Health Shocks and Household Poverty: Evidence from South Africa

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

A Research Report submitted in partial fulfilment of the Degree of
Master of Commerce (Economic Science)
in the School of Economic and Business Sciences,
University of the Witwatersrand

by

Pheeha Morudu
Student No: 495536

Supervised by Professor Uma Kollamparambil

Word Count: 14762

March 2019

"The financial assistance of the National Treasury towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at are those of the author and should not be attributed to the National Treasury"

ABSTRACT

Although South Africa is classified as an upper middle income country, over 50% of its population live under the poverty line. Together with very low level of medical insurance penetration rate, this implies that there are probably over 80 percent of South Africans who are likely to be affected negatively by a sudden health shock. It is therefore important to study the relationship between health shocks and poverty in the South African context. Using food expenditure shock as an indicator for poverty incidence, the study utilises non-linear econometric models to assess the association between households' heads experiencing health shock and food poverty shock. The main finding of the study is that households in the upper quintile are observed to sacrifice food consumption to access healthcare highlighting their vulnerability to health shocks in the absence of medical insurance. In contrast, the results show that lower income households may forgo healthcare, as they potentially cannot trade off their already constrained food consumption for medical expenses. The results indicate that healthcare is a privilege in South Africa, available to the minority of the population. The vast majority of population in SA is vulnerable due to lack of healthcare insurance underlining the need for a National Health Insurance that would enable universal access to health care in the country.

ACKNOWLEDGEMENT

I am grateful to Professor Uma Kollamparambil for positively challenging me during the period of writing this thesis. The coding skills and writing skills that I have learned during the process equipped me and gave me the edge.

I am also thankful for the help I received from my professional mentor Nompumelelo Radebe who constantly made sure that I remembered and prioritised my goals.

I would like to thank my family: My parents Johannes and Fridah and my siblings Maomela and Sontaha, for their prayers and their love throughout this academic journey.

Last but not list, I am thankful for every individual who assisted and supported me during the process of writing this thesis.

1. INTRODUCTION

Health shocks are any injuries and/or illnesses that result in an abrupt deterioration of a person's health (Novignon et al., 2012). These health shocks are often characterised as unforeseeable diseases that result in deterioration of individuals' health status, they are probably one of the most vital components related to poverty (Leive & Xu, 2008). Health shocks result in increased burden of out of pocket health spending and thereafter inducing poverty incidence in low and middle-income countries (Ghosh, 2011; Van Minh et al., 2013). They also decrease labour supply in terms of both hours and labour participation rate (Rocco et al., 2011). Furthermore, they decrease income and reduce non-medical food consumption, food consumption and non-food consumption (Gertler & Gruber, 2002; Lindelow & Wagstaff, 2005; De Weerdt & Dercon, 2006; Khan, 2010). These imply that health shocks can have adverse economic outcomes.

Health shocks mostly increase the vulnerability of the poorer and uninsured households (Novignon et al., 2012). This implies that medical insurance, which covers individuals against health shocks, is important in preventing the extreme economic outcomes that may results in out of pocket spending due to health shocks. The United Nations' 2030 agenda for sustainable development lead to majority of developing country states committing to achieve Goal 3, which intends to ensure that all people have access to quality health services without risks of financial hardships (Keiny et al., 2017). As a result, countries like Togo are in the process of implementing universal health coverage (UHC) even though it will only cover public servants, public servants who are retired and their dependents (Atake, 2018). South Africa is no exception because its government is planning to implement the National Health Insurance (NHI) that will ensure that all South Africans have access to health services. The NHI seeks to ensure that vulnerable and poor populations' welfare is maximised at every step of the implementation process towards the UHC (DoH, 2017).

In 2017, the General Household Survey (GHS) revealed that only 16.9 percent of individuals in South Africa were covered by medical aid (StatsSA, 2017a). According to the GHS, coverage only increased by 1 percent between 2002 and 2017 (StatsSA, 2017a). This implies that over 80 percent of South Africans are likely to be negatively affected by a sudden health shock because they do not have medical aid with a possibility of falling into poverty.

