University of Pretoria
Department of Economics Working Paper Series

On Health Insurance and Household Decisions: An Econometric Policy Analysis

Olufunke A. Alaba
University of Pretoria
Steven F. Koch
University of Pretoria

Working Paper: 2008-24
July 2008

Department of Economics
University of Pretoria
0002, Pretoria
South Africa
Tel: +27 12 420 2413
Fax: +27 12 362 5207

ON HEALTH INSURANCE AND HOUSEHOLD DECISIONS: AN ECONOMETRIC POLICY ANALYSIS

OLUFUNKE A. ALABA[†] AND STEVEN F. KOCH[‡]

Abstract. In developing countries, health insurance is not a commonly purchased financial instrument. Recent debates have revolved around extending health insurance coverage to a wider range of the population, primarily via compulsory insurance schemes. However, the debate rarely considers the competing demands placed on the family budget. In this paper, we have examined expenditure substitution patterns for both insured and uninsured households in a highly unequal developing country allowing for selection on insurance status. Our analysis suggests that expansion of health insurance coverage via compulsory schemes will create additional burdens for households, especially in terms of food purchases, and are, therefore, likely to require simultaneously implemented welfare or subsidy policies in order to be effective. It is not clear, then, that the benefits of a compulsory insurance scheme will outweigh the additional costs in terms of behavioural constraints, fiscal constraints and public sector service delivery capacity constraints.

Date: June 2008.
J.E.L. Classification. I11, G22, D14.
Key words and phrases. Treatment Effects, Hazard.
This paper represents work in progress; therefore, we ask that you do not quote the results without the authors' permission. The authors would like to thank Gauthier Tshiswaka-Kashalala for his comments on this research. All remaining errors are the sole responsibility of the authors.

1. Introduction

Health care delivery is highly segregated in developing countries, as well as in many developed countries. In these segregated health economies, a large proportion of health seekers are uninsured and purchase their health care from a subsidized public sector,[1] while a smaller proportion purchases their health care from a highly developed private sector, where purchases are often, at least partially, covered by a third-party insurance contract. For those who are uninsured, efforts to obtain health care will create out-of-pocket expenditures, and these expenditures affect their ability to purchase other items; uninsured health seekers may need to draw down their savings, sell assets, or substitute away from other household goods.

Despite potential out-of-pocket expenditures, many consumers remain uninsured. One possible explanation for the lack of insurance coverage observed in many countries is that information asymmetries result in insurance market failure, while another possible explanation is that insurance is unaffordable given current conditions. Regardless of the reason for low health insurance coverage, the initial recourse for uninsured health-seekers is the public sector, placing pressure on service delivery, which, in turn, leads other uninsured health-seekers to bypass the public sector for better quality private sector care, despite the higher costs.

In response to these pressures, governments have been pushing to increase insurance coverage. Governments and health officials have employed, or are contemplating the employment of, numerous policies, such as compulsory insurance or subsidized voluntary insurance. These policies are generally meant to ensnare formal sector employees, although the inclusion of informal sector employees in these schemes is likely to be an important feature in developing countries, given the size of the informal sector. The primary benefit of compulsory insurance is the inclusion of all consumers within the same pool, resulting in cross-subsidization. By treating everyone equally, compulsory insurance reduces the pay-off associated with assortative matching, and, thus, alleviates adverse selection, while simultaneously reducing the cost of insurance for high-risk consumers.

[1]Although there is often no direct fee for service, especially in poorer developing countries, consumers are subject to opportunity costs related to transport and queueing.

However, any policy meant to increase coverage, must also take cognizance of the underlying determinants of low coverage, as well as the subsequent consumer responses to that policy.[2] Although potential efficiency gains can arise from compulsory or social health insurance schemes, the mitigation of adverse selection may lead to the proliferation of moral hazard, because cross-subsidization reduces the individual's incentives to improve their own health, i.e., there is an economic trade-off associated with cross-subsidization. Furthermore, given that the decision to purchase insurance as a consumption good in voluntary schemes depends on the discretion of the buyer, who faces various economic and social constraints (while compulsory schemes will create additional economic burdens on those facing compulsory schemes). It is imperative for policymakers to understand the potential effects of those policies on consumer behaviour.

In this paper, we consider the possible household expenditure effects of a compulsory health insurance scheme by analysing the substitution patterns associated with the purchase of health insurance. Although gross substitution is an obvious implication of a binding budget constraint, very little evidence on substitution patterns related to health insurance exists. We apply the analysis within a developing country, South Africa, that is considering various combinations of voluntary and compulsory health insurance programmes in an effort to alleviate pressure on the public health sector.[3]

Several distinct features of the health system in South Africa make it particularly relevant to analyzing the effect of health insurance on consumer behaviour. The fact that South African health insurance is predominantly geared towards the employed creates problems for individuals who are not employed, and are thus precluded from participating in the insurance market.[4] In an attempt to fill this gap, an eclectic mix of other institutions has been introduced by the government: disability grants, old age pension grants and grants for children. However, these

[2]As shown by both Levy and DeLeire (2003) and Bundorf and Pauly (2006), health insurance coverage is not solely determined by income or poverty; rather, other factors, in addition to income, matter. These factors include, but are not limited to: the prices of other goods, preferences and the expected net benefit of insurance coverage.

[3]These pressures have been observed, for example, by Grobler and Stuart (2007), who find that private facility utilization across all income levels is extensive, despite the relatively high cost associated with those facilities.

[4]Söderland and Hansl (2000) suggest that South Africans are uninsured, because medical schemes only cater to the employed.

grants were not designed to deal directly with the costs of health care; thus, the government is gradually expanding the Employees Medical Scheme, created in 2005, to low income households, in the form of a low income medical scheme, along with other health and tax reforms. Although all of these changes are likely to help increase access, a more targeted reform would likely be more successful. Bloom and McIntyre (1998), for example, argue that the inequalities in the system require a definitive and targeted response, and that the appropriate response is a social health insurance system. However, as noted by Gilson, Doherty, Lake, McIntyre, Mwikisa and Thomas (2003) a social health insurance program was not favoured by either the Minister of Health, National Treasury or the Congress of South African Trade Unions (COSATU), albeit for different reasons.[5] Despite the failure, heretofore, to adopt a social health insurance program, the recent obligatory pension program put forth by National Treasury, suggests that requiring all South Africans to participate in a social insurance program may no longer be sacrosanct, (National Treasury, 2007). Given the recent willingness to consider compulsory programs, an analysis of the potential household level effects of a social insurance scheme can provide very valuable information for the design and implementation of such a program.

The analysis is predicated on data from the South African Income and Expenditure Survey of 2000. However, a number of empirical matters arise while considering substitution patterns, and each of these are considered in the paper. Firstly, given the fact that expenditures are part of a demand system, an appropriate system analysis must be employed for the representation of expenditure patterns, and, therefore, we make use of the Quadratic Almost Ideal Demand System (QUAIDS) developed by Banks, Blundel and Lewbel (1997). Secondly, insurance purchase is a choice for many households, and, therefore, we consider the factors that determine the household's insurance status. Finally, given that insurance status is endogenous to the expenditure system, our analysis controls for that endogeneity

[5]According to Gilson *et al* (2003), COSATU felt that it was inappropriate to expect their members to pay for something that they currently received for free, while the Minister of Health at the time thought it would be difficult to convince people to buy the service from a costly industry (the medical schemes were meant to be the primary delivery vehicle for the social insurance program) and was ideologically opposed to the industry in the first place. National Treasury, on the other hand, was concerned about the effect such a program might have on their efforts to reduce the budget deficit and the tax burden.

via regime-switching. We treat insurance status as endogenous and estimate separate expenditure systems, one for the insured and another for the uninsured, via a two-stage procedure similar to that proposed by Heckman (1979). This parametric model is then used to estimate treatment effects associated with a compulsory insurance scheme, although the results could also be used to inform a voluntary scheme, based on differences in substitution patterns that might arise from the policy.

The calculated treatment effects show that a compulsory health insurance policy would impact uninsured household expenditure behaviour in a myriad of ways, by raising expenditure shares for certain items and lowering the expenditure shares for other items. Most worryingly, compulsory insurance would reduce household food expenditure shares by 47%. Therefore, any compulsory health insurance programme designed to increase access must be implemented alongside a food subsidy scheme, otherwise, the health of poor households might actually worsen instead of improve, which is not the intended effect of the health insurance programme. However, the analysis does assume that the proposed voluntary or compulsory schemes are based upon insurance contracts that are already in force. In other words, the analysis does not assume that there is a missing health insurance market, and, therefore, the results should be taken as an upper bound to the subsidies or the effects of the policies.

The remaining paper is organized in the following fashion. We continue by examining the relevant theoretical and empirical literature within which this paper is set. The empirical model used to estimate the expenditure systems and treatment effects are developed in Section 3, while Section 4 provides a discussion of the data and an analysis of the characteristic determinants of insured and uninsured households, in terms of their structure and expenditure patterns. The empirical results and implications are discussed in Section 5. We conclude in Section 6 with a summary and discussion of policy recommendations and limitations of the current study.

2. The Demand for Insurance

The theory of demand has widely been applied to determining individual and household consumption behavior, with expenditure and price elasticity providing valuable information on consumers' reactions to price and income changes. Ehrlich and Becker (1972) note that, like any other good purchased in the market, the demand for health insurance is derived from the needs it satisfies, and, thus could be modelled within the demand framework. The key point of demand theory applied to insurance is that the demand for insurance does not only depend upon income; instead, prices and preferences also matter. Abdulla and Aubert (2004) argue that the income, price and other elasticities that can be estimated in insurance demand models are capable of providing valuable information for policymakers. However, our data, which does not include price information, requires us to find proxies for prices, allowing us to focus on the overall pattern of expenditure effects associated with health insurance, instead of elasticities.

In order to test the predictions of Ehrlich and Becker's (1972) model of insurance demand, a number of studies have applied choice models to determine health insurance coverage amongst various populations. Along those lines, Ying, Hu, Chen, Xu and Huang (2007) find that education, income, employment, and age affect the demand for health insurance in urban China, while Doiron, Jones and Savage (2008) find that the uninsured population in Australia is younger, less educated, more likely to be a single parent and to smoke, but less likely to be employed or to undertake regular exercise. Kirigia, Sambo, Nganda, Mwabu, Chatora and Mwase (2005) find that health insurance cover is more likely for older South African women with children. Marquis, Beeuwkes, Escarce, Kapur and Yegian (2004), who made use of a number of different US population, income and employment surveys, find that health insurance is affected by the premium, employment status, the poverty ratio, education, race, health status and the availability of safety nets for the family, while the uninsured were primarily younger, poorer and in poorer health. In terms of demand concepts, they also find that the income elasticity is low. Bundorf and Pauly (2006) also show that income, alone, cannot explain insurance choices. They define behavioural and normative affordability, and use data from the 2000

Medical Expenditure Survey in the United States to show that health insurance is affordable for 82% of the uninsured population, although that percentage is not robust to changes in the underlying poverty threshold. In addition to choice models, the actual value of total insurance coverage, inclusive of health, auto, home and life insurance, has been considered by Showers and Shotick (1994). In their Tobit model, they find that household income, the age of the household head, the size of the household and the number of income earners in the household matter; similar to Marquis et al's (2004) results, their estimated income elasticity was low. These studies provide the basis for the discrete choice model that we estimate for household insurance status.

Although insurance has many advantages; by pooling risks, higher risk individuals can be subsidized by lower risk individuals. Gao, Tang, Tolhurst and Rao (2001) suggest that health insurance is one of the most effective means of risk-pooling and risk-sharing capable of affecting health care utility and health care financing. However, these benefits come at a cost. Rothschild and Stiglitz (1976) show that in markets with imperfect information regarding ex ante health risks, adverse selection and incentives for assortative matching can arise, causing the market to fail. Adverse selection in health insurance obtains when there is a positive relationship between health risk and health insurance coverage. However, much of the empirical literature suggests that there is a negative relationship between health risk and insurance coverage; see, for example, Doiron *et al* (2008), Shumeli (2001) and Propper (1989). On the other hand, Arrow (1963) shows that markets in which agent actions are not completely observable or verifiable are also characterised by market failures. In the case of health insurance, Zweifel and Breyer (1997) point to two types of hidden actions: ex ante moral hazard, under which insufficient preventative care is sought, and ex post moral hazard, under which overconsumption of health care is one outcome. Zweifel and Manning (2000) argue that efficiency gains can result from insurance coverage, because optimal health care moral hazard is positive. Empirically, Manning, Newhouse, Duan, Liebowitz and Marquis (1987) use data from the RAND health insurance experiment to show that health services consumption increased in the USA, when levels of out-of-pocket expenditures decreased. In Australia, Cameron, Trivedi, Milne and Piggot (1988) find that insurance coverage

is associated with an increase in service utilization. Research in Switzerland, by Holly, Gardiol and Domenighetti (1998), in Ireland, by Harmon and Nolan (2001), and in France, by Chiappori, Durand and Geoffard (1998) have found similar results to those already reported for the USA and Australia. Insurance moral hazard, although widely studied in developed countries, has not been widely studied in developing countries; however, Jowett, Deolalikar and Martinsson (2004) find that insured individuals are more likely to use more costly inpatient services and public providers in Vietnam. They also find that lower income is associated with increased service usage.

Although the effects of insurance, as highlighted by the preceding literature focus primarily upon health care utilization and the risk profile of the insured, other aspects of Ehrlich and Becker's (1972) model of insurance have received less attention. One of those effects, and the effect that is the focus of the analysis presented in this paper, is the effect that the purchase of insurance might have on other consumption activities. For example, Chou, Liu and Hammit (2003) estimate the effect of a change in Taiwan's insurance regulations on precautionary savings. In their paper, they are able to estimate treatment effects via a difference-in-difference estimator, due to the fact that the change in the law was isomorphic to a social experiment. In their analysis, they find that savings decreased from between 5.1% and 13.7%, while consumption increased by up to 5.7%, depending upon the controls included in the regressions. Levy and DeLeire (2003), on the other hand, consider the entire expenditure system, and examine the differences in expenditure behaviour of insured households compared to uninsured households. They show that expenditure patterns for the uninsured poor differ from the insured poor in a way that does not arise when comparing the wealthier insured and uninsured. The main criticism to be leveled at Levy and DeLeire's (2003) analysis is that insurance status is treated exogenously, such that the reported differences are hard to interpret. Below, we modify Levy and DeLeire's analysis to take account of the potential endogeneity associated with insurance status to estimate the possible effect that a compulsory insurance policy in an unequal developing country, like South Africa, might have on households.

3. Identifying Selection in a Demand System

3.1. **The Empirical Specification.** The model developed for this analysis focuses on two issues at the level of the household; the first is general consumption expenditure, and the second is health insurance status. Although households are a collection of individuals, and there are likely to be interesting internal decision processes within the household, the data collection process treats the household as the unit of analysis. Since the data does not allow for the empirical identification of the intra-household economy, a unitary household model is presumed here.[6]

The key assumption underlying the model is that household preferences differ by insurance status, such that the empirical model must account for potential selection effects. However, the underlying model of behaviour is predicated on QUAIDS developed by Banks *et al* (1997). Therefore, selection effects must be accommodated within an expenditure share system. Therefore, the empirical model is a demand system that accounts for regime-switching across health insurance status, similar to the model developed by Roy (1951).[7]

The insurance regime is determined by a standard binary choice model, where the latent health insurance demand function is assumed to be linear in the parameters.

$$h^* = z\delta + \mu \tag{1}$$

Following standard convention, we observe $h = 1$, if $h^* \geq 0$. Each expenditure share, on the other hand, is assumed to be generated from the following stochastic relationship, modified from Banks *et al* (1997), equation (10).[8]

$$w_\ell^h = A_\ell^h + B_\ell^h \ln x^h + C_\ell^h (\ln x^h)^2 + d\psi_\ell^h + \epsilon_\ell^h \qquad \forall \ell \in \{1, \ldots, J\}, h \in \{0, 1\} \tag{2}$$

[6]In other analysis with this data, Koch (2007) has shown that single person households do not significantly differ from multiple person households, such that a unitary model is a reasonable approximation in South Africa.

[7]Tshiswaka-Kashala and Koch (2008) and John (2008) use QUAIDS to estimate crowding-out associated with tobacco expenditure in South Africa and India, respectively, based on Vermeulen's (2003) model of conditional demand. The model presented here is more general, since it allows for regime-switching within the system, rather than just an intercept effect.

[8]Although the data generating process for the shares can be written in a more succinct matrix representation, singling out one equation within the system and across the binary response function is useful for further discussion regarding the correlations between the unobserved determinants of both expenditure shares and insurance status.

In the preceding equation, d represents various household and community level factors affecting autonomous expenditure, while $\ln x^h = \ln m^h - \ln P_c$ is total household expenditure, $\ln m^h$, adjusted for cluster-level heterogeneity in minimum expenditure, P_c.[9] The parameters A_ℓ^h, B_ℓ^h, C_ℓ^h, and ψ_ℓ^h are to be estimated for each share, ℓ, for both insured and uninsured households.

For any single share, there are three potentially correlated stochastic variables: the unobserved component affecting insurance status, the unobserved component affecting the observed expenditure share for the insured households, and the unobserved component affecting the observed expenditure share for the uninsured households. The within-share error structure is assumed to be multivariate normally distributed. The description of that structure is provided below, and is limited to just one share within the system, for notational convenience. The remaining unexplained share components are assumed to have the same structural relationship with the unexplained components of insurance status.

$$\begin{pmatrix} \mu \\ \varepsilon_\ell^0 \\ \varepsilon_\ell^1 \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{\ell\mu 0} & \sigma_{\ell\mu 1} \\ \sigma_{\ell\mu 0} & \sigma_{\ell 0}^2 & \sigma_{\ell 01} \\ \sigma_{\ell\mu 1} & \sigma_{\ell 01} & \sigma_{\ell 1}^2 \end{pmatrix} \right] \tag{3}$$

In a switching model, such as this, it is assumed that insured and uninsured households are identifiably separable. In particular, it is assumed that the insured expenditure shares are only observed for the insured, while the uninsured shares are only observed for the uninsured. For that reason $\sigma_{\ell 01} = 0$, i.e., no correlation between uninsured and insured expenditures is observed. In order to examine the effect of insurance on the uninsured, treatment effects will be calculated. These effects are based on the parametric specification outlined in equations (1), (2) and (3).

Treatment effects are becoming increasingly common in the literature;[10] however, the analysis presented here is focussed upon two specific within-group effects:

[9]Banks *et al* (1997) stress that the quadratic expenditure component cannot be independent of prices in the QUAIDS model. However, our data does not contain price information; therefore, we use minimum expenditure within survey clusters to control for price differences within each cluster. Although minimum expenditure does control for the fact that higher income households are likely to purchase more expensive items, the quality of the items purchased by these households is also likely to be better. In other words, although the control is not perfect, it is likely to incorporate the same problems as other price indexes, and is, therefore, a reasonable proxy.

[10]Heckman and Vytlacil (2005) present a unifying theory of the disparate treatment definitions and their interpretations, and provide an excellent summary of the literature to date.

the average effect of treatment on the untreated (ATEUT), which is our primary interest, and the average effect of treatment on the treated (ATET), which is presented for symmetry. In expectation, $\text{ATET}_\ell = E[w_\ell^1|x, h = 1] - E[w_\ell^0|x, h = 1]$, where the last term is the expected value of uninsured expenditure share ℓ for those currently insured, which is an unobserved value. On the other hand, $\text{ATEUT}_\ell = E[w_\ell^1|x, h = 0] - E[w_\ell^0|x, h = 0]$, where the first term measures the expected value of the insured expenditure share ℓ for those currently uninsured, which is also unobserved. Mathematically,[11]

$$\begin{aligned}
\text{ATET}_\ell &= E\left[w_\ell^1|x, h = 1\right] - E\left[w_\ell^0|x, h = 1\right] \\
&= x\beta_1 + \sigma_{\ell\mu 1}\frac{\phi\left(z\delta\right)}{\Phi\left(z\delta\right)} - \left(x\beta_0 + \sigma_{\ell\mu 0}\frac{\phi\left(z\delta\right)}{\Phi\left(z\delta\right)}\right) \\
&= x\left(\beta_1 - \beta_0\right) + \left(\sigma_{\ell\mu 1} - \sigma_{\ell\mu 0}\right)\frac{\phi\left(z\delta\right)}{\Phi\left(z\delta\right)}
\end{aligned} \tag{4}$$

Via a similar derivation,

$$\text{ATEUT}_\ell = x\left(\beta_1 - \beta_0\right) + \left(\sigma_{\ell\mu 0} - \sigma_{\ell\mu 1}\right)\frac{\phi\left(z\delta\right)}{1 - \Phi\left(z\delta\right)} \tag{5}$$

The empirical foundation of these treatment effects is a switching regime model, though it is classified as one of many sample selection models resulting from Roy's (1951) model. Importantly, failure to control for self-selection bias in the estimating equations will lead to biased and inconsistent parameter estimates, and will further bias the underlying estimated treatment effects, which is our primary concern.

The estimation method adopted in this paper is the two-step procedure advanced by Heckman (1979) and Lee (1978) to account for sample selection, but adapted to include the entire demand system.[12] In our case, sample selection bias arises, because the choice of insurance status may have a common set of unobservable factors, in which case the error term in the insurance status choice model and the error term in each of the consumption shares would be correlated.[13] In the

[11]For derivation of these effects, see, for example, Maddala (1983).

[12]Shonkwiler and Yen (1999) consider a tobit demand system that underpins our modelling approach. In our analysis, we ignore the fact that many households do not buy certain commodities, and, instead, focus on expenditure share averages inclusive of zeroes, due to the difficulty in separately identifying both insurance status and positive consumption within each share equation. In calculating our treatment effects, however, we truncate any negative predictions at zero, in order to keep all shares within the unit interval.

[13]Although we can control for insurance status, it is possible that households are not exogenously formed, which would affect the interpretation of the results; however, consistent with many other

first stage, a probit is estimated, based on equation (1), where prob$(h = 1|z) =$ prob$(\mu \geq -z\delta) = \Phi(z\delta)$.[14] The probit results are used to create estimates of the insurance and non-insurance hazard rates, $\phi(z\delta)/\Phi(z\delta)$ and $\phi(z\delta)/(1-\Phi(z\delta))$, respectively; see equations (4) and (5). To control for selectivity, these estimated hazard rates are placed in equation (2) for the appropriate insured or uninsured share, which is then estimated as part of a system, and that system is estimated for both the insured and uninsured households. The estimated systems, share by share, are written:

$$w_\ell^1 = A_\ell^1 + B_\ell^1 \ln x^1 + C_\ell^1 (\ln x^1)^2 + d\psi_\ell^1 + \rho_\ell^1 \frac{\phi(z\delta)}{\Phi(z\delta)} + \epsilon_\ell^1 \tag{6}$$

$$w_\ell^0 = A_\ell^0 + B_\ell^0 \ln x^0 + C_\ell^0 (\ln x^0)^2 + d\psi_\ell^0 - \rho_\ell^0 \frac{\phi(z\delta)}{1-\Phi(z\delta)} + \epsilon_\ell^0 \tag{7}$$

Given the fact that the model is estimated across two stages, the underlying standard errors are incorrectly estimated. Therefore, the model standard errors are predicted via a 1000-replication bootstrap.

The selectivity terms, represented by $\rho_\ell^h[\cdot]$, measure the covariance between the unexplained components of insurance status and each expenditure share. Therefore, the statistical significance of the estimated coefficient, ρ_ℓ^h, provides useful information about the extent to which the two decisions are interrelated. Although, according to McManus (1992), identification can be secured by non-linearity of the functional form, exclusion restrictions, if available, ensure a more robust identification of the causal simultaneous model (Heckman, 2000). The exclusion restrictions applied in this analysis are discussed below.

3.2. **Variables and Identification Strategy.** In the IES data, an insured household is identified by reported expenditure on any kind of health insurance, while an uninsured household does not report any expenditure on health insurance. In terms of expenditure categories, an aggregation decision had to be made, due to the fact that the IES contains information on expenditures for hundreds of detailed categories. For this analysis, we aggregated these detailed expenditures into 14 monthly expenditure categories. These are: (1) Food, (2) Housing, (3) Tobacco

demand analyses, for example Handa (1996), this study takes household composition as exogenous; hence all results are conditional upon this assumption.

[14] Φ represents the normal cdf, which is symmetric, and, therefore, $1 - \Phi(-z\delta) = \Phi(z\delta)$.

and alcoholic beverages, (4) Non-alcoholic drinks, (5) Education, (6) Health Care, (7) Household utilities, (8) Communication, (9) Recreation, (10) Transport, (11) Appliances and furniture, (12) Clothing, (13) Personal care items and (14) Other expenditures. For a description of these categories and the trimmed sample (see below) descriptive statistics, see Table A1 in the appendix.

The independent variables used in the analysis are classified into three categories, namely: basic individual and household variables, capability controls and health risk controls, the latter of which are used primarily to identify selection into health insurance coverage. The basic individual and household variables include: the household size, the proportion of dependents in the household (proportion of children under one, proportion of children between one and five and the proportion of children between six and sixteen), the sex of the household head, and dummy variables indicating the population group of the household head. The capability factors control for the economic status of the household, including employment and financial empowerment, which are associated with financial security. Included in these factors are: the number of working adults in the household (dummy variables for whether the household has zero, one or two or more workers), the proportion of working adults in the household, a set of dummies reflecting home ownership (homes owned without mortgage, with mortgage, rents and occupying a home without paying rent), and household income, although total expenditure is used to proxy for income.[15] Descriptive statistics for the variables included in the analysis, by household insurance status, are contained in Table A2.

In selection models, such as the one conducted here, it is necessary to identify several potential exclusion restrictions that are correlated with household insurance status, but not with the underlying expenditure shares. In our analysis, we have identified four such potential restrictions, all of which address either preferences for health insurance, health production or the price of insurance. The expenditure data that we use neither contains price data nor obvious inputs to the production of health. However, the data does contain information on the quality of sanitation

[15]Deaton (1997) argues that expenditure is a better measure of household resources than income, since it is more likely to be correctly recorded. Labour Force Surveys in South Africa, which explicitly try to collect income information are fraught with non-responses for income (Daniels, 2008), further validating our choice of expenditure over income.

infrastructure embedded in the household's dwelling, while additional information can be gleaned from other sources, and merged into the data, to be used as instruments. From the survey, we are able to include a dummy variable for whether or not the household contains a flush toilet, and another for whether or not the dwelling contains internal piped water. Each of these variables is related to household sanitation, and, therefore, potentially measures household preferences for health or inputs into the household's production of health. If sanitation provides an accurate measure of household preferences for health or if sanitation and health insurance are complementary inputs to the production of health, we would expect improved sanitation to point towards an increased likelihood of also purchasing health insurance, after controlling for other effects. However, household infrastructure variables are also likely to be related to household wealth, and, therefore, correlated with share purchases. In other words, embedded infrastructure variables may be weak instruments. Humphreys, Phibbs and Moos (1996) suggest that it would be better to look beyond the individual to programs and community features that may be related to participation, since these features are less likely to be correlated with, in our case, expenditure shares. In accordance with this principle, we make use of provincial information describing health status and human development. In particular, we include the rate of HIV prevalence from antenatal clinic data and the human development index from the province to proxy for the underlying health status and the actuarial relationship between community health-risk ratings and health insurance premiums.[16] We expect that increases in the HIV prevalence rate would be associated with generally higher insurance premiums in the province, and thus, a lower probability of insurance purchase. Furthermore, we expect that higher levels of human development are observed in provinces with better health status, overall, and, therefore, a reduced need for health insurance for households in the province.

4. The South African Income and Expenditure Survey, 2000

The data for our analysis is sourced from the Income Expenditure Survey (IES) of South Africa, conducted in October 2000 by Statistics South Africa. The IES is a quinquennial cross-sectional survey based on the master sample of South African

[16]Efforts are underway to directly source actuarial data from the Council of Medical Schemes in South Africa that would allow us to predict the household's expected health insurance premium.

census enumeration areas; in this case, the sample is based upon the 1996 census. Weights are included in the survey to make the estimation results nationally representative; however, the weights are a source of contention in the survey, as highlighted by Van Walbeek (2005) and Koch (2007), amongst others. Due to the contention surrounding the weights, the weights are not used in the analysis; thus, the results can only be interpreted relative to the sample.

The data set provides detailed expenditure and demographic information on 26249 households. Given the nature of the data, it has been used extensively for consumption and income studies by Burger, Van der Berg and Nieftagodien (2004), Simkins (2004), and Koch (2007), as well as for the analysis of poverty in South Africa; see, for example, Özler and Hoogeveen (2005). The data has also been used to consider the consumption patterns of specific products. For example, tobacco and alcohol consumption have been examined by Van Walbeek (2005), Ground and Koch (2008), Ground, Koch and Van Wyk (2008), as well as Tshiswaka-Kashala and Koch (2008). To the best of our knowledge, the data has not been used to examine the effects of health insurance.

Given the data quality issues identified by various authors, including Burger et al (2004), Simkins (2004) and Van Walbeek (2005), and in order to minimize measurement errors, the top and bottom one percent of basic need expenditures, based on food expenditure, were trimmed. The result of trimming was a remaining sample of 22494 households; 3300 insured households and 19194 uninsured households.[17]

The main shortcoming of the data set, relevant to this analysis, is the lack of information on health status or other indicators of risky health behaviour related to insurance coverage. These shortcomings are to be expected, because the IES was designed for calculating the consumption basket used to underpin the South African Consumer Price Indexes, CPI and CPI-X. Despite the targeted purpose of the IES, some health indicator proxies are available in the data set, and are used in the analysis. Therefore, the IES is a suitable data source for the analysis of policies that might affect consumption behaviour, as is done here.

[17]The analysis was also performed for the entire sample. Symmetric trimming lowers average expenditure in all commodities, but does not systematically affect the estimated coefficients. Results are available from the authors.

5. The Results

5.1. **Household Insurance Status.** Households are assumed to select their insurance status, either by virtue of being employed, or by virtue of their own optimising decisions. Insurance status selection is estimated via a binary response model, based upon a probit specification. The results of the probit estimates are reported in Table 1. These results are generally as expected, given the descriptive statistics in Table A2. In line with Söderland and Hansl (2000), employment is an important determinant of health insurance; a household with an employed household head, as well as households with a greater proportion of working adults is more likely to purchase health insurance.[18] Furthermore, insurance status is concave in total expenditure, which is a proxy for household income, while urban households are more likely to purchase health insurance. Given South Africa's segregationist history, it is not surprising that there are racial differences in health insurance status, and that the differences place Africans at a disadvantage, compared to other races. Finally, households with more young children were more likely to purchase insurance than households with older children, which is consistent with younger children needing more access to health care than older children. Furthermore, the results are consistent with estimates presented by Kirigia *et al* (2005) and Marquis *et al* (2004), who find that households with younger children are more likely to have health insurance.

In an effort to account for selection, our analysis employed multiple exclusion restrictions to separately identify insurance status from household expenditure behaviour.[19] Each of the exclusion restrictions had the expected effect. Households with better sanitation infrastructure were more likely to purchase insurance, suggesting that inputs to health production are complementary. Furthermore, the provincial HIV prevalence rate, which we use to proxy for health-risk, and, thus,

[18]The employment measure in the Income and Expenditure Survey is crude. Rather than using the ILO definition of employment that requires someone to not be working, but interested enough in a job to search, the IES only asks whether or not each individual in the household undertook any activities for pay in the week preceding the survey. A measure of the proportion of working household members, however, was unsuccessful in explaining access to insurance.

[19]In the table, we only report one probit result, although many other specifications were estimated, including separate estimations for models containing only one of each of the exclusion restrictions. The choice of excluded variables did not affect the overall results in either the probit equations or share equations. Other results are available from the authors, upon request.

provincial health insurance premiums, reduces the probability of health insurance.[20] Finally, human development is negatively correlated with health insurance status, implying that, once other features are controlled for, a generally healthier population has less need for health insurance.

5.2. **Selection within the Demand System.** The preceding probit estimates were used to predict the hazard of participation, which was included in the demand system for the insured, as well as the hazard of non-participation, which was included in the demand system for the uninsured. The estimated hazard rate indicates the degree of correlation between the unobserved components of expenditure and health insurance status. The expenditure share estimates, inclusive of the participation hazard, i.e., for the insured, are presented in Table 2. In 10 of the 14 share equations, the participation hazard is significant, although the direction of correlation differs from share to share. For housing, education, household utilities, communication and transport, the correlation is negative; for food, tobacco and alcohol, nonalcoholic beverages, clothing and personal care, the correlation is positive; for the remainder, the correlation is insignificant. These correlations, however, are driven by unexplained variables, such as prices and health status. Therefore, one interpretation is that these correlations explain substitution or complementarity between household expenditure shares and health insurance. Given the estimated direction of correlation, insured households, on average, could view health insurance as complementary to housing, education, household utilities, communication, recreation and transport, while viewing health insurance, food, tobacco and alcohol, clothing, furnishings, personal care and miscellaneous items as substitutes.[21]

The uninsured expenditure share equations, after correcting for selection, are presented in Table 3. The correlation patterns between unobserved components are broadly similar for the insured and uninsured households, although the magnitudes differ. If the driving force behind the unobserved components is the relative price of insurance, the uninsured households view miscellaneous items as substitutes

[20]Given that prevalence rates are derived from antenatal clinics, which are public facilities, high prevalence rates could be revealing reduced access to private health care. If so, high prevalence rates would be a proxy for lower expected private health care benefits, and, therefore, a reduced demand for insurance. Our results cannot separate these interpretations.

[21]An increase in the price of product reduces the demand for a complementary good and raises the demand for a substitute good.

for health insurance, while the insured households view miscellaneous items as complements. Otherwise, patterns of substitution and complementarity are the same across the insured and uninsured households. Therefore, analysing insured and uninsured households separately is reasonable.[22]

The observed patterns of substitution and complementarity are not all that surprising, once one considers the developing country backdrop. The complementarity between housing and insurance agrees with the notion that only a small percentage of the population has access to health insurance, and that these same individuals tend to have both better jobs and nicer houses. Complementarity between health insurance, education, household utilities, communication and transport can be supported with a similar argument. Namely, these products are more likely to be purchased by those who are employed, and the employed also have better access to health insurance. On the other hand, the substitutability between health insurance, food and clothing, is also understandable. For those without employment based access to health insurance, it is necessary to purchase health insurance directly from earnings. Given that public health care is available, those without insurance, especially the poorest, will prefer to spend their income on other necessities. One additional substitution pattern is worth highlighting, and that is substitutability between health insurance and tobacco and alcohol products. Substitution between these products is expected if, for example, tobacco and alcohol consumption is reflective of risky behaviour, and people with preferences for risky behaviour are less likely to purchase risk reducing financial instruments, like health insurance.

5.3. **Treatment Effects Related to Insurance Status.** The preceding results are consistent with the assumptions made in the analysis, that insured and uninsured households are different, and can be treated separately. In other words, the estimated effect of a compulsory insurance programme is likely to differ by insurance status. For that reason, we have estimated two treatment effects, the effect of treatment on the untreated (uninsured), which is the more relevant effect, and

[22]For comparison purposes, the non-selection insured and uninsured household expenditure share estimates are provided in Tables A3 and A4. The most noticeable effect of including the participation hazard is the effect on the estimated parameters; see either Table 2 and Table A3 or Table 3 and Table A4. As expected the non-corrected estimates are attenuated, due to omitted variables bias.

the effect of treatment on the treated (insured). These estimated treatment effects are reported in Table 4, based on equations (4) and (5) in the text.[23] In terms of considering the economic consequences of a national health insurance programme, the average effect of treatment on the untreated is the more relevant measure, since they are currently without insurance; however, we present both treatment effects in the interest of symmetry. As can be seen in the table, the effect of treatment on the insured is opposite the effect on the uninsured.

We begin by considering the share of expenditure devoted to items that are normally classified as necessities: food, clothing and housing. For those currently insured, the average effect of insurance, i.e., the average effect of treatment on the treated, is estimated to be -16.0% for housing, 30.8% for food and 5.5% for clothing.[24] In other words, if insurance were taken away from currently insured households, the share of expenditure that they devoted to housing would be 16% higher, the share devoted to food would be nearly 30% less and the share devoted to clothing would be nearly 6% less. For uninsured households, on the other hand, compulsory insurance would raise the share of expenditure devoted to housing by 23.5%, lower the share devoted to food by 47.1% and lower the share devoted to clothing by 2.7%. In other words, requiring uninsured households to purchase insurance would lower expenditure on two of the three basic necessities, one of them by an economically significant amount. Although not all uninsured households are poor, the majority are. Requiring these households to participate in an insurance scheme would have negative effects on their ability to feed and clothe themselves. As shown by Levy and DeLeire (2003), health insurance is one of many choices that households can make, and, therefore, constraining behaviour, in such a way that insurance must be purchased, is a burden to many households.

On the other hand, extending insurance to the uninsured would also reduce the expenditure share devoted to tobacco and alcohol by 6.9%, raise the share of expenditure for transport by 6.5% and raise the education expenditure share by 3.1%. In other words, compulsory insurance might have other positive effects that

[23]Standard errors for these treatment effects were bootstrapped to accommodate for the fact that they are calculated from estimated parameters.
[24]Due to the fact that shares were normalized to fall between 0 and 100, these treatment effects can be interpreted as share percentages.

counteract the negative effect insurance has on food expenditure. However, for the poor uninsured, whose tobacco and alcohol, education and transport shares are small, the negative food impacts are likely to swamp these positive effects. However, taking health insurance away from those already insured would raise the tobacco and alcohol expenditure share by 10.2%, and lower the education and and transport expenditure shares by 6.4 and 6.0%, respectively. As noted in the preceding subsection, health insurance does substitute for certain household purchases.

To summarize, the effect of compulsory insurance on the uninsured is complicated. Due to the fact that these estimates are incorporated within a system, the treatment effects generally offset each other, in percentage terms. In other words, the analysis highlights the fact that household consumption behaviour is determined by economic constraints and opportunities, changing these constraints and opportunities, as would occur if a compulsory health insurance system were to become law, will affect behaviour. Most worryingly, if such a policy were to be put into place, our model suggests that it would have dire consequences for household food consumption amongst uninsured households, further clouding the supposed benefits of a national health insurance policy. Assuming that food consumption is positively related to the health of the populace, it is possible that compulsory health insurance would have the unintended consequence of lowering the general level of health in the population, which could also reduce labour productivity, lower the capacity to learn and, most distressingly, harm future growth prospects for South Africa.

6. Conclusions and Recommendations

In this paper, we have presented estimates of the treatment effects associated with health insurance on household expenditure behaviour. In order to estimate those treatment effects, we have employed a parametric specification, via a regime switching model, based on multivariate normality. These estimated treatment effects are calculated based on counterfactuals, i.e., we consider expenditure behaviour for the insured compared to their expenditure under the counterfactual of not having access to insurance, as well as the expenditure behaviour for the uninsured compared to their predicted expenditure under the counterfactual of

being able to access insurance. In the analysis, we found that the treatment effects differed by current insurance status. The most worrying result arising from the analysis is the estimated reduction in food expenditure amongst the currently uninsured households, assuming that they were forced to purchase health insurance. If the estimated effects are correct, a national health insurance programme would have unintended negative consequences related to the productivity of human capital in the country, which could negatively impact economic growth in the long-run. Despite the aforementioned potential negative consequences of requiring health insurance for those currently uninsured, the analysis also points to a reduction in tobacco and alcohol expenditure, as well as an increase in education expenditure. Each of these changes would partially offset the food share reduction, although not by enough to counterbalance it.

The results reported in this paper are another reminder that even good-intentioned social policies can have negative consequences, and that those consequences must be weighed up against the expected benefits of the social policy. The research presented here cannot speak to the net benefits of a national health insurance programme, since it only focusses upon household expenditure behaviour. However, the results show that not all effects are necessarily positive. The analysis suggests that compulsory insurance would need to be implemented along with subsidies devoted to reducing the price of food, which will have consequences for the nation's fiscus. Therefore, additional research is needed to provide a more complete picture of the potential effects of a national health insurance policy, especially one based on compulsory participation, by examining other aspects of a national health insurance programme. In particular, future research should address the effect of health insurance on other measures of health capital, such as risky sexual behaviour and healthcare utilization. In addition, additional detailed calculations are necessary, including an estimate of the underlying food subsidy that would be required to offset the effect of treatment on the untreated's food expenditure, and a comparison of the negative consequences of decreased food expnediture with the potential positive impacts from increased education and reduced tobacco and alcohol consumption.

There are limitations in our analysis. Importantly, we assume that the proposed voluntary or compulsory schemes are based upon insurance contracts that are already in force. In other words, the analysis cannot directly consider the generosity of the proposed insurance plan, and, therefore, the results should be taken as an upper bound to the subsidies or the effects of the policies. As already noted, there are two very important missing variables in the analysis. The first of which is prices, the second of which is household health status. Although we have used a number of proxies to control for these missing variables, there is no substitute for better information. Future research must strive to uncover additional data to provide a more complete picture of proposed national health insurance plans in both developing and developed countries.

References

[1] Abdulai, A. and Aubert, D. 2004. A cross-section analysis of household demand for food and nutrients in Tanzania. *Agricultural Economics* , 31(1):67–79.

[2] Arrow, K. 1963. Uncertainty the Welfare Economics of Medical Care. *The American Economic Review*, 53(5):941-967.

[3] Banks, J., Blundell, R. and Lewbel, A. 1997. Quadratic Engel curves and consumer demand. *The Review of Economics and Statistics*, 79(4):527–39.

[4] Bloom, G. and McIntyre, D. 1998. Towards equity in Health in an Unequal Society *Social Science and Medicine* , 47(10):1529–38.

[5] Bundorf, M. and Pauly, M. 1997. Is health Insurance affordable for the uninsured? *Journal of Health Economics* , 25(4):650–73.

[6] Burger, R., van der Berg, S. and Nieftagodien, S. 2004. Consumption patterns of South Africa's rising black middle-class: Correcting for measurement errors. Paper delivered at the conference of the Centre for the study of African Economies (CSAE) on *Poverty reduction, Growth and Human development in Africa* Oxford, March 2004.

[7] Cameron, A., Trivedi, P., Milne, F. and Piggott, J. 1988 A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia. *Review of Economic studies*, 55(1):85-106.

[8] Chiappori, P., Durand, F. and Geoffard, P. 1998. Moral hazard and the demand for physician services: First lessons from a French natural experiment. *European Economic Review*, 42(3-5):499-511.

[9] Chou, S., Liu, J. and Hammitt, J. 2003. National health Insurance and precautionary saving:evidence from Taiwan. *Journal of Public Economics*, 87(9-10):1873-1894.

[10] Daniels, R. 2008. The income distribution with coarse data. *Economic Research Southern Africa Working Paper 82*, University of Cape Town.

[11] Deaton, A. 1997 The Analysis of Household Surveys.John Hopkins University Press, Baltimore, MD, USA.

[12] Dorion, D., Jones, G. and Savage, E. 2008. Healthy, Wealthy and Insured? The role of Self-Assessed health in the Demand for private Health Insurance. *Health Economics* , 17(3):317–34.

[13] Ehrlich, I. and Becker, G. 1972. Market Insurance, self-Insurance and Self-Protection. *The Journal of Political Economy*, 80(4):623-648.

[14] Gao, J., Tang, S., Tolhurst, R. and Rao, K. 2001. Changing access to health services in urban China: implications for equity. *Health Policy and Planning* , 16(10):302–312.

[15] Gilson, L., Doherty, J., Lake, S., McIntyre, D., Mwiska, C. and Thomas, S. 2003. The SAZA study: implementing health financing reform in South Africa and Zambia. *Health Policy and Planning*, 18(1):31–46.

[16] Grobler. C. and Stuart, I. C. 2007. Health care provider choice. *South African Journal of Economics*, 75(2):327–350

[17] Ground, M. and Koch, S.F. 2008. Hurdle Models of Alcohol and Tobacco Expenditure in South African Households. *South African Journal of Economics*, 76(1):132–143.

[18] Handa, S. 1996. Expenditure behaviour and children's welfare: An analysis of female headed households in Jamaica. *Journal of Development Economics* , 50(1):165–187.

[19] Harmon, C. and Nolan, B. 2001 Health insurance and health services utilization in Ireland.*Health Economics*, 10(2):135-145.

[20] Heckman, J.J. 1979. Sample selection bias as a specification error. *Econometrica* , 47(1):153–161.

[21] Heckman, J.J. 2000. Causal parameters and policy analysis in economics: A twentieth century retrospective. *Quarterly Journal of Economics* , 115(1):45–97.

[22] Heckman, J.J. and Vytlacil, E. 2005. Structural Equations, Treatment Effects, and Econometric Policy Evaluation. *Econometrica* , 73(3):669–738.

[23] Holly, A., Gardiol, L. and Domenighetti,G. 1998 An econometric model of health care utilization and health insurance in Switzerland. *European Economic Review*, 42(3-6):513–522.

[24] Humphreys, K., Phibbs, C. and Moos, R. 1996. Addressing self-selection effects in evaluations of mutual help groups and professional mental health services: an introduction to two-stage sample selection models. *Evaluation and Program Planning* , 19(4):301-308.

[25] John, R. 2008. Crowding out effect of tobacco and its implications on household resource allocation in India. *Social Science & Medicine*, 62(6):1414–1426.

[26] Jowett, M., Deolalikar, A. and Martinsson. P., 2004. Health Insurance and treatment seeking behaviour: evidence from a low-income country. *Health Economics*, 13(9):845–857.

[27] Kirigia, J., Sambo, L., Nganda, B., Mwabu, G., Chatora, R. and Mwase, T. 2005. Determinants of health insurance ownership among South African Women. *BMC Health Services Research* , 5(17).

[28] Koch, S.F. 2007. The AID and MAID Systems: South African Household Data Pitfalls. *Studies in Economics and Econometrics*, 31(1):1–28.

[29] Koch, S.F., Ground, M. and van Wyk, D. 2008. South African Household Expenditure Patterns: Alcohol Products in 1995 and 2000. *Acta Academica* 40(1).

[30] Lee, L-F. 1978. Unionism and wage rates: a simultaneous equation model with qualitative and limited dependent variables. *International Economic Review* , 19(2):415–433.

[31] Levy, H. and DeLeire, T. 2003. What do people buy when they don't buy health insurance and what does that say about why they are uninsured? *National Bureau of Economic Research (NBER) Working paper 9826*, Cambridge.

[32] Maddala, G. 1983. Limited dependent and qualitative variables in econometrics. Cambridge University press, Cambridge, UK.

[33] Manning, W., Newhouse, J., Duan, N., Liebowitz, A. and Marquis, S. 1987. Health Insurance and the demand for medical care: results from a randomized experiment. *American Economic Review*, 77(3):251–276.

[34] Marquis, M., Beeuwkes, M., Escarce, J., Kapur, K. and Yegian, J. 2004. Subsidies and the demand for health insurance in California. *Health Services Research*, 39(5):1547–72.

[35] McManus, D.A. 1992. How Common is Identification in Parametric Models. *Journal of Econometrics*, 53(1-3):5–23.

[36] National Trasury. 2007. Social Security and Retirement Reform. *Second Discussion Paper*, pp 39.

[37] Propper, C. 1989. An econometric analysis of the demand for private health insurance in England and Wales. *Applied Economics*,21(6):777–792.

[38] Özler, B. and Hoogeveen, J.G. 2005. Not separate, not equal: poverty and inequality in post-apartheid South Africa. William Davidson Institute.

[39] Rothschild, M. and Stiglitz, J. 1976. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *Quarterly Journal of Economics* , 90:630–649.

[40] Roy, A. 1951. Some thoughts on the distribution of earnings. *Oxford Economic Papers* , 3(2):135–146.

[41] Shonkwiler, S. and Yen, S. 1999. Two-step estimation of a censored system of equations. *American Journal of Agricultural Economics* , 81(4):972–982.

[42] Showers, V. and Shotick, J. 1994. The effects of household characteristics on demand for insurance: A tobit analysis. *The Journal of Risk and Insurance*, 61(3):492–502.

[43] Shmueli, A. 2001. The effect of health on acute care supplemental insurance ownership: an empirical analysis. *Health Economics*, 10(4):341–350.

[44] Simkins, C. 2004. *What happened to the distribution of income in South Africa between 1995 and 2001?* University of the Witwatersrand.

[45] Söderlund, N. and Hansl, B. 2000. Health Insurance in South Africa: an empirical analysis of trends in risk-pooling and efficiency following deregulation. *Health policy and Planning*, 15(4):378–385.

[46] Tshiswaka-Kashala, G. and Koch, S.F. 2008. Tobacco's crowding out of expenditure in South Africa. *University of Pretoria Manuscript.*

[47] Van Walbeek, C. 2005. The Economics of Tobacco Control in South Africa, *PhD Thesis*, University of Cape Town, Cape Town, South Africa.

[48] Vermeulen, F. 2003. Do smokers behave differently? A tale of zero exenditures and separability concepts. *Economics Bulletin* 4(6):1-7.

[49] Ying, X., Hu, T., Chen, W., Xu, K. and Huang, J. 2007. Demand for private health insurance in Chinese Urban Areas. *Health Economics* 16(10):1041-1050.

[50] Zweifel, P. and Manning, W. 2000 Moral Hazard and Consumer Incentives in Health Care. In Handbook of Health Economics, Newhouse, J.P. and Culyer, A.J., eds. Elsevier Science B.V., North-Holland: Amsterdam, pp. 409-459.

[51] Zweifel, P. and Breyer, F. 1997. Health Economics. Oxford University Press: Oxford.

† Postdoctoral Fellow, Health and Development Policy Research Group, Department of Economics, University of Pretoria, Pretoria, Republic of South Africa; (O) 27-12-420-4504, (F) 27-12-362-5207.

E-mail address: funkeal@yahoo.com

‡ Director, Health and Development Policy Research Group, Professor of Economics, Department of Economics, University of Pretoria, Pretoria, Republic of South Africa; (O) 27-12-420-5285, (F) 27-12-362-5207.

E-mail address: steve.koch@up.ac.za

Table A4: Coefficient estimates for base regression- UNINSURED

Variables	Housing (a)	Food (a)	Tobacco & alchohol (a)	Non-alcoholic drinks (a)	Education (a)	Miscellaneous (a)	Health Care (a)
Adjusted Expenditure (b)	-0.80 **	-6.52 **	0.54 **	0.26 **	0.27 *	0.75 **	0.17 **
	(0.20)	(0.30)	(0.14)	(0.04)	(0.12)	(0.25)	(0.04)
Adjusted Expenditure Sq. (b)	0.54 **	-0.44 **	-0.25 **	-0.10 **	0.03	0.83 **	-0.05 **
	(0.06)	(0.09)	(0.04)	(0.01)	(-0.03)	(0.08)	(0.01)
Household Size	-0.23 **	0.13 **	-0.10 **	-0.05 **	0.16 **	-0.10 **	-0.02 **
	(0.03)	(0.04)	(0.02)	(0.01)	(0.01)	(0.03)	(0.01)
Households with child(< 5)	-0.57	9.84 **	-1.26	0.00	-5.37 **	-2.63	0.59 *
	(-1.24)	(1.82)	(-0.88)	(-0.26)	(0.69)	(-1.54)	(0.27)
Households with child (5-15)	0.30	5.66 **	-1.74 **	-0.02	-2.81 **	-1.08	-0.23
	(-0.57)	(0.84)	(0.40)	(-0.12)	(0.32)	(-0.71)	(-0.12)
Proportion of working adults	-2.38 **	-5.01 **	3.43 **	0.60 **	-2.92 **	4.08 **	-0.22 **
	(0.30)	(0.44)	(0.21)	(0.06)	(0.16)	(0.37)	(0.06)
R-squared	0.25	0.38	0.06	0.06	0.05	0.09	0.03

Variables	Household Utilities (a)	Communication (a)	Recreation (a)	Transport (a)	Furniture (a)	Clothing (a)	Personal care (a)
Adjusted Expenditure (b)	-0.55 **	1.44 **	0.11 **	1.17 **	1.93 **	1.85 **	-0.63 **
	(0.12)	(0.14)	(0.02)	(0.15)	(0.11)	(0.13)	(0.08)
Adjusted Expenditure Sq. (b)	0.00	0.20 **	-0.03 **	0.19 **	-0.33 **	-0.54 **	-0.05
	(-0.03)	(0.05)	(0.01)	(0.05)	(0.03)**	(0.04)	(-0.03)
Household Size	0.06 **	-0.09 **	0.01	0.03	-0.01	0.18 **	0.05 **
	(0.02)	(0.02)	(-0.00)	(-0.02)	(-0.02)	(0.02)	(0.01)
Households with child(< 5)	-2.07 **	-0.35	-0.18	-3.20 **	0.38	2.21 **	2.61 **
	(0.71)	(-0.89)	(-0.13)	(0.92)	(-0.69)	(0.79)	(0.50)
Households with child (5-15)	-1.57 **	0.11	-0.06	-0.45	1.08 **	0.45	0.35
	(0.33)	(-0.41)	(-0.06)	(-0.42)	(0.32)	(-0.36)	(-0.23)
Proportion of working adults	-1.79 **	0.22	-0.13 **	1.64 **	0.99 **	1.39 **	0.10
	(0.17)	(-0.21)	(0.03)	(0.22)	(0.16)	(0.19)	(-0.12)
R-squared	0.15	0.07	0.06	0.09	0.04	0.06	0.04

Observations: 21885

* significant at 5%; ** significant at 1%. Standard errors in parentheses

a: Regression also includes: provincial dummies, race dummies, urban dummy, headworker dummy and house ownership dummy

b: Adjusted Expenditure is log expenditure net of log minimum expenditure within the local cluster.

Table A1: Average expenditure by categories and insurance status in Rands

Expenditure on:	Description	Insured	Uninsured
Housing	Final consumption expenditure on rent, mortgage excluding expenditure on utilities	1413.16	179.35
		(3765.88)	(917.76)
Food	All types of food excluding non-alchoholic beverages	1029.45	488.77
		(517.76)	(344.43)
Tobacco & alcohol	Cigarette and tobacco and alchoholic drinks	126.49	52.21
		(233.11)	(146.32)
Non-alchoholic drinks	Non-alchoholic drinks and beverages consumed	52.26	20.63
		(68.39)	(47.00)
Education	Every expenditure pertaining to education	246.30	48.90
		(573.63)	(199.88)
Healthcare	Final out of pocket health expenses	72.53	17.12
		(263.81)	(92.24)
Utilities	Consumption of household utilities	261.58	71.52
		(310.50)	(371.20)
Communication	Expenditures on computer and telecommunication equipment, communication services, e.g pay phone, postals	816.69	100.71
		(2268.75)	(682.41)
Recreation	Expenditure that has to do with recreation	97.48	9.19
		(400.95)	(53.11)
Transport	Expenditure that has to do with vehicle purchase and maintainance	880.37	130.10
		(1762.31)	(556.44)
Furniture	Expenditure on furniture, electrical appliances and household equipment	181.54	48.15
		(339.75)	(155.81)
Clothing	All types of clothing and footwear	251.73	90.70
		(258.96)	(141.56)
Personal care	All types of personal care items. E.g. bath soap, hair dressing, beauty care etc	208.96	72.97
		(179.02)	(93.06)
Miscellaneous	Summation of other categories of goods consumed not mentioned above	3162.04	313.61
		(10833.37)	(2680.94)
No. of observations		3492	21885

Standard deviation in parentheses

Table A2: Means and standard deviations of independent variables by insurance status

Independent Variable	Insured	Uninsured
Log of adjusted expenditure	1.87	1.00
	(1.12)	(0.81)
Log of adjusted expenditure Sq.	4.75	1.67
	(5.08)	(2.49)
Household size	3.65	4.06
	(2.12)	(2.76)
Households with child(< 5)	0.01	0.01
	(0.05)	(0.05)
Households with child (5-15)	0.06	0.08
	(0.12)	(0.12)
proportion of working adults	0.48	0.35
	(0.31)	(0.35)
Race-coloured	0.15	0.10
	(0.36)	(0.29)
Race-Asian	0.04	0.02
	(0.20)	(0.12)
Race-White	0.34	0.03
	(0.47)	(0.18)
urban	0.87	0.56
	(0.34)	(0.49)
House ownership	0.37	0.60
	(0.48)	(0.48)
headwker	0.80	0.57
	(0.39)	(0.49)
Provincial HIV prevalence	24.64	25.87
	(8.27)	(7.81)
Human development	0.66	0.64
	(0.05)	(0.05)
Sanitation-flushed toilets	0.99	0.88
	(0.07)	(0.32)
Water source-Piped -in	0.79	0.27
	(0.41)	(0.44)
No. of observations	3492	21885

Standard deviation in parentheses

Table A3. Coefficient estimates for base regression- INSURED

Variables	Housing (a)	Food (a)	Tobacco & alchohol (a)	Non-alcoholic drinks (a)	Education (a)	Miscellaneous (a)	Health Care (a)
Adjusted Expenditure (b)	-0.96	-2.58 **	-0.14	-0.12 *	0.47 *	1.90 *	0.16
	(-0.53)	(0.45)	(-0.16)	(0.05)	(0.23)	(0.76)	(-0.11)
Adjusted Expenditure Sq. (b)	0.24 *	-0.24 *	-0.06	-0.01	-0.07	0.28	-0.03
	(0.11)	(0.10)	(-0.03)	(-0.01)	(-0.05)	(-0.16)	(-0.02)
Household Size	-0.12	0.69 **	0.04	-0.02	0.25 **	-1.13 **	-0.06 *
	(-0.11)	(0.10)	(-0.03)	(-0.01)	(0.05)	(0.17)	(0.02)
Households with child(< 5)	-7.93 *	2.15	-1.66	0.29	-6.21 **	6.13	1.17
	(3.44)	(-2.91)	(-1.05)	(-0.32)	(1.50)	(-4.90)	(-0.71)
Households with child (5-15)	5.67 **	-2.26	-0.73	0.21	-1.75 *	0.81	0.29
	(1.76)	(-1.49)	(-0.54)	(-0.16)	(0.77)	(-2.51)	(-0.36)
Proportion of working adults	-4.01 **	-2.30 **	1.29 **	0.15	-1.50 **	1.20	-0.32
	(0.886)	(0.75)	(0.27)	(-0.08)	(0.38)	(-1.26)	(-0.18)
R-squared	0.24	0.30	0.05	0.08	0.06	0.11	0.03

Variables	Household Utilities (a)	Communication (a)	Recreation (a)	Transport (a)	Furniture (a)	Clothing (a)	Personal care (a)
Adjusted Expenditure (b)	-0.42 **	1.46 **	0.07	0.18	0.41	-0.17	-0.26 *
	(0.16)	(0.48)	(-0.10)	(-0.43)	(-0.21)	(-0.18)	(0.12)
Adjusted Expenditure Sq. (b)	-0.03	0.03	0.00	0.15	-0.10 *	-0.07	-0.07 *
	(-0.03)	(-0.11)	(-0.02)	(-0.09)	(0.05)	(-0.04)	(0.03)
Household Size	0.05	-0.17	-0.02	0.16	0.02	0.22 **	0.10 **
	(-0.03)	(-0.10)	(-0.02)	(-0.09)	(-0.05)	(0.04)	(0.02)
Households with child(< 5)	-3.75 **	3.52	-0.45	-0.83	2.38	1.70	3.48 **
	(1.04)	(-3.08)	(-0.64)	(-2.77)	(-1.38)	(-1.16)	(0.78)
Households with child (5-15)	-1.46 **	-1.77	-0.41	-1.03	1.86 **	-0.12	0.68
	(0.53)	(-1.58)	(-0.33)	(-1.42)	(0.71)	(-0.59)	(-0.40)
Proportion of working adults	-2.11 **	2.47 **	-0.26	2.46 **	0.87 *	1.58 **	0.47 *
	(0.26)	(0.79)	(-0.16)	(0.71)	(0.35)	(0.30)	(0.20)
R-squared	0.15	0.06	0.07	0.04	0.05	0.15	0.11

Observations: 3492

* significant at 5%; ** significant at 1%. Standard errors in parentheses

a: Regression also includes: provincial dummies, race dummies, urban dummy, headworker dummy and house ownership dummy

b: Adjusted Expenditure is log expenditure net of log minimum expenditure within the local cluster.

Table 4: Treatment Effects on the Treated and the Untreated

Treatment Effects	Housing	Food	Tobacco & alchohol	Non-alcoholic drinks	Education	Miscelleous	Health Care
Treatment on the untreated(TEUT)	23.46	-47.11	-6.91	-1.65	3.07	19.02	0.33
	(0.25)	(-0.06)	(-0.01)	(-0.01)	(0.01)	(0.03)	(0.01)
Treatment on the treated(TET)	-16.03	30.79	10.22	1.88	-6.40	-9.51	0.14
	(-0.08)	(0.14)	(0.03)	(0.01)	(-0.02)	(-0.07)	(0.01)

	House Utilities	Communi-cation	Recrea-tion	Transport	Furniture	Clothing	Personal care
Treatment on the untreated(TEUT)	4.57	5.36	0.04	6.47	-0.46	-2.73	-3.60
	(0.01)	(0.01)	(0.00)	(0.01)	(-0.01)	(-0.01)	(-0.01)
Treatment on the treated(TET)	-8.58	-6.78	-0.86	-5.97	3.17	5.53	2.33
	(-0.04)	(-0.03)	(-0.00)	(-0.03)	(0.01)	(0.03)	(0.01)

Bootstrapped (1000 replications) standard errors in parentheses

Table 1: Parameter estimates - Selection of Insurance

VARIABLES	Coefficients	
Expenditure	0.60	**
	(0.04)	
Expenditure Sq.	-0.05	**
	(-0.01)	
Household Size	-0.004	
	(0.01)	
Households with child(< 5)	0.47	*
	(0.21)	
Households with child (5-15)	-0.04	
	(-0.11)	
Race-coloured	0.27	**
	(0.05)	
Race-Asian	0.16	
	(0.08)	
Race-White	1.09	**
	(0.05)	
proportion of working adults	0.15	**
	(0.05)	
urban	0.39	**
	(0.03)	
House ownership	-0.12	**
	(-0.03)	
headwker	0.33	**
	(0.03)	
EXCLUSION VARIABLES		
HIV prevalence	-0.01	**
	(-0.00)	
Human development	-1.07	**
	(-0.35)	
Sanitation-flushed toilets	0.52	**
	(0.08)	
Water source-Piped -in	0.67	**
	(0.03)	
Constant	-2.44	**
	(-0.24)	
Wald chi2(16)	3143.27	
Pseudo R2	0.32	
Log pseudolikelihood	-6866.63	
No of observations	25377	

Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table 2: Coefficient Estimates of Corrected Regression- INSURED

Variables	Housing (a)	Food (a)	Tobacco & alchohol (a)	Non-alcoholic drinks (a)	Education (a)	Miscellaneous (a)	Health Care (a)
Adjusted Expenditure (b)	-5.17 **	3.42 **	0.94 **	0.15 *	-0.02	0.90	0.08
	(-0.86)	(0.59)	(0.23)	(0.07)	(-0.30)	(1.19)	(0.10)
Adjusted Expenditure Sq. (b)	0.61 **	-0.77 **	-0.15 **	-0.03	-0.03	0.37 *	-0.03
	(0.18)	(-0.10)	(-0.03)	(-0.01)	(-0.05)	(0.24)	(-0.02)
Household Size	-0.08	0.62 **	0.03	-0.02	0.26 **	-1.12 **	-0.05 *
	(-0.11)	(0.11)	(0.03)	(-0.01)	(0.04)	(-0.16)	(-0.01)
Households with child(< 5)	-11.13 **	6.71 *	-0.83	0.50	-6.58 **	5.37	1.11
	(-2.79)	(2.80)	(-0.88)	(0.33)	(-1.05)	(4.89)	(0.75)
Households with child (5-15)	6.57 **	-3.54 *	-0.96	0.16	-1.64 *	1.03	0.30
	(1.90)	(-1.39)	(-0.43)	(0.18)	(-0.71)	(2.49)	(0.30)
Proportion of working adults	-4.28 **	-1.91 **	1.36 **	0.17 *	-1.53 **	1.14	-0.32
	(-0.90)	(-0.78)	(0.26)	(0.08)	(-0.39)	(1.24)	(-0.17)
Inverse Mills Ratio	-10.06 **	14.33 **	2.58 **	0.65 **	-1.18 *	-2.40	-0.17
	(-0.94)	(1.00)	(0.45)	(0.11)	(-0.46)	(-1.62)	(-0.10)
R-squared	0.26	0.35	0.06	0.09	0.06	0.11	0.03

Variables	Household Utilities (a)	Communication (a)	Recreation (a)	Transport (a)	Furniture (a)	Clothing (a)	Personal care (a)
Adjusted Expenditure (b)	-1.34 **	0.61	0.08	-0.84	0.71 *	0.26	0.23
	(-0.21)	(0.60)	(0.11)	(-0.60)	(0.28)	(0.23)	(0.16)
Adjusted Expenditure Sq. (b)	0.05	0.10	-0.01	0.24 *	-0.13 **	-0.11 **	-0.11 **
	(0.04)	(0.12)	(-0.03)	(0.12)	(-0.04)	(-0.03)	(-0.02)
Household Size	0.05	-0.17	-0.02	0.17	0.02	0.22 **	0.09 **
	(0.03)	(-0.09)	(-0.01)	(0.10)	(0.04)	(0.04)	(0.03)
Households with child(< 5)	-4.45 **	2.87	-0.44	-1.61	2.62	2.03	3.85 **
	(-0.70)	(2.90)	-(0.46)	(-2.28)	(1.38)	(1.03)	(0.75)
Households with child (5-15)	-1.26 *	-1.58	-0.41	-0.81	1.79 *	-0.21	0.57
	(-0.45)	(-1.50)	(-0.32)	(-1.31)	(0.68)	(-0.50)	(0.37)
Proportion of working adults	-2.17 **	2.42 **	-0.26	2.40 **	0.89 *	1.61 **	0.50 *
	(-0.28)	(0.83)	(-0.18)	(0.71)**	(0.34)	(0.30)	(0.21)
Inverse Mills Ratio	-2.18 **	-2.04 *	0.02	-2.45 **	0.73	1.03 **	1.15 **
	(-0.31)	(-0.89)	(0.20)	(-0.97)	(0.43)	(0.40)	(0.25)
R-squared	0.16	0.06	0.07	0.04	0.06	0.16	0.12

Observations: 3492

* significant at 5%; ** significant at 1%. Bootstrapped (1000 replications) standard errors in parentheses

a: Regression also includes: provincial dummies, race dummies, urban dummy, headworker dummy and house ownership dummy

b: Adjusted Expenditure is log expenditure net of log minimum expenditure within the local cluster.

Table 3: Coefficient Estimates of Corrected Regression- UNINSURED

Variables	Housing (a)	Food (a)	Tobacco & alchohol (a)	Non-alcoholic drinks (a)	Education (a)	Miscellaneous (a)	Health Care (a)
Adjusted Expenditure (b)	-2.06 **	-4.78 **	1.21 **	0.37 **	-0.20	1.03 **	0.15 **
	(-0.30)	(-0.33)	(0.15)	(0.04)	(-0.12)	(0.33)	(0.05)
Adjusted Expenditure Sq. (b)	0.50 **	-0.38 **	-0.23 **	-0.09 **	0.02	0.84 **	-0.05 **
	(0.12)	(-0.10)	(-0.04)	(-0.01)	(0.04)	(0.13)	(-0.01)
Household Size	-0.23 **	0.12 **	-0.10 **	-0.05 **	0.16 **	-0.10 **	-0.02 **
	(-0.03)	(0.04)	(-0.01)	(-0.01)	(0.01)	(-0.03)	(-0.01)
Households with child(< 5)	-1.24	10.76 **	-0.91	0.06	-5.62 **	-2.48	0.58 *
	(-1.23)	(1.80)	(-0.76)	(0.20)	(-0.48)	(-1.34)	(0.30)
Households with child (5-15)	0.63	5.20 **	-1.92 **	-0.05	-2.69 **	-1.15	-0.23
	(0.55)	(0.83)	(-0.32)	(-0.10)	(-0.28)	(-0.63)	(-0.11)
Proportion of working adults	-2.50 **	-4.84 **	3.49 **	0.61 **	-2.96 **	4.11 **	-0.22 **
	(-0.32)	(-0.44)	(0.23)	(0.07)	(-0.19)	(0.39)	(-0.06)
Inverse Mills Ratio	-9.54 **	13.23 **	5.08 **	0.86 **	-3.58 **	2.16 **	-0.15
	(-0.71)	(0.91)	(0.43)	(0.11)	(-0.37)	(0.87)	(-0.13)
R-squared	0.25	0.38	0.06	0.06	0.05	0.09	0.03

Variables	Household Utility (a)	Communication (a)	Recreation (a)	Transport (a)	Furniture (a)	Clothing (a)	Personal care (a)
Adjusted Expenditure (b)	-1.31 **	0.98 **	0.06 *	0.70 **	2.15 **	2.20 **	-0.52 **
	(-0.14)	(0.24)	(0.02)	(0.21)	(0.11)	(0.14)	(-0.08)
Adjusted Expenditure Sq. (b)	-0.02	0.18 **	-0.03 **	0.17 **	-0.32 **	-0.53 **	-0.05
	(-0.05)	(0.09)	(-0.01)	(0.09)	(-0.03)	(-0.04)	(-0.02)
Household Size	0.06 **	-0.09 **	0.01	0.03	-0.01	0.18 **	0.05 **
	(0.02)	(-0.02)	(0.00)	(0.02)	(-0.02)	(0.02)	(0.01)
Households with child(< 5)	-2.47 **	-0.59	-0.21	-3.44 **	0.50	2.39 **	2.67 **
	(-0.63)	(-0.87)	(-0.10)	(-0.78)	(0.67)	(0.81)	(0.52)
Households with child (5-15)	-1.37 **	0.23	-0.05	-0.33	1.02 **	0.36	0.33
	(-0.30)	(0.37)	(-0.06)	(-0.39)	(0.32)	(0.35)	(0.23)
Proportion of working adults	-1.86 **	0.17	-0.13 **	1.60 **	1.01 **	1.43 **	0.11
	(-0.16)	(0.23)	(-0.03)	(0.23)	(0.16)	(0.19)	(0.12)
Inverse Mills Ratio	-5.78 **	-3.53 **	-0.41 **	-3.51 **	1.72 **	2.65 **	0.81 **
	(-0.33)	(-0.56)	(-0.06)	(-0.54)	(0.33)	(0.39)	(0.25)
R-squared	0.16	0.08	0.06	0.10	0.04	0.06	0.05

Observations: 21885

* significant at 5%; ** significant at 1%. Bootstrapped (1000 replications) standard errors in parentheses

a: Regression also includes: provincial dummies, race dummies, urban dummy, headworker dummy and house ownership dummy

b: Adjusted Expenditure is log expenditure net of log minimum expenditure within the local cluster.