma}$, $\boldsymbol{\theta}$, and $\boldsymbol{\eta}$ are simply the probit estimates of the parameter vectors in the sickness, consultation and positive medical expenditure equations. Conditional medical expenditure follows a lognormal distribution (as shown in Figures A1 and A2), thus the MLE estimator of $\boldsymbol{\beta}$ is simply the OLS estimator from a regression of log (M) on $\mathbf{x}$ using only those observations where individuals report positive medical expenditure.

Joint consideration of the coefficients in each of the four stages on the relevant variables allows us to extract the *unconditional* expectation in the Hurdle Model and thus allows for comparison with the simple OLS unconditional model. Any significant difference in these two reflects that gender inequality may be present in one or more if not all of the decision stages. Using properties of the lognormal distribution, we can show that:

$$E(M \mid \mathbf{x}, S=1, D=1, M>0) = \exp(\mathbf{x}'\boldsymbol{\beta} + \sigma^2/2) \qquad (6)$$

$$E(M \mid \mathbf{x}) = \exp(\mathbf{x}'\boldsymbol{\beta} + \sigma^2/2)\Phi(\mathbf{x}'\boldsymbol{\gamma})\Phi(\mathbf{x}'\boldsymbol{\theta})\Phi(\mathbf{x}'\boldsymbol{\eta}) \qquad (7)$$

Given $\hat{\beta}$, $\hat{\gamma}$, $\hat{\eta}$ and $\hat{\theta}$ it is easy to calculate these expectations. It is then possible to calculate the marginal effect of a change in any continuous independent variable x_i on medical expenditure, *M*, in both its conditional and unconditional expressions by differentiating equations (6) and (7) respectively:

$$\frac{\partial E(M \mid \mathbf{x}, S=1, D=1, M>0)}{\partial x_i} = \beta_i \cdot \exp(\mathbf{x}'\boldsymbol{\beta} + \sigma^2/2) \qquad (8)$$

$$\begin{aligned}
\frac{\partial E(M \mid \mathbf{x},)}{\partial x_i} = {} & \beta_i \cdot \exp(\mathbf{x}'\boldsymbol{\beta} + \sigma^2/2)\Phi(\mathbf{x}'\boldsymbol{\gamma})\Phi(\mathbf{x}'\boldsymbol{\theta})\Phi(\mathbf{x}'\boldsymbol{\eta}) \\
& + \gamma_i \phi(\mathbf{x}'\boldsymbol{\gamma})\exp(\mathbf{x}'\boldsymbol{\beta} + \sigma^2/2)\Phi(\mathbf{x}'\boldsymbol{\theta})\Phi(\mathbf{x}'\boldsymbol{\eta}) \\
& + \theta_i \phi(\mathbf{x}'\boldsymbol{\theta})\exp(\mathbf{x}'\boldsymbol{\beta} + \sigma^2/2)\Phi(\mathbf{x}'\boldsymbol{\gamma})\Phi(\mathbf{x}'\boldsymbol{\eta}) \\
& + \eta_i \phi(\mathbf{x}'\boldsymbol{\eta})\exp(\mathbf{x}'\boldsymbol{\beta} + \sigma^2/2)\Phi(\mathbf{x}'\boldsymbol{\gamma})\Phi(\mathbf{x}'\boldsymbol{\theta}) \qquad (9) \\
= {} & \exp(\mathbf{x}'\boldsymbol{\beta} + \sigma^2/2)[\beta_i \Phi(\mathbf{x}'\boldsymbol{\gamma})\Phi(\mathbf{x}'\boldsymbol{\theta})\Phi(\mathbf{x}'\boldsymbol{\eta}) \\
& + \gamma_i \phi(\mathbf{x}'\boldsymbol{\gamma})\Phi(\mathbf{x}'\boldsymbol{\theta})\Phi(\mathbf{x}'\boldsymbol{\eta}) + \theta_i \phi(\mathbf{x}'\boldsymbol{\theta})\Phi(\mathbf{x}'\boldsymbol{\gamma})\Phi(\mathbf{x}'\boldsymbol{\eta}) \\
& + \eta_i \phi(\mathbf{x}'\boldsymbol{\eta})\Phi(\mathbf{x}'\boldsymbol{\gamma})\Phi(\mathbf{x}'\boldsymbol{\theta})]
\end{aligned}$$

As before, Φ represents the standard normal distribution function, while ϕ represents the standard normal density function. It is the marginal effect from Equation (9) that is directly comparable to the coefficient estimate from the simple unconditional medical expenditure equation. However, the variable of primary interest to us in these individual-level regressions is the gender indicator, *male*, and we wish to calculate the marginal effect of *male*. Since taking derivatives of a binary variable is problematic, we instead evaluate the unconditional expectation of medical expenditure for *male*=1 and then for *male*=0, substituting in mean values for the remaining dependent variables, and then take the difference between the two, i.e.

$$E(M \mid x_i = 1, \mathbf{x}_{-i} = \overline{\mathbf{x}}_{-i}) - E(M \mid x_i = 0, \mathbf{x}_{-i} = \overline{\mathbf{x}}_{-i}) \qquad (10)$$

where x_i refers to the *male* dummy, and $\mathbf{x}_{\text{-i}}$ to all remaining explanatory variables. $E(M|\mathbf{x})$ is given by Equation (7).

A potential problem in constructing the hurdle model is the assumption of independence between decisions at each stage of the process. In principle, this could be accounted for by estimating sample selectivity corrected equations for the consultation decision, positive expenditure decision, and the conditional log of medical expenditure decision, but in practice it is difficult to find convincing variables to identify the selectivity term.

Regressions were run both at the individual level, and at the household level. For the latter we assume an extended Engel curve relationship between medical expenditure and household income, as first formulated by Working (1934) and expanded by Deaton (1997) to allow for the inclusion of household demographics and other characteristics.[9] Our household level specification is thus:

$$m_i = \alpha + \beta \ln(x_i / n_i) + \eta \ln n_i + \sum_{k=1}^{K-1} \gamma (n_{ki} / n_i) + \boldsymbol{\tau}.\mathbf{z}_i + u_i \qquad (11)$$

where m_i represents the share of household *i's* budget dedicated to medical expenditure, x_i is total expenditure of household *i*, and n_i is household size (such that x_i / n_i represents per capita expenditure). In Working's analysis, goods are defined as either necessities or luxuries, and have the appropriate sign on the household expenditure coefficient, i.e. $\beta_i < 0$ indicates a necessity, since consumption declines as income rises, while $\beta_i > 0$ indicates a luxury. The variable n_{ki} reflects the number of people in age-sex class *k* where there are *K* such classes in total (such that n_{ki} / n_i reflects the proportion of household members in each class). As outlined previously, we include five age-sex categories, namely males and females in each of the age groups 0 - 5 years, 6 - 15 years, 16 - 40 years, 41 - 64 years, and 65 years and older. Males in the youngest age cohort are taken as the base category. We also allow for the inclusion of a vector of socioeconomic characteristics, $\mathbf{z}_\text{i}$, (such as race, the dependency ratio, gender of the household head), as well as an error term, u_i. For comparability with Case and Deaton (2002), we also run a more parsimonious household specification as outlined in Equation (1).

[9]In conducting this investigation, we assume that a unitary model of the household applies, i.e. that all household income is pooled and allocated among members according to a joint utility function. However, recent theoretical models indicate that it is more realistic to assume a degree of intra-household conflict arising from different individual preferences. There is certainly evidence to suggest that in the South African context, the unitary model of the household is not applicable (see for example Bookwalter & Warner, 2001). Some caution should thus be exercised when interpreting individual medical expenditure as a household level decision.

Much of the existing literature on household-level discrimination employs Deaton's household composition methodology in an attempt to evaluate the extent of gender bias in household expenditures. However, we are fortunate that our dataset also allows us to examine the question of gender differentiation in medical expenditure at the level of the *individual*. Deaton (1997: 240-241) discusses an apparent puzzle of the household literature on gender bias: even where measured outcomes show clear differences between male and female groups, expenditure patterns generally fail to demonstrate significant gender differences. Following the work of Kingdon (2005), examining the gender question at both the individual and household levels allows us to examine whether there is some form of aggregation bias, such that sex differentiation is more detectable with individual level than with household level data.

A final note is in order: so as to retain a degree of consistency with the household level model, in our individual level expenditure equations, our dependent variable is defined as an individual's medical expenditure divided by per capita household expenditure (or the logged version of this, in case of the conditional equation). Therefore, in considering our regression results we must be careful to bear in mind how the dependent variable is specified. We expect medical expenditure to be a necessity, therefore a smaller coefficient on the household income (explanatory) variable should be understood as reflecting two forces driving in somewhat contradictory directions: either a lower concern for health or higher household income. We will consider both.

5. Discussion of results

We first discuss the results of household level regressions, i.e. those using household-level data, in Section 5.1, and then discuss the results using individual level data in Section 5.2. Section 5.3 examines the age structure of gender differences using individual level data. Section 5.4 reports household fixed effects estimation of our various equations to obtain results that circumvent the problem of endogeneity bias in certain household level variables such as household size and total household expenditure. Section 5.5 expands the interpretation of the hurdle model.

5.1 Household level analysis

Regressions of the household budget share of both food (a necessity) and non-food expenditure (not presented) showed no clear evidence of gender bias and, if anything, showed slight *pro-female* bias in food expenditure in two of the age groups. This suggests that women are not generally discriminated against in the South African context.

The household level results for medical care are presented in Table 4. The first column presents OLS estimates of the household medical budget share, i.e. the proportion of total household expenditure spent on healthcare. It is called 'unconditional' because households with zero medical budget shares are also included. The next four columns represent the Hurdle Model. Column 2 is a binary probit equation of anyone in the household reporting being sick in the past two weeks; column 3 is a probit of anyone in the household consulting a medical practitioner, conditional on reporting sick; column 4 is a probit of the household incurring positive medical expenditure, conditional on seeing a medical practitioner; and the last column is the OLS of the natural log of the conditional medical budget share, i.e. conditional on having positive medical budget share. The reason for using logs is because the kernel density of household medical budget share shows that this variable is log-normally rather than normally distributed (Figures 1 and 2).

[Table 4 about here]

The last four rows present the P-values of the F-test of the null that the coefficients on the male and female variables within each age cohort are equal. There is evidence of significant pro-female gender difference in health care only in the 'reporting sick' decision but only in age groups 16-40 (p-value of 0.05) and 65+ years (p-value 0.03). In other words, women in child-bearing age and elderly women are statistically significantly more likely to report sick than their male counterparts. There is no evidence of gender differentiation in the unconditional medical budget share equation or indeed in the other stages of the hurdle model, other than the reporting sick decision, though conditional medical expenditure is higher (at the 9% level of significance) for men than women in the prime working age 41-64 years. The result for elderly women in the 'reporting sick' decision may be attributable to a sample selection issue rather than a gender effect: since women generally outlive men, there are considerably more elderly females (1,008) than males (682) in our sample. They are on average older and thus more prone to sickness than men. If we assume that women and men are equally stoic and equally likely to suffer morbidity, then higher incidence of 'reporting sick' among women in the child-bearing age (16-40) compared with men in that age group could represent the lower opportunity cost of women's time or that in this child-bearing and maternal period, women's health is regarded as particularly important. With household data it is not possible to adjudicate between these two explanations; we revisit them later with individual level data.

In summary, the household level regressions do not show much gender disparity in health behavior apart from higher sickness reporting by women than men in the prime-age and elderly age categories. Conditional expenditure results do provide weak suggestion of pro-male medical expenditures in the older working age-group, suggesting the operation of economic imperatives in this age group.

The advantage of hurdle model estimation becomes clear when we compare the findings of the commonly estimated Engel Curve equation of medical budget share (column 1) with the more disaggregated hurdle model in columns 2-5[10]. While the Engel Curve equation of unconditional medical budget share (column 1) finds no significant gender difference in any age group, the hurdle model reveals significant pro-female gender differences in the reporting sick decision (column 2) in the 16-40 and 65+ age groups and weakly significant pro-male gender difference in the conditional medical expenditure decision (column 5) in the 41-64 age group. In other words, modeling only the unconditional medical expenditure decision masks the fact that in one or more of the underlying constituent decisions (reporting sick, consulting a doctor, incurring positive medical expenditure and conditional medical expenditure) there may be significant gender differentiation, and that it may be in opposing directions in the different decisions.

5.2 Individual level analysis, taking all ages together

With individual level data, it is possible to disaggregate household medical spending with more precision, and one would expect that any gender bias would be revealed more clearly than in the household regressions discussed in section 5.1.

The primary variable of interest in the individual level regressions is the *male* dummy variable. Recall that women's labor force participation rates and wages were both markedly lower than men’s (Table 1). This provides some support for an *a priori* expectation of gender differentiation, based either on a household `investment motive' or on an opportunity cost explanation, rather than a pure gender bias explanation.

Table 5 presents community fixed effects models using individual-level data. Results without community fixed effects were similar. The gender dummy *male* is insignificant in column 1, i.e. there is no gender difference in unconditional medical expenditure (all ages taken together). There are interesting results on the labor market variables: compared to those out of the labor force, the unemployed and those in regular employment both spend less on medical care (less by roughly 1 percentage point of the household budget; this is large in light of mean medical budget share in Table 2). For the unemployed, this likely reflects low resources, which limits potential expenditure on medical care; for those in regular employment, this may reflect lower utilization of medical care due to their higher opportunity cost of lost time and foregone wages.

[Table 5 about here]

Labor market status is clearly endogenous to health status: those more prone to sickness are arguably less likely

[10] To compare our Engel curve equation results with those of Case and Deaton (2002) for South Africa, we estimated the unconditional medical budget share equation using the same parsimonious specification and sample (African and colored families only) as in Case and Deaton. The F-test results (not reported) showed no significant gender difference in any of the age groups, a conclusion common with Case and Deaton.

to obtain stable employment or even be in the labor force. More `motivated' or stoic individuals are both more likely to retain stable employment and also less likely to report sickness at a given health status, due to a greater unwillingness to take time off from productive activities. Such a positive association remains a potential source of endogeneity bias. We attempt to ameliorate such bias via household fixed effects estimation in Section 5.4.

While we are controlling for employment status, men and women in the same form of employment may still have different earnings: within the `regular wage employed' group, average earnings for women are R894 compared with R1,283 for men perhaps because women are in inferior jobs. A man will therefore on average suffer greater earnings-loss upon reporting sick and taking time off work, an effect not fully captured by the relevant employment variables and therefore subsumed in part by the gender dummy.

Introducing the healthcare decision as a sequential process in the Hurdle model allows us to begin to separate out some of these effects, and is more revealing in terms of gender differences. Men are 1.3 percentage points less likely to report sickness than women (column 2), and this effect is statistically highly significant. It could reflect that sick time is a relative luxury for household breadwinners (predominantly male) compared with those engaged in household production (predominantly female), although we cannot discount the pure bias or biological explanations at this level of aggregation, where we have pooled all ages[11]. The fact that those in regular employment are 2 percentage points less likely to report sick than those out of the labor force lends credence to the ‘higher opportunity cost’ explanation. There is a positive linear relationship between age and the probability of reporting sick, with each 10 years of age increasing that probability by about 1 percentage point. Years of education has a large and robust negative association with the likelihood of reporting sick: each extra year of education is associated with a nearly 0.4 percentage point reduction in the probability of reporting ill; thus increasing education by 5 years reduces the chances of reporting sick by nearly 2 points. This could be because the better-educated are more informed about health risks and take more preventative measures, or because they have higher opportunity cost of reporting sick.

There is no significant gender differentiation in the consultation decision (column 3, Table 5): conditional on reporting sick, men and women are equally likely to seek medical care. There is a convex relationship with respect to age, which may reflect the lower opportunity cost of time for the young and elderly. Unemployed people are 10 percentage points less likely and people in regular employment 9 points more likely to seek treatment than the base category (persons out of the labor force). This difference is likely to reflect unemployed and waged workers’ differential ability to pay for medical treatment, conditional on reporting being sick.

There is a large gender difference of 5 percentage points in the positive medical expenditure decision (column

4), though the marginal effect of *male* is only weakly significant. The marginal effect of 5 points is not small considering that 66% of all sick individuals report positive medical expenditure. Pro-female bias in the positive-expenditure decision, taking the sample as a whole, is a somewhat surprising result especially given that we have already allowed for women's higher rate of reporting sick. One possibility could have been the high costs of obstetric and gynaecologic care associated with child-bearing but data do not support this explanation since, among women (of all ages) reporting sick, less than 2% cite a pregnancy-related sickness, and in the child-bearing age (16 - 40) only 5% report a pregnancy-related sickness. A related possibility is that child-bearing may produce greater ill-health among women due to maternal health-depletion if children are closely-spaced (Strauss *et al* 1993: 806). Nonetheless, this effect would be subsumed in the first stage hurdle (reporting sick) unless it was the case that the nature of such illness would render it more expensive to treat. While the health effects of child-bearing most likely contribute in part to the result outlined here, we do not feel it is the only factor driving this result. The quadratic effect of age could reflect the more specialized nature of pediatric care and the expense of treating chronic conditions in the elderly.

Conditional on deciding to spend on healthcare, there is no gender differentiation in the conditional medical budget share equation (column 5). An examination of types of practitioners consulted by sample males and females shows virtually identical consultation patterns by gender. Thus, there are no systematic gender differences in the nature of care that warrants higher charges for any gender. Those regularly employed spend significantly more on treatment than those out of the labor force. The impact of age is linear, possibly explained by increased dependence on expensive prescription medication for chronic conditions, the incidence of which increase with age.

5.3 Individual level analysis: The age structure of gender differences

Having established in section 5.2 that, taking all ages together, there is gender differentiation in the decision to report sick, as well as the decision to engage in positive medical expenditure, we now explore the age pattern of gender differences. We also attempt to understand what is driving gender differences by building up each component of our hurdle model. For each decision, we first include only a *male* dummy, then include age cohorts and interaction terms between these and the *male* dummy (using male age 0 - 5 as the base category), next we add employment and education variables, and finally, we add a number of household level variables.

In light of our hypotheses of Section 2, we anticipate that gender differences in health behavior should manifest primarily in the working age population, i.e. those aged 16 - 40 and 41 – 64. The key variables of interest therefore are those related to gender and work status.

---

[11] In section 5.3 where we probe the age structure of the gender differences, we examine this issue further. The biological

As before, we control for community fixed effects in all regressions. Results are reported in Table 6. We begin by considering the unconditional medical expenditure decision (last main column, labeled column 5). Here few of the gender variables are significant but results for the elderly (over 65) suggest a small pro-female bias. We anticipate that biology and the age structure of the elderly group are the key explanatory factors for this result, as well as higher sickness reporting rates among elderly women. Certain of the age and labor market variables are significant.

[Table 6 about here]

Turning to the hurdle model in Table 6, we consider each of the binary stages in turn, starting with the reporting sick decision (column 1). In the simplest specification, we see the *male* dummy is highly significant, suggesting women are nearly 2 percentage points more likely to report sick, in line with raw data from Table 1. However, once age cohorts and interaction terms are included in the regression, the age pattern of gender differentiation becomes clear. In particular, men in both the working-age groups are significantly less likely to report sick than their female counterparts, which aligns with our hypothesis that much of the gender difference in health behavior relates to greater likelihood of reporting sick among women of child-bearing and maternal ages. The fact that this pro-female gender difference exists in the elderly age group is likely to be because of a selection effect: women in this age group are on average older than the men in this age group (due to their longer life-expectancy than men) and the incidence of morbidity /medical complications is greater among older persons.

Introducing labor market status variables causes the coefficient on 'male aged 41 – 64' to fall and become statistically insignificant. This implies that higher opportunity cost of time is an explanation for men's lower probability to reporting sick (than women) only in the age group 41-64. In the 16-40 age group, the introduction of labour market variables – which capture opportunity cost of time – does not significantly alter the coefficient on the gender dummy, suggesting that women's health is highly valued over the maternal period. It is not immediately apparent why the labor market variables have a strong effect in the 41 - 64 age cohort. One potential explanation may be that those in this group are at the peak of their earnings potential (Table 1 shows men in this age category to be the highest mean earners), and thus the opportunity cost argument may be particularly pertinent.

The second stage consultation hurdle shows that in the younger working group (ages 16 - 40) – also the child-bearing age group – women are 14 percentage points more likely to consult a medical practitioner than men, conditional on reporting sick. The decision to incur positive medical expenditure (third stage hurdle) also shows important gender differentials only in the two working-age groups. Women in the younger working-age group

---

explanation was discussed earlier in Section 2 when discussing our three hypotheses.

are 16 percentage points more likely than men to engage in positive expenditure conditional on consultation, while women aged 41 - 64 are 26 points more likely to do so. Controlling for labor market status (and thus for opportunity cost of time) makes little difference to the results. This clear pro-female bias in both the consultation and positive expenditure decisions for individuals aged 16 – 40 years old – even after controlling for opportunity cost of time – is persuasive evidence that particular value is placed on household production and child-rearing, with implicit recognition of women's non-market contribution to household welfare.

Finally we consider the conditional medical budget share results (column 4), noting that once again gender and labor market effects are insignificant, as in the pooled sample results of Table 5. We conclude that the majority of gender differentiation occurs in the binary decision stages, rather than in the conditional expenditure decision.

5.4 Individual-level analysis: Results with household fixed effects

A final consideration is whether the observed gender differences are largely accounted for by unobserved household heterogeneity. If health behavior is more similar within than across households, then family fixed effects estimation gives results that are less contaminated by omitted variable bias. We therefore repeat the hurdle specification employing a household fixed effects model. Only individual level variables are included, and a linear probability model (LPM) is used for the binary decision stages.[12] Family fixed effects is also a means of circumventing bias on the gender variable in all our equations upto now due to endogeneity of household variables such as household size and household per capita expenditure.

[Table 7 about here]

The results, reported in Table 7, show strong and precisely determined gender effects of the same type as observed before. Pro-female gender differentiation is most evident in the reporting sick decision and, to a lesser extent, in the decision to incur positive medical expenditure. These effects are particularly strong for working age groups and the elderly. Controlling for unobserved household heterogeneity, working age women are 3 – 4 percentage points more likely (and elderly females about 8 points more likely) to report sick than their male counterparts, although the gender differential with respect to positive medical expenditure is much reduced compared with the large marginal effects in the corresponding regressions in Table 6, column 3.

[12] In running a household fixed effects model, sample size is much reduced. However, there remain enough observations to conduct meaningful regressions, with two or more binary observations occurring in 507 households for the reporting sick decision, 377 households for the consultation decision, and 296 households for the positive expenditure decision.

5.5 Individual-level analysis: Expanding the interpretation of the hurdle model

We now consider whether the hurdle methodology has added significantly to our understanding of the dynamics of healthcare behavior. To do so, we have presented the marginal effect (ME) on the *male* dummy variable for a number of regressions based on different sub-samples of individuals. Table 8 reports these results, with column 6 reporting the combined marginal effect on *male* in the four stages of the Hurdle model, calculated using the approach in Equation (10). Standard errors are calculated by bootstrapping and the appropriate z-statistics reported in parentheses.[13] The ME is directly comparable to that in column 2, since in manipulating the hurdle regression results, we transform the conditional log value reported in column 5 such that it now represents an `unlogged' medical budget share.

Some broad initial observations are in order. First, the calculated combined ME appear extremely low (we have multiplied them by 100 to avoid showing lots of zeros after decimal places), however in light of the fact that the mean of unconditional medical expenditure as a proportion of household per capita expenditure is only 1.05% (Table 2, bottom panel), such small figures are unsurprising. Second, while most of the unconditional estimates calculated for the conventional model (column 1) are insignificant, the sign on these estimates differs to that in the calculated coefficients (column 6) in roughly half the cases presented. This provides evidence for the importance of the hurdle model in helping to unravel the contribution of various components of health behavior.

[Table 8 about here]

The key finding from Table 8 is that the male dummy variable is significant in a number of sub-samples for which a simple unconditional regression predicts a non-significant gender effect. While there is generally little significance for the consultation and conditional expenditure equations (columns 3 and 5), it is consistently the higher probability with which females tend to report sick and incur positive medical expenditure that has an impact on the unconditional estimate in column 6, and which is not apparent in the conventional model[14].

---

[13]In order to avoid a cumbersome computational process given the bootstrapping methodology, we chose to employ the specification without community fixed effects. This is justified in light of similarity of the coefficients in regressions with and without community fixed effects.

14 The only hint of pro-male differentiation is in the medical expenditure decision for the very young, ages 0 - 5, where, conditional on reporting sick, boys are 6.1 percentage points more likely than girls to be taken to consult a healthcare practitioner. In this age category one cannot invoke a labor market explanation for gender differentiation, thus to the extent that pure gender bias exists, it is most likely to show up within this group. However, in the absence of pure gender bias in any other age group, including for the intermediate juvenile group, this appears an unlikely explanation. Instead, there is support for the notion of gender differentiation in the nature of illnesses suffered by boys and girls. A study by Van den Bosch et al (1992) finds that young boys aged 0 and 4 are significantly more likely to suffer respiratory diseases, behavioral disorders, gastroenteritis and accidents, resulting in higher rates of referrals to specialists and hospital admissions than girls of the same age. An explanation along these lines seems more plausible than one based on pure gender bias.

In order to further assess the nature of gender differentiation, apart from the pooled sample in the first row of Table 8, a number of regressions were run on sub-samples: different age cohorts; individuals in households that are above and below the poverty line[15]; individuals in households with none or some unemployed members; those resident in a former homeland and those not; and by race group. These equations demonstrate the value of the hurdle specification.

We begin by analyzing the combined estimate (final column) for the pooled sample in the first row. This suggests that women's unconditional medical expenditure budget share is roughly 0.16 percentage points higher on average than men's, after accounting for the various channels through which health behavior operates. Conversely, the unconditional OLS coefficient in the first column, while insignificant, suggests that the unconditional medical budget share is 0.08 percentage points higher for men, in line with what might be anticipated from the raw summary statistics (Table 2). It appears that the components driving the result are not significant when aggregated, a result only revealed when the sequential hurdle model is employed.

For the remaining split sample regressions, it is of particular interest to note the cases in which the unconditional coefficient recovered from the individual hurdles is statistically significant, and consider how this contributes to our understanding of the gender differentiation uncovered thus far. First looking at age cohorts, we note that the results clearly align with our previous findings: elderly females are likely to spend a (large) 1.7 percentage points more of the household per capita budget on medical care, while females in the older working age group are likely to spend a somewhat smaller but still relatively substantial 0.38 percentage points more than their male counterparts. It is somewhat surprising that the combined coefficient on the *male* dummy for the 16 - 40 age group (though relatively large in size) is not statistically significant, in spite of large and significant coefficients at the reporting sick and positive expenditure decision stages. We suspect small and insignificant coefficients on the *male* dummy at the remaining two stages to be at the root of this.

The *male* variable is significantly negative both among black individuals and among former homeland residents (in these groups, women spend respectively 0.18 and 0.20 percentage points more of the per capita household budget on medical care). Black and former homeland households' more gendered division of labor means that adult women are likely to be more responsible for childcare and household maintenance than in their counterparts (non-black and non-homeland households), In light of our maintained hypothesis that gender differentials are explained largely via an implicit recognition of the non-market value of women's health, we attribute this higher pro-female gender effect in these types of households to the more traditional household structure prevalent in the black and homeland categories.

[15] Following May (1998), we use a value of Rands 353 income per adult equivalent as the poverty line in the SAIHS data.

Our hypothesis is also reinforced by the significant negative gender effect in households above the poverty line and in households where all labor force participants are employed, where we observe a pro-female bias of 0.13 and 0.21 percentage points respectively. Here we are examining households for whom we presume income or wealth is less likely to act as a constraint on health behavior, and we would thus anticipate that underlying gender effects would be more readily evident.

6. **Conclusion**

Our central question has been to examine the prevalence of gender differentiation in the allocation of household health expenditure. While patterns of pro-male bias in household expenditures are common in developing country contexts, there is little evidence of gender bias in consumption expenditures in South Africa, so it would not be unsurprising to find a correspondingly limited role for gender in explaining *medical* expenditure. A priori, there are two opposing directions of gender difference in medical expenditure: (1) the investment motive hypothesis predicts higher spending on male health; (2) the higher opportunity cost or reporting sick for men predicts lower unconditional spending on male health. We explored both the relative explanatory power of these conjectures, as well as the possibility that women's non-market contribution to household production is viewed as an important investment in household welfare.

Modeling a sequence of health behavior decisions, we observe clear pro-female gender differentiation particularly in the binary decision stages: reporting sick and incurring positive medical expenditure, conditional on consulting a medical practitioner. An unconditional model of healthcare expenditure tends to dilute the differentiation evident in these phases, and we therefore favor the use of a hurdle model to explain health behavior. In line with a broad international literature, women report sick with higher frequency than men across almost all age groups. However, our most compelling finding relates to evidence suggesting that women are favored in the treatment decision, in particular, the positive medical expenditure decision. Individual level results also indicate some pro-female bias in the consultation decision. The particularly strong results in these decisions for prime age women (aged 16 - 40) provide support for our hypothesis that household health allocation favors women over men in the child rearing period, and aligns well with Gao and Yao's (2006) findings for China.

Additionally, an examination of labor market variables suggests that both labor market mechanisms outlined above are valid in different contexts. The sequential model is of particular value in unraveling these effects. As anticipated, while the regular wage employed are less likely to report sick than others, conditional on doing so and on consulting a healthcare practitioner, their expenditure is higher relative to that of other labor market groups. Inclusion of employment status variables allows us to see the purer effect of gender, without the confounding influence of opportunity cost of time. The fact that women's higher sickness reporting and higher likelihood of incurring positive medical expenditure (conditional on reporting ill and seeking medical consultation) than men persists even after controlling for the economic / labor market variables suggests that it is not women's lower opportunity cost of time that is responsible for their being more likely to seek/get medical attention, and it suggests instead that families recognize the value of women in household production.

This paper has also considered the relative explanatory power of individual and household level regressions, and the related implications for the analysis of gender bias. We find that the extent and dimensions of gender differentiation are revealed far more clearly in individual-level regressions, despite the use of similar specifications and identical data in both models. This addresses Deaton's (1997) 'puzzle' and affirms the conclusions of Kingdon (2005). Since results for household level regressions are more muted than those for individual level regressions, it appears there is something in the aggregation procedure that makes it more difficult to observe gender differences in expenditure. We advance this as a potential explanation for the failure to detect gender differentiation in medical expenditures in other contexts, e.g. in India where there is a strong *a priori* expectation of gender bias and yet no gender differences are detected (see Subramanian and Deaton 1990). The use of individual level data is therefore advocated in preference to household level data for the evaluation of gender bias in household expenditure.

References

Africa, M., Budlender, D. and Mpetsheni, Y. 2001. *Education in South Africa: Selected findings from Census 1996.* Statistics South Africa, Pretoria.

Bookwalter, J.T and Warner, J.M. 2001. *Power and Bargaining in Intra-Household Resource Allocation: Race, Class and Gender in South Africa*. University of Montana Department of Economics Working Paper. http://www.umt.edu/econ/papers.htm

Case, A. and Deaton, A. 2002. *Consumption, health, gender and poverty.* Research Programme in Development Studies, Princeton University, New Jersey.

Case, A. and Paxson, C. 2004. *Sex differences in morbidity and mortality.* NBER Working Paper No. W10653, Cambridge (Massachusetts).

Clearly, P.D., Mechanic, D. & Greenley, J.R. 1982. `Sex differences in medical care utilization: an empirical investigation.' *Journal of Health and Social Behavior,* Volume 23(2), pp.106-119

Coale, A. J. and Banister, J. 1994. `Five decades of missing females in China.' *Demography,* Volume 31(3), pp. 459-479.

Deaton, A. 1997. *The Analysis of Household Surveys: a microeconometric approach to development policy.* Johns Hopkins University Press for the World Bank, Washington, DC.

Dow, W. H. 1996. *Unconditional demand for health care in Côte d'Ivoire: does selection on health status matter?* Living Standards Measure Study Working Paper No. 127, The World Bank.

Gao, M. & Yao, Y. 2006. `Gender Gaps in Access to Health Care in Rural China.' *Economic Development and Cultural Change,* Volume 55, pp.87-107.

Gangadharan, L. & Maitra, P. 2003. `Testing for Son Preference in South Africa.' *Journal of African Economies,* Volume 12(3), pp.371-416.

Haddad, L., J. Hoddinott and H. Alderman (1997) *Intrahousehold Resource Allocation in Developing Countries: Models, Methods and Policy*, Johns Hopkins University Press.

Kingdon, G. 2005. `Where has all the bias gone? Detecting gender bias in the household allocation of educational expenditure.' *Economic Development and Cultural Change*, Volume 53 pp.409-451.

Kingdon, G. and J. Knight. "The measurement of unemployment when unemployment is high", *Labour Economics*, 13, No. 3: 291-315, June 2006.

May, J. 1998. *Poverty and Inequality in South Africa - Summary Report.* Report prepared for the Office of the Executive Deputy President and the Inter-Ministerial Committee for Poverty and Equality, 13 May 1998. www.welfare.gov.za/Documents/1998/Pov.htm

Pitt, M., Rosenzweig, M. & Hassan, N. 1990. `Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries.' *American Economic Review*, 80(5), pp.1139--56

Puoane, T, Steyn, K, Bradshaw, D, Laubscher, R., Fourie, J., Lambert, V. & Mbananga, N. 2002. `Obesity in South Africa: The South African Demographic and Health Survey.' *Obesity Research,* 10, pp.1038-1048

Ray, R. 2000. `Poverty and expenditure pattern of households in Pakistan and South Africa: a comparative study.' *Journal of International Development*, 12(2), pp.241 - 256

Rose, E. 1999. `Consumption Smoothing and Excess Female Mortality in Rural India.' *Review of Economics and Statistics,* Volume 81(1), pp.41--49

Rosenzweig, M. R. & Schultz, T. P. 1982. `Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India.' *American Economic Review*, 72 (September), pp.803-15

Sen, A. 1992. `Missing women.' *British Medical Journal,* Volume 304 (March), pp.5-6

Subramanian, S. & Deaton, A. 1990. *Gender Effects in Indian Consumption Patterns*. Discussion Paper No. 147, Woodrow Wilson School, Princeton University.

South Africa Labor Development Research Unit. 1994. *South Africans Rich and Poor: Baseline Household Statistics.* University of Cape Town, South Africa

Strauss, J., Gertler, P. J., Rahman, O. & Fox, K.. 1993. `Gender and Life-Cycle Differentials in the Patterns and Determinants of Adult Health.' *The Journal of Human Resources*, 28(4), pp.791-837

Van den Bosch, W.J.M., Huygen F.J., van den Hoogen H.J. & van Weel C. 1992. `Morbidity in early childhood: differences between girls and boys under 10 years old.' *British Journal of General Practice*, 42, pp.366--369

Waldron, I. 1983. `Sex Differences in Illness Incidence, Prognosis and Mortality: Issues and evidence.' *Social Science and Medicine,* Volume 17(16), pp.1107-1123

Wittenberg, M. 2005. *The School Day in South Africa.* Centre for Social Science Research Working Paper No. 113, University of Cape Town

Wooldridge, J. 2002. *Econometric Analysis of Cross Section and Panel Data.* MIT Press, Boston.

Working H. 1934. `Statistical laws of family expenditure.' *Journal of the American Statistical Association*, 34, pp.43-56.

Figure 1: Epanechnikov kernel density function
Unconditional medical expenditure as a proportion of per capita household expenditure

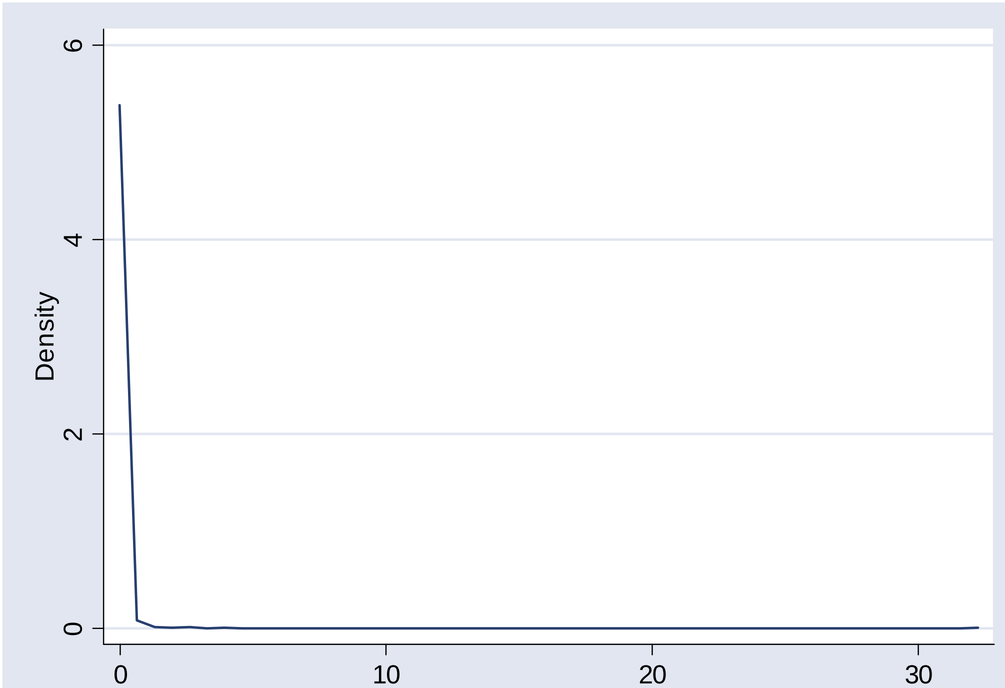

Figure 2: Epanechnikov kernel density function
Log of conditional medical expenditure as a proportion of per capita household expenditure

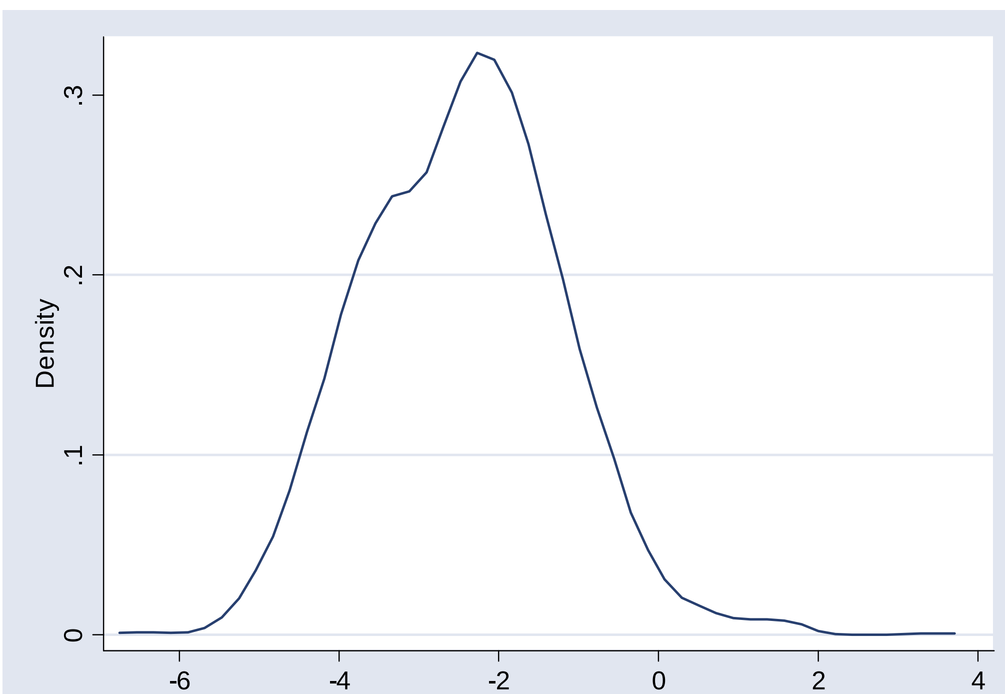

Table 1: Summary of labor market variables by gender

	MALE	FEMALE	DIFFERENCE
Age 16-40			
labor force participation rate (%)	58.69	47.16	11.53
average earnings	1,013.59	758.53	255.06
	(1,319.94)	*(798.22)*	
Age 41-64			
labor force participation rate (%)	65.43	58.69	6.74
average earnings	972.73	664.53	308.20
	(1,316.99)	*(889.47)*	
Combined (age 16-64)			
labor force participation rate (%)	60.43	46.06	14.37
average earnings	1,109.53	785.70	323.83
	(1,391.95)	*(851.13)*	

Notes: Standard deviations for earnings reported in parentheses. Earnings averaged over working individuals only

Table 2: Gender differences in various health decisions, pooled and by age group

	Age group				
	0 - 5	**6 - 15**	**16 - 40**	**41 - 64**	**65 +**
Sample size					
men	3,104	5,353	8,213	2,855	947
women	2,999	5,319	9,035	3,491	1,236
% Reporting sick					
men	8.09	4.11	5.48	12.08	12.35
women	8.50	3.89	7.35	14.90	20.06
difference	-0.41	0.22	-1.87***	-2.82***	-7.71***
% Consulting medical practitioner, conditional on reporting sick					
men	87.25	77.27	78.44	84.93	85.47
women	82.35	77.29	79.82	84.04	83.06
difference	4.90	-0.02	-1.38	0.89	2.41
% Incurring positive medical expenditure, conditional on consultation					
men	88.58	79.41	71.95	69.28	85.00
women	84.29	74.38	84.15	86.73	86.41
difference	4.29	5.03	-12.20***	-17.45***	-1.41
Unconditional medical expenditure					
men	2.22	1.94	2.46	10.53	6.39
women	2.09	0.91	2.97	7.57	18.88
difference	0.13	1.03*	-0.51	2.96	-12.49
Conditional medical expenditure					
men	35.58	77.11	79.59	148.11	71.19
women	35.38	40.75	60.09	69.74	131.10
difference	0.20	36.36	19.50	78.37*	-59.91
Unconditional medical expenditure as a proportion of household per capita expenditure (%)					
men	0.80	0.39	0.74	2.91	1.74
women	0.70	0.32	0.91	2.31	3.42
difference	0.10	0.07	-0.17	0.60	-1.68**
Conditional medical expenditure as a proportion of household per capita expenditure (%)					
men	12.94	15.68	23.73	41.63	19.90
women	11.81	14.27	18.40	21.29	24.17
difference	1.13	1.41	5.33	20.33+	-4.27

POOLED (ALL AGES)	**MALE**	**FEMALE**	**Difference**	*t-statistic*	
sample size	20,472	22,080	-1,608		
% reporting sick	6.76	8.58	-1.82	-7.05	***
% reporting consultation	82.07	81.47	0.60	0.44	
% reporting positive expenditure	76.74	84.19	-7.45	-4.78	***
Unconditional medical expenditure	3.60	3.97	-0.37	-0.47	
Conditional medical expenditure	84.55	67.50	17.05	1.07	
Unconditional medical expenditure as a proportion of household per capita expenditure (%)	1.00	1.10	-0.10	0.51	
Conditional medical expenditure as a proportion of household per capita expenditure (%)	23.80	18.73	5.07	-1.41	

Note: +, *, ** and *** signify statistically significant gender differences at the 15%, 10%, 5% and 1% levels respectively.

Table 3: Descriptive statistics for independent variables (N=42,552)

INDIVIDUAL LEVEL VARIABLES	Mean	SD
male (%)	48.11	
black (%)	80.68	
coloured (%)	7.76	
Indian (%)	2.63	
white (%)	8.94	
age	24.28	*(18.79)*
unemployed (%)*	10.39	
engaged in casual labor or self-employed (%)	14.51	
regular employment (%)	4.37	
out of labor force (%)	70.72	
years of education	5.41	*(4.31)*
HOUSEHOLD LEVEL VARIABLES		
household size	4.50	*(2.97)*
log of household size	1.28	*(0.72)*
household expenditure per capita	607.43	*(989.65)*
log of household expenditure per capita	5.73	*(1.13)*
dependency ratio	0.41	*(0.32)*
male household head (%)	58.98	
COMMUNITY ENVIRONMENT VARIABLES		
community income per capita (R - monthly)	618.63	*(843.63)*
community unemployment rate (%)	38.21	*(24.22)*
proportion of households subject to personal crime (%)	7.04	*(8.77)*
proportion of households subject to property crime (%)	4.11	*(5.92)*
COMMUNITY HEALTH VARIABLES		
proportion of households with flushing toilet installed	44.38	*(47.27)*
proportion of households with electricity installed	48.02	*(43.54)*
distance to closest hospital (km)**	20.03	*(22.04)*

NOTE: mean reported except for binary variables where sample proportions are given
Standard deviation reported in parentheses for continuous variables

**Implies an unemployment rate of 35.5% among adults aged 16-64*
*** Over 7,000 observations were missing for this variable. In order to address this problem, missing observations were assigned a value of 0, and a new dummy variable,* DIST_MISS, *was created to account for the effect of this adjustment*

Table 4: Household level regression results - Controlling for Community fixed effects

	CONVENTIONAL MODEL		HURDLE MODEL		
	unconditional OLS	probit	probit	probit	conditional OLS
	MED_PROP	SICK	CONSULT	POS_MEXP	LNMPROP
HOUSEHOLD LEVEL VARIABLES					
black	0.038	-0.007	0.200	0.208	0.482
	(1.15)	*(0.10)*	*(1.25)*	*(1.09)*	*(1.29)*
coloured	0.036	-0.024	0.182	0.314	0.786
	(1.27)	*(0.22)*	*(1.63)*	*(1.73)**	*(0.76)*
indian	0.021	0.004	-0.886	0.343	-0.705
	(0.98)	*(0.03)*	*(9.96)****	*(1.77)**	*(7.53)****
log of household size	-0.034	0.176	0.185	0.034	-0.540
	(0.77)	*(10.36)****	*(4.58)****	*(0.58)*	*(3.77)****
log of household pce	0.030	0.030	0.159	0.024	-0.167
	(1.33)	*(2.41)***	*(5.22)****	*(0.58)*	*(1.62)*
dependency ratio	-0.043	0.107	-0.287	0.198	0.307
	(0.86)	*(1.48)*	*(1.56)*	*(0.53)*	*(0.54)*
male household head	0.065	-0.013	-0.011	0.074	-0.007
	(1.27)	*(0.84)*	*(0.32)*	*(1.36)*	*(0.05)*
household unemployment rate	0.142	0.015	-0.044	0.040	0.483
	(1.32)	*(0.82)*	*(1.05)*	*(0.61)*	*(3.15)****
HOUSEHOLD DEMOGRAPHIC VARIABLES					
proportion of females 0-5	0.004	0.069	-0.052	-0.327	0.033
	(0.18)	*(0.90)*	*(0.28)*	*(1.23)*	*(0.06)*
proportion of males 6-15	0.013	-0.122	-0.250	-0.103	0.097
	(0.61)	*(1.69)**	*(1.43)*	*(0.46)*	*(0.18)*
proportion of females 6-15	0.227	-0.102	-0.219	-0.103	0.226
	(0.97)	*(1.33)*	*(1.40)*	*(0.42)*	*(0.39)*
proportion of males 16-40	-0.065	0.022	-0.451	0.009	0.383
	(1.05)	*(0.23)*	*(1.95)**	*(0.02)*	*(0.50)*
proportion of females 16-40	-0.003	0.099	-0.464	0.132	0.791
	(0.06)	*(0.97)*	*(1.83)**	*(0.30)*	*(0.93)*
proportion of males 41-64	-0.156	0.159	-0.293	-0.140	1.251
	(1.28)	*(1.59)*	*(1.24)*	*(0.33)*	*(1.39)*
proportion of females 41-64	-0.022	0.215	-0.485	0.261	-0.054
	(0.44)	*(2.18)***	*(1.94)**	*(0.60)*	*(0.06)*
proportion of males 65+	-0.221	0.026	-0.089	0.665	-0.699
	(1.54)	*(0.20)*	*(0.25)*	*(1.44)*	*(0.55)*
proportion of females 65+	-0.042	0.375	-0.127	0.014	-0.361
	(1.13)	*(3.40)****	*(0.53)*	*(0.05)*	*(0.41)*
Observations	6,573	6,663	1,253	963	1,321
R-squared	0.04	0.13	0.13	0.20	0.36
P-values: age 6-15	0.34	0.74	0.94	0.75	0.80
age 16-40	0.23	0.05	0.74	0.42	0.38
age 41-64	0.14	0.54	0.14	0.44	0.09
age 65+	0.20	0.03	0.61	0.27	0.87

Notes: Absolute value of robust t or z statistics in parentheses. *significant at 10%, **significant at 5%, *** significant at 1% level. Marginal effects and pseudo R-squared reported for probit equations. Base category for race is '*white*' and for demographic variables is '*proportion of males aged 0 – 5*'.

Table 5: Individual level regression results - Controlling for Community fixed effects

	CONVENTIONAL MODEL	HURDLE MODEL			
	unconditional OLS	probit	probit	probit	conditional OLS
	MED_PROP	SICK	CONSULT	POS_MEXP	LNMPROP
INDIVIDUAL LEVEL VARIABLES					
male	0.001	-0.013	-0.011	-0.050	0.022
	(0.24)	*(5.62)****	*(0.55)*	*(1.80)**	*(0.41)*
black	-0.005	-0.023	0.061	0.137	-0.513
	(1.30)	*(1.15)*	*(0.65)*	*(0.71)*	*(1.62)*
coloured	-0.001	-0.026	0.118	0.261	-0.112
	(0.10)	*(0.95)*	*(1.29)*	*(1.49)*	*(0.15)*
indian	-0.013	0.016	0.025	0.287	-1.027
	*(2.92)****	*(0.40)*	*(0.39)*	*(2.34)***	*(29.17)****
age	0.001	0.001	-0.004	-0.007	0.009
	*(2.18)***	*(3.48)****	*(2.01)***	*(2.37)***	*(1.60)*
age squared	-1.52e-06	5.65e-06	6.48e-05	9.75e-05	-3.98e-05
	(0.46)	*(1.50)*	*(2.41)***	*(2.68)****	*(0.56)*
years of education	-0.0004	-0.0037	0.0007	-0.0015	0.0044
	*(2.17)***	*(9.91)****	*(0.22)*	*(0.35)*	*(0.52)*
unemployed	-0.010	-0.020	-0.101	0.114	-0.079
	*(2.80)****	*(4.76)****	*(2.29)***	*(1.94)*	*(0.70)*
regular wage employment	-0.009	-0.020	0.089	-0.011	0.161
	*(2.19)***	*(5.04)****	*(2.45)***	*(0.22)*	*(1.65)**
casual employment	-0.006	-0.001	-0.070	0.140	0.076
	(1.08)	*(0.18)*	*(1.45)*	*(2.38)***	*(0.61)*
HOUSEHOLD LEVEL VARIABLES					
log of household size	0.001	-0.015	0.126	0.026	0.152
	(0.45)	*(4.88)****	*(5.46)****	*(0.65)*	*(2.02)***
log of household pce	-0.003	0.006	0.139	0.048	-0.599
	*(2.05)***	*(2.28)***	*(6.62)****	*(1.31)*	*(9.27)****
dependency ratio	-0.004	-0.020	-0.071	0.038	0.051
	(1.26)	*(3.63)****	*(1.79)**	*(0.60)*	*(0.44)*
male household head	0.005	-0.002	0.012	-0.010	0.150
	*(1.65)**	*(0.59)*	*(0.61)*	*(0.28)*	*(2.04)***
Observations †	41,530	41,000	2,463	1,665	2,120
R-squared	0.01	0.09	0.14	0.22	0.40

Notes: Absolute value of robust t or z statistics in parentheses. * Significant at 10%,**significant at 5%, *** significant at 1%. Marginal effects and pseudo R-squared reported for probit equations. Base category for race and labor market variables is *white* and *out of the labor force* respectively.
† Apparent discrepancy in number of observations used is due to observations dropped due to collinearity between the community identifier and binary decision variables

Table 6: Individual level regressions - Controlling for community fixed effects (continues on next page)

	SICK DECISION (1) probit				CONSULTATION DECISION (2) probit				POSITIVE EXPENDITURE DECISION (3) probit			
male	-0.019	-0.004	-0.004	-0.005	0.003	0.060	0.058	0.063	-0.047	0.079	0.079	0.079
	*(8.1)****	*(0.7)*	*(0.7)*	*(0.9)*	*(0.2)*	*(1.2)*	*(1.1)*	*(1.2)*	*(1.8)**	*(1.2)*	*(1.2)*	*(1.2)*
age 6-15		-0.044	-0.034	-0.032		-0.076	-0.095	-0.094		-0.109	-0.112	-0.112
		*(8.0)****	*(5.7)****	*(5.5)****		*(1.5)*	*(1.9)**	*(1.9)**		*(1.4)*	*(1.3)*	*(1.3)*
age 16-40		0.012	0.022	0.019		-0.069	-0.100	-0.079		0.013	-0.004	-0.007
		*(2.5)***	*(3.3)****	*(3.0)****		*(1.7)**	*(2.0)***	*(1.6)*		*(0.3)*	*(0.1)*	*(0.1)*
age 41-64		0.049	0.082	0.072		-0.012	-0.038	-0.024		0.032	0.021	0.031
		*(7.1)****	*(9.9)****	*(9.0)****		*(0.3)*	*(0.8)*	*(0.5)*		*(0.6)*	*(0.3)*	*(0.5)*
age 65+		0.103	0.120	0.111		-0.019	-0.043	-0.009		-0.004	-0.003	0.006
		(9.4)***	(10.3)***	(9.5)***		*(0.4)*	*(0.8)*	*(0.2)*		*(0.1)*	*(0.0)*	*(0.1)*
male age 6-15		0.008	0.007	0.007		-0.081	-0.079	-0.078		-0.018	-0.019	-0.044
		(1.0)	*(0.8)*	*(0.8)*		*(1.1)*	*(1.1)*	*(1.1)*		*(0.2)*	*(0.2)*	*(0.5)*
male age 16-40		-0.017	-0.015	-0.015		-0.109	-0.138	-0.143		-0.183	-0.168	-0.159
		*(2.4)***	*(2.1)***	*(2.2)***		*(1.6)*	*(2.0)***	*(2.0)***		*(2.1)***	*(1.9)**	*(1.8)**
male age 41-64		-0.015	-0.009	-0.008		-0.045	-0.067	-0.083		-0.245	-0.246	-0.260
		*(2.2)***	*(1.2)*	*(1.2)*		*(0.7)*	*(1.0)*	*(1.2)*		*(2.6)****	*(2.6)****	*(2.7)****
male age 65+		-0.027	-0.025	-0.024		-0.028	-0.023	-0.037		-0.068	-0.072	-0.108
		*(3.2)****	*(3.0)****	*(2.9)****		*(0.4)*	*(0.3)*	*(0.4)*		*(0.5)*	*(0.6)*	*(0.9)*
years of education			-0.003	-0.003			0.006	0.003			0.0001	0.001
			*(7.2)****	*(7.7)****			*(2.0)***	*(0.8)*			*(0.0)*	*(0.3)*
unemployed			-0.024	-0.023			-0.115	-0.109			0.084	0.082
			*(6.3)****	*(5.9)****			*(2.55)***	*(2.43)***			*(1.4)*	*(1.3)*
regular employment			-0.021	-0.023			0.096	0.085			-0.04	-0.05
			*(5.0)****	*(5.9)****			*(2.7)****	*(2.3)***			*(0.8)*	*(1.0)*
casual employment			-0.004	-0.005			-0.060	-0.076			0.127	0.115
			(0.7)	*(0.8)*			*(1.3)*	*(1.6)*			*(2.1)***	*(1.9)**
black				-0.023				0.054				0.103
				(1.1)				*(0.6)*				*(0.6)*
coloured				-0.026				0.122				0.202
				(0.9)				*(1.4)*				*(1.2)*
indian				0.011				0.030				0.281
				(0.3)				*(0.5)*				*(2.4)***
log of HH size				-0.017				0.116				0.024
				*(5.5)****				*(4.9)****				*(0.6)*
log HH expenditure pc				0.006				0.134				0.053
				*(2.13)***				*(6.40)****				*(1.48)*
dependency ratio				-0.014				-0.051				0.048
				*(2.56)***				*(1.28)*				*(0.75)*
male HH head				0.001				0.012				0.008
				(0.30)				*(0.55)*				*(0.23)*
Observations	42,021	42,021	42,021	41,489	2,551	2,551	2,551	2,494	1,720	1,720	1,720	1,689
Pseudo R-squared	0.05	0.08	0.09	0.09	0.10	0.11	0.12	0.14	0.20	0.20	0.21	0.22

Table 6 - continued	CONDITIONAL EXPENDITURE (4)				UNCONDITIONAL EXPENDITURE (5)			
	OLS				OLS			
male	0.016	0.059	0.058	0.027	-0.001	0.001	0.001	0.001
	(0.28)	*(0.51)*	*(0.50)*	*(0.24)*	*(0.33)*	*(0.43)*	*(0.46)*	*(0.43)*
age 6-15		0.060	0.108	-0.012		-0.004	-0.002	-0.003
		(0.40)	*(0.68)*	*(0.08)*		*(3.63)****	*(1.81)**	*(1.91)**
age 16-40		0.163	0.268	0.202		0.003	0.009	0.009
		(1.60)	*(2.05)***	*(1.62)*		*(2.05)***	*(3.41)****	*(3.36)****
age 41-64		0.229	0.29	0.266		0.017	0.022	0.022
		*(2.03)***	*(2.19)***	*(2.18)***		*(5.62)****	*(5.72)****	*(5.58)****
age 65+		0.432	0.471	0.512		0.028	0.029	0.030
		*(3.26)****	*(3.38)****	*(3.93)****		*(4.53)****	*(4.76)****	*(5.25)****
male age 6-15		-0.17	-0.179	-0.003		0.001	0.001	0.001
		(0.89)	*(0.94)*	*(0.01)*		*(0.49)*	*(0.37)*	*(0.42)*
male age 16-40		0.017	0.012	-0.027		-0.002	-0.001	-0.001
		(0.10)	*(0.07)*	*(0.17)*		*(0.81)*	*(0.40)*	*(0.55)*
male age 41-64		0.156	0.160	0.093		0.006	0.008	0.006
		(1.00)	*(1.02)*	*(0.60)*		*(0.46)*	*(0.59)*	*(0.52)*
male age 65+		-0.192	-0.192	-0.254		-0.017	-0.016	-0.019
		(0.92)	*(0.92)*	*(1.20)*		*(2.39)***	*(2.35)***	*(2.70)****
years of education (x100)								
			(1.17)	*(0.51)*			*(1.31)*	*(0.94)*
unemployed			-0.113	-0.082			-0.009	-0.010
			(0.93)	*(0.72)*			*(2.96)****	*(3.04)****
regular employment			0.019	0.150			-0.008	-0.008
			(0.18)	*(1.52)*			*(2.05)***	*(2.00)***
casual employment			-0.012	0.080			-0.005	-0.005
			(0.09)	*(0.65)*			*(0.97)*	*(1.02)*
black				-0.521				-0.005
				(1.61)				*(1.13)*
coloured				-0.142				-0.001
				(0.19)				*(0.11)*
indian				-0.998				-0.013
				*(24.34)****				*(3.26)****
log of HH size				0.160				0.0002
				*(2.12)***				*(0.09)*
log of HH expenditure pc				-0.588				-0.003
				*(9.20)****				*(2.28)***
dependency ratio				0.058				-0.003
				(0.48)				*(1.08)*
male HH head				0.151				0.005
				*(1.98)***				*(2.02)***
Observations	2,136	2,136	2,136	2,136	42,035	42,035	42,035	42,020
R-squared	0.31	0.32	0.32	0.40	0.01	0.01	0.01	0.01

Note: *significant at 10%,**significant at 5%, *** significant at 1%

Table 7: Individual level regression results - Controlling for Household fixed effects

	CONVENTIONAL MODEL		HURDLE MODEL							
	unconditional OLS		LPM		LPM		LPM		conditional OLS	
	MED_PROP		SICK		CONSULT		POS_MEXP		LNMPROP	
male	-0.001	-0.000	0.003	0.003	-0.086	-0.088	0.027	0.027	0.121	0.115
	(0.10)	*(0.09)*	*(0.43)*	*(0.48)*	*(1.57)*	*(1.61)*	*(0.67)*	*(0.66)*	*(0.78)*	*(0.74)*
age 6-15	-0.006	-0.004	-0.045	-0.030	-0.149	-0.150	0.005	-0.009	-0.092	-0.128
	(1.25)	*(0.72)*	*(7.21)****	*(4.62)****	*(2.59)****	*(2.54)***	*(0.12)*	*(0.20)*	*(0.53)*	*(0.70)*
age 16-40	0.002	0.010	-0.014	0.024	-0.152	-0.169	0.010	-0.031	-0.002	-0.046
	(0.50)	*(1.89)*	*(2.41)***	*(3.48)****	*(3.22)****	*(2.91)****	*(0.27)*	*(0.68)*	*(0.01)*	*(0.25)*
age 41-64	0.019	0.024	0.064	0.089	-0.117	-0.140	0.029	-0.002	0.210	0.168
	*(3.50)****	*(4.25)****	*(9.40)****	*(12.21)****	*(2.24)***	*(2.44)***	*(0.72)*	*(0.05)*	*(1.31)*	*(0.95)*
age 65+	0.024	0.025	0.124	0.132	-0.137	-0.147	0.112	0.090	-0.053	-0.099
	*(3.15)****	*(3.30)****	*(12.87)****	*(13.65)****	*(2.00)***	*(2.09)***	*(2.16)***	*(1.68)**	*(0.26)*	*(0.47)*
male age 6-15	0.004	0.003	0.001	-0.001	0.014	0.015	-0.001	0.003	0.108	0.122
	(0.56)	*(0.51)*	*(0.08)*	*(0.12)*	*(0.18)*	*(0.20)*	*(0.02)*	*(0.05)*	*(0.49)*	*(0.55)*
male age 16-40	-0.001	0.000	-0.025	-0.022	0.040	0.042	-0.018	-0.020	0.054	0.056
	(0.09)	*(0.03)*	*(3.12)****	*(2.73)****	*(0.59)*	*(0.61)*	*(0.36)*	*(0.38)*	*(0.27)*	*(0.28)*
male age 41-64	0.007	0.009	-0.038	-0.028	0.037	0.031	-0.098	-0.108	-0.032	-0.033
	(0.93)	*(1.20)*	*(3.91)****	*(2.90)****	*(0.52)*	*(0.44)*	*(1.86)**	*(2.03)***	*(0.15)*	*(0.15)*
male age 65+	-0.015	-0.014	-0.080	-0.077	0.129	0.130	-0.128	-0.129	0.059	0.064
	(1.35)	*(1.30)*	*(5.82)****	*(5.62)****	*(1.45)*	*(1.45)*	*(1.94)**	*(1.95)**	*(0.23)*	*(0.24)*
years of education		-0.001		-0.003		0.001		0.004		0.009
		(1.42)		*(6.75)****		*(0.12)*		*(1.28)*		*(0.70)*
unemployed		-0.010		-0.026		-0.016		-0.018		-0.227
		*(2.59)****		*(5.14)****		*(0.29)*		*(0.39)*		*(1.27)*
regular wage employment		-0.008		-0.036		0.033		0.030		0.033
		*(2.10)***		*(7.13)****		*(0.74)*		*(0.89)*		*(0.23)*
casual employment		-0.004		-0.003		0.078		0.040		-0.011
		(0.69)		*(0.37)*		*(1.40)*		*(0.86)*		*(0.06)*
Constant	0.007	0.007	0.085	0.087	0.956	0.959	0.789	0.790	-2.576	-2.570
	*(1.80)**	*(1.89)*	*(17.31)****	*(17.72)****	*(23.62)****	*(23.60)****	*(25.36)****	*(25.34)****	*(20.87)****	*(20.66)****
Observations	42,035	42,035	42,552	42,552	3,277	3,277	2,689	2,689	2,136	2,136
Number of households	8,534	8,534	8,662	8,662	2,594	2,594	2,199	2,199	1,765	1,765
R-squared	0.00	0.00	0.02	0.03	0.03	0.04	0.03	0.03	0.03	0.03

Notes: Absolute value of robust t or z statistics in parentheses. *significant at 10%,**significant at 5%, *** significant at 1%. Base category for age is 0-5 years; for gender interaction terms, it is male aged 0-5 years; for labor market status, it is those out of the labor force. LPM is Liner Probability Model.

Table 8: Marginal effect of the male dummy (x100)

	CONVENTIONAL MODEL	HURDLE MODEL				
Sub-sample	Unconditional OLS (1)	Reporting sick probit (2)	Consultation probit (3)	Positive medical expenditure probit (4)	Conditional medical expenditure OLS (5)	Combined Hurdle (6)
	MED_PROP	SICK	CONSULT	POS_MEXP	LNMPROP	
pooled sample	0.08	-1.36	0.43	-4.23	5.13	-0.157
	(0.39)	*(5.42)****	*(0.31)*	*(2.69)***	*(0.97)*	*(2.40)***
age 0-5	0.12	-0.38	6.12	4.31	-1.92	0.044
	(0.88)	*(0.55)*	*(2.02)***	*(1.43)*	*(0.19)*	*(1.39)*
age 6-15	0.09	0.02	1.78	3.51	-3.43	0.012
	(0.66)	*(0.06)*	*(0.45)*	*(0.72)*	*(0.23)*	*(0.57)*
age 16-40	-0.18	-2.11	-1.60	-7.51	6.25	-0.288
	(1.00)	*(5.65)****	*(0.63)*	*(2.61)****	*(0.64)*	*(1.31)*
age 41-64	0.95	-1.70	0.75	-14.48	8.55	-0.387
	(0.66)	*(1.91)**	*(0.27)*	*(4.35)****	*(0.68)*	*(2.08)***
age 65+	-1.96	-8.57	-0.85	3.28	-12.22	-1.667
	*(1.97)***	*(5.15)****	*(0.17)*	*(0.85)*	*(0.66)*	*(1.79)**
black	0.12	-1.54	-0.12	-3.46	7.22	-0.184
	(0.40)	*(5.83)****	*(0.07)*	*(2.14)***	*(1.17)*	*(2.62)****
coloured	-0.73	-1.26	-1.12	6.02	19.92	-0.311
	(1.60)	*(1.25)*	*(0.41)*	*(1.03)*	*(1.33)*	*(0.40)*
indian†	0.03	-2.65	---	8.38	-16.25	---
	(0.08)	*(1.15)*	---	*(0.62)*	*(0.65)*	---
white	0.47	1.14	2.24	-8.02	15.93	-0.181
	(1.04)	*(1.03)*	*(0.56)*	*(1.49)*	*(0.94)*	*(0.22)*
unemployed HH member	0.56	-1.00	1.14	-2.61	9.22	-0.103
	(1.07)	*(2.82)****	*(0.66)*	*(1.21)*	*(1.38)*	*(1.25)*
no unemployed HH member	-0.34	-1.82	1.03	-6.45	3.47	-0.213
	*(2.04)***	*(4.92)****	*(0.39)*	*(2.61)****	*(0.36)*	*(3.53)****
homeland resident	0.22	-1.69	-1.07	-2.58	9.24	-0.196
	(0.52)	*(5.44)****	*(0.52)*	*(1.61)*	*(1.26)*	*(3.18)****
non-homeland resident	0.07	-0.78	2.56	-6.39	1.73	-0.104
	(0.33)	*(1.85)**	*(1.35)*	*(2.39)***	*(0.24)*	*(0.24)*
below poverty line	0.10	-1.43	-0.37	3.10	4.76	-0.193
	(0.31)	*(5.25)****	*(0.20)*	*(1.85)**	*(0.74)*	*(0.00)*
above poverty line	0.02	-1.19	1.10	-6.73	2.79	-0.128
	(0.07)	*(2.01)***	*(0.53)*	*(2.11)***	*(0.30)*	*(2.29)***

Notes: Absolute value of robust t or z statistics in parentheses. *significant at 10%,**significant at 5%, *** significant at 1%. Marginal effects reported for probit equations. Specification is as per individual results, not controlling for community fixed effects. † No ME reported for Indian families, since all males who reported sick consulted a healthcare practitioner. A coefficient on the *male* dummy thus cannot be calculated for the consultation equation.