Furthermore, the poverty rate in South Africa seems to be persistently above 50 percent. The proportion of the population living in poverty declined from 66.6 percent (31,6 million individuals) in 2006 to 53.2 percent (27,3 million individuals) in 2011, and then increased to 55.5 percent (30,4 million individuals) in 2015 (StatsSA, 2017b). Persistent existence of poverty remains a concern in recent years in South Africa. For this reason, understanding the mechanisms behind the effect of health shocks on poverty in South Africa is a major focus of this study. This study investigates the relationship between health shocks and households falling into poverty.

Most studies only show the negative effects of the health shocks on income, food expenditure and non-food expenditure (Wagstaff, 2005; De Weerdt & Dercon, 2006; Wagstaff, 2007; Wagstaff and Lindelow, 2010). Therefore, this research goes further than previous studies by investigating the relationship between health shocks and food expenditure shock, which represents a shock that can lead to household poverty. Therefore, this investigation introduces a new measurement variable of poverty shock. The study uses the National Income Dynamics Study (NIDS) panel data to investigate this relationship. The investigation utilized waves three to wave five, which are between the periods 2012 to 2017. The independent variable (health shock) used in this study is the change in body mass index (BMI) and the dependent variable that measures poverty is food expenditure shock. This food poverty shock is a good indicator of poverty shock because food consumption is the general measurement of poverty (StatsSA, 2017b). This food expenditure shock is a binary variable indicating whether the household experienced a significant decrease in growth rate of logged real total food expenditure or not. This implies that this study employs non-linear models to solve the research problem.

The non-linear models of estimation used are panel data logit model, non-linear Difference in Differences (DID) model and Propensity Score Matching (PSM) with DID model. The main model of estimation is the DID model whereby the control group will consist of households that did not experience a health shock and the treatment group consists of household which experienced a health shock. The panel logit model uses the first dataset that allowed the presence of health shocks from period one to period three. The DID estimations used the second dataset (quasi-experimental dataset) which only allowed the health shocks in period three. The utilization of the DID model allows inference of 'Post treatment period' x 'Treatment' effect which holds both Treatment and Post treatment period constant (Puhani, 2012). The PSM is a method to reduce the bias in the estimation of treatment effects with observational data (Rosenbaum & Rubin, 1983). This implies that the PSM with DID method employed will reduce the possible selection bias that may arise in original DID method.

However, there still exist a problem of reverse causality (endogeneity). Instrumental Variable (IV) approach to explicitly address the endogeneity of illness or health shock (Gertler & Gruber, 2002; De Weerdt & Dercon, 2006; Genoni, 2012). Literature suggested the prices of health inputs, individual reports about physical functioning abilities, medical expenditure and reduced labour hours as instruments of health shocks to use in performing instrumental variables (Gertler & Gruber, 2002; De Weerdt & Dercon, 2006; Genoni, 2012), respectively. However, due to NIDS dataset limitations, the instrumental Variable approach was not employed in this study. Therefore, the study can only claim association rather than causal effect.

Nevertheless, using change in BMI as a health shock, the analysis showed that higher income households (quintile 2) respond to the health shocks by decreasing their food expenditure while the lower income households (quintile 1) models' estimations were not significant. Households living on the margin do not have room to bring down food expenditure any further (Islam & Maitra, 2012). This indicates that the healthcare is not accessible to a large number of households. Therefore, the individuals that access health care services have to do so by reducing food consumption. Other covariates e.g. higher number of household sizes (4-6; 7-9 and 10-24) are associated with an increased probability of a significant food expenditure decrease versus lower household size of 1-3 (Atake, 2018). On the other hand, receiving a pension grant and having a spouse are associated with a decrease in probability of a household experiencing a food expenditure shock.

The structure of the research will be as follows: the health and poverty theoretical framework is covered in section 2. The measurement of health shocks and poverty shocks is included in section 3. The econometric strategy is presented in section 4 while the data, including descriptive statistics analysis, is presented in section 5. The analysis of the results and conclusion are presented in section 6 and 7, respectively.

2. LITERATURE REVIEW

The existing literature shows that there are two distinct ways through which households can be impacted by health shocks. Firstly, illness or death can affect households through a loss of productive labour time and earnings (Alam & Mahal, 2014). Secondly, incurring out of pocket health expenditures when seeking treatment due to a health shock, this can result in households' spending above a prescribed capacity to pay (Alam & Mahal, 2014).

Death of an adult, changes in self-reported health, an indicator of activities of daily living and specific disease indicators are the four main indicators of health shocks in literature. The first method considers the effect of health shocks on household labour supply and income using the former stated health shocks indicators. Considering the death of an adult as an indicator: Khan (2010) finds that in Bangladesh, work participation of household members lowers (by an average of 8.6 hours in the last two weeks) due to the death of a household member in the prior two years. This shows that health shocks can affect labour supply. In addition, one can investigate the direct impact of adult mortality on household income or earnings. A study based on rural Zambia concluded that the effect of the death of an adult on household income per capita is not statistically significant (Mahmoud & Thiele, 2013). Khan (2010) reached a similar conclusion when he found no statistically significant effect of the death of an adult on household income. In contrast, results from investigations for Vietnam and Kenya show that there is a significant effect of adult mortality on income and the studies suggest an income reduction between 25 percent and 40 percent respectively (Wagstaff, 2007; Yamano & Jayne, 2004).

In addition, the impact of other health shock indicators, namely changes in the level of disability, self-assessed health, body mass index, on household earnings and labour supply are study specific. Therefore, it is difficult to generalize a conclusion. For example, Bales (2013) states that annual workdays decreased by 8 percent in households where an adult was bedridden for two or more weeks in Vietnam. The following health indicator is the specific disease indicator, for instance, HIV. Both Yamano and Jayne (2004) and Beegle (2005) analyses based on Kenya and Tanzania respectively resulted in showing a significant reduction in labour force participation and earnings on account of the death of an adult due to AIDS.

Similarly, with non-communicable diseases, in general, chronic non-communicable diseases are associated with decreased household's income by 4.8 percent (Abegunde & Stanciole, 2008). Furthermore, Murphy et al. (2013) and Rocco et al. (2011) show that lower rates of work participation are likely associated with households consisting of a member with a chronic condition.

A number of studies also investigate the effect of health shocks on out of pocket health expenditures. Firstly, considering death of an adult as a health shock, Wagstaff (2007) finds that the death of working age member 2 years before the survey leads to a non-significant 27 percent increase in households' per capita medical expenditure in Vietnam. However, Khan's (2010) analysis from Bangladesh leads to the conclusion that there is a significant 54 percent decrease in medical expenditure due to the death of a household member for the previous 2 years. The analysis also finds a significant 62 percent increase in medical expenses for serious illness of any household member in the previous one year. Therefore, there is a contradiction from a general hypothesis that death of an adult causes an increase in medical expenditure, and as we have seen from Khan (2010) which shows that only health expenditure due to illness resulted in an increase of medical expenditure.

In addition, there is literature, which also focuses on the specific disease indicator that is HIV. Using the case of Nigerian households, Mahal et al. (2008) finds that out of pocket health expenses of households with HIV patients are significantly higher than the per capita out of pocket health expenditures of similar households without HIV patient as a member. Abegunde and Stanciole (2008) shows that households containing members with chronic diseases incur higher levels of out of pocket health expenditure than those without members with chronic diseases. The population of South Africa suffers from high levels of premature death and burden of illness (World Bank, 2012). Spending on preventative programmes is one of the activities that are intentionally regarded an as important investment in human capital and a key component of a poverty reduction programme (World Bank, 2012).

The above shows that intensive investigations of the impact of health shocks on household's income shocks have been done internationally. However, there is limited literature that focuses on South Africa despite South Africa's exposure to poverty that may be associated with some form of health shocks. The World Bank (2012) finds that 23 percent of the South African population use private sector health services on a regular basis and 17 percent of the population are members of a medical scheme. This implies that those not covered by a medical scheme face high out of pocket expenses for many health services. South Africans' spending on medication and private hospitals over the past 10 years has increased (World Bank, 2012). Furthermore, StatsSA (2017a) recently reports that the percentage of individuals covered by a medical aid scheme increased marginally by 1 percent (15.9 percent to 16.9 percent) between 2002 and 2017.

A medical poverty trap refers to a phenomenon whereby households are driven into poverty and there is a further increase in poverty for those who already fall below poverty line due to rises in out of pocket costs for public and private health services (Whitehead et al., 2001; Ramaraj & Alpert, 2008). This implies that health trap and poverty trap can occur interchangeably. Banerjee et al. (2011) study on Indonesia reports on the case of a household that experienced a health shock in the form of blindness of the head of the household. The household had to borrow money to cover medical expenses which resulted in the household incurring more debt. However, the household remained poor and when a child became sick, the household could not afford medical treatment of the child. Therefore, the child stayed sick and hence the child had higher risk of being poor in the future because he could not go to school due to poor health.

Novignon et al. (2012) and Somi et al. (2009) agree that impacts on households' welfare can be in both the short and long run by health shocks and their associated costs. For example, a household can substitute the consumption and production expenditure for healthcare when it faces health shocks in the short run (Somi et al., 2009). Therefore, there exists a possibility that a household can fall into poverty or become poorer due to a health shock in the process of accessing a healthcare especially when it is uninsured. This leads to the concept of vulnerability. Hoddinott and Quisumbing (2010) define vulnerability as the likelihood that an individual will have a level of welfare below some norm or benchmark at a given time in the future. The following section elaborates on health shocks, vulnerability concept and poverty shocks including the focus of the study and additions to existing literature.

3. MEASUREMENT OF HEALTH SHOCKS AND POVERTY SHOCKS

Existing literature has used different combinations of variables as measures of health shocks. For example, death of a working age adult household member due to illness (Yamano & Jayne, 2004; Wagstaff, 2007; Khan, 2010; Wagstaff & Lindelow, 2010; Mahmoud & Thiele, 2013). In addition, some studies used self-reported serious illness that prevented a household member from being able to work (Lindow & Wagstaff, 2005; Bridges & Lawson, 2008; Khan, 2010; Rocco et al., 2011). Other studies used changes in limitations of ability to perform activities of daily living (Gertler & Gruber, 2002; Genoni, 2012).

However, there are limitations to the NIDS data regarding the above suggested health shocks variables. Serious illness as a health shock is defined as any illness that prevented a household member from doing normal activities, serious illness can include any health problem, for example, disability, disease, injury or any other chronic disease (Abegunde & Stanciole, 2008; Khan, 2010 & Rocco et al., 2011). The NIDS questionnaire changed and the household questionnaire discontinued to capture the question that asks whether there was "serious illness or injury of a household member" from wave 4. Therefore, the household dataset used no longer captures the information indicating if any member (household head, child or any other member) has had a serious illness. Due to this data limitation, this ruled out the possibility to assess the impact of health shocks using "serious illness or injury of a household member" as a health shock indicator.

Nevertheless, this research investigates how health shocks push households in margins to poverty. This study focuses on how a significant decrease in body mass index (BMI) as a health shock affects households in margins. BMI is equal to the individual's weight in kilograms divided by the square of the individual's height in meters. James et al. (1988) states that BMI is a recognized and reliable measure of the current nutritional status of adults. According to

Waaler (1984), extremely low BMI below 17 and 18 have been associated with high mortality even though there is still substantial debate on the cut-off points for identifying the degree of malnutrition. Nevertheless, Dercon and Krishnan (2000) states that the link between illness and the BMI is less clear. Alternatively, Gertler and Gruber (2002) use changes in ability to perform activities of daily living to capture severe illness events and they find that labour supply and earnings are significantly compromised by illness. However, the NIDS data does not capture information on ability to perform activities of daily living.

Wagstaff (2005) investigated a decrease in log average (BMI) among household members aged 18 and above between 1993 and 1998. The author then adjusted the former study and used just a BMI as a dummy variable, which takes the value of 1 if BMI is less than 18.5 in those two waves in the study or a drop in BMI between waves exceeding one standard deviation of the BMI-change distribution (Wagstaff, 2007). This research focuses only on the decline in the change of BMI that is more than one standard deviation as an indicator of health shock because this indicator is suitable and can be calculated using NIDS data on weight and height information. The persistent BMI below underweight does not fit the criteria of a health shock because it implies that the person was ill to start with.

Similarly, poverty shock can be measured using different measurements. For instance, one way can use vulnerability as a measure of poverty. Hoddinott and Quisumbing (2010) discusses three methods which are used to estimate vulnerability, that is, Vulnerability as Low Expected Utility (VEU), Vulnerability as Expected Poverty (VEP) and Vulnerability as Uninsured Exposure to Risk (VER). Ligon and Schechter (2003) define VEU with reference to the difference between the utility derived from a certain level of certainty-equivalent consumption at and above which the household would not be considered vulnerable. Chaudhuri et al. (2002) define VEP as the probability that expected consumption expenditure of a household will fall into poverty in the future. Hoddinott and Quisumbing (2010) and Ligon and Schechter (2003) define VER as a method that assesses welfare loss in the absence of effective risk management tools. The VER approach allows an ex-post evaluation of the scope of the negative shock causing a loss of well-being. This allows the effective use panel data study. According to Hoddinott and Quisumbing (2010) illustrated the general VER model as follows:

$$\Delta \ln C_{htv} = \sum_i \lambda_i S(i)_{tv} + \sum_i \beta_i S(i)_{htv} + \sum_{tv} \delta_v (D_v) + \delta X_{hv} + \Delta\varepsilon_{hvt}$$

which denotes that a change in log consumption (the growth rate in total consumption) per capita of household h, in period t is a function of covariate shocks ($S(i)_{tv}$), idiosyncratic shocks ($S(i)_{htv}$), set of binary variables identifying each community separately (D_v) and a vector of household or household head’s characteristics (X). The above general method can be altered. For example, Tesliuc and Lindert (2002) have a single cross sectional survey, which resulted in estimation of the following model:

$$\ln C_{htv} = \alpha + \sum_i \lambda_i S(i)_{tv} + \sum_i \beta_i S(i)_{htv} + \sum_{tv} \delta_v (D_v) + \delta X_{hv} + \varepsilon_{hvt}$$

The above equation is a situation where the coefficients λ and β are both equal to zero which implies that the investigated household is fully insured against covariate and idiosyncratic shocks respectively. Higher estimated values of coefficients are interpreted as signifying a higher covariance between the independent variable and dependent variable changes and thus a higher vulnerability of dependent variable to independent variable (risk). The VER method is an alternative method to measure vulnerability of households to risks and it is different from the VEP method used by Atake (2018) in measuring the extent of vulnerability of poor and uninsured households.

Furthermore, literature shows that others used logarithmic changes in non-medical consumption (C), income (Y) and medical consumption (M) to analyse the impact of health shocks on these three variables of interest (Gertler & Gruber, 2002; Wagstaff, 2005). Khan (2010) also used logged values of the former variables of interest using fixed effects regression model to estimate the impact of health shocks. The variable of interests are continuous variable, thus ordinary fixed effects was appropriate. Food consumption per capita is considered to be a preferred measure of absolute poverty for developing countries and income per capita is second option (Ravallion & Chen, 2007). Statistics South Africa also measures poverty using the food poverty per person. Therefore, this research focuses on the adverse decline in food expenditure per person as a food expenditure shock, which represents a poverty shock. This implies that models that are suitable for continuous variables will not be suitable. Thus, the following section elaborates on the models of estimations that differ from the work of Hoddinott and Quisumbing (2010) and Wagstaff (2005).

4. ECONOMETRIC STRATEGY

The study uses non-linear models in the analysis because the main dependent variable (food expenditure shock) is a categorical variable. Table A.1 in the Appendix has detailed descriptions of the variables used in the following models. The methodological strategy uses the panel structure of the data that allows the time-invariant unobservables and individual specific effects to be controlled for. The conditional fixed effects logit or traditional random effects logit model takes into account the unobserved heterogeneity at household level (Bartels, 2008; Dhanaraj, 2014). Unlike a situation whereby the dependent variable is continuous, the food expenditure shock is a binary variable.

4.1. Model I

The following method will use the first dataset, which allows the health shock to vary throughout the three investigation periods. The panel logit model using latent variable framework is as follows:

$$y_{it} = \beta_0 + \beta_1(health\ shock) + \beta_3 X_{it} + \varepsilon_{it}$$

Where:

- Individual and time identifiers are subscripted as i and t respectively;

- y_{it} is a food expenditure shock dummy variable taking value 1 if there was a significant decrease in food expenditure and 0 otherwise. This is the main dependent variable;

- $health\ shock$ is an identifier of whether the head of the household experienced a significant decrease in BMI change which takes the value of 1, and 0 otherwise. This is the main independent variable;

- X_{it} is the vector of control variables; and
- ε_{it} is a non-zero residual variable containing both conditional error and time invariant unobservables.

The upcoming results from the panel logit regressions are extensions from above stated equation.

4.2. Model II

The above panel logistic model estimator does not take into account the time effect and treatment effect because the health shock is allowed to vary throughout three periods. Therefore, Method II uses the second dataset (quasi-experimental dataset) because the health shock (treatment) in this data was allowed to exist only in the post treatment period.

Yamano and Jayne (2004) uses fixed effects to account for endogeneity in their analysis of the impact of working-age adult mortality on households in Kenya. The Difference in Difference (DID) estimation is identical to household fixed effects analysis used by some authors (Ainsworth and Semali (2000); Deininger et al. (2003); Yamono & Jayne (2004)) because unobservable time-invariant household characteristics are differenced out. Wagstaff (2007) used logged income, consumption and non-food consumption (medical spending) as a dependent variable in his analysis and used Fixed Effects model. However, this investigation focus is on the movement of households to poverty, which led to the creation of a binary variable food expenditure shock.

The DID model can be used to estimate the effect that the health shock has on the treated because the households belonging to treatment group and control group are observed in both the pre-treatment (wave 3 and wave 4) and post treatment period (wave 5). Post treatment variable (Post) is a dummy variable that is equal to one if the observable is from the post treatment period and zero if the observable is from the pre-treatment period. Karaca-Mandic et al. (2012) suggested that a non-linear DID model should have the conditional probability that the dependent variable equal to 1 be expressed as a function of the usual DID function with a continuous dependent variable. Therefore, the non-linear DID model is as follows:

$$P(y = 1|x) = F(\beta_0 + \beta_1 Post + \beta_2 Treat + \beta_{12}(Treat * Post) + X\beta)$$

Where:

- $P(y = 1|X, Treat = 0, Post = 0) = F(\beta_0 + X\beta)$;
- $P(y = 1|X, Treat = 0, Post = 1) = F(\beta_0 + \beta_1 + X\beta)$;
- $P(y = 1|X, Treat = 1, Post = 0) = F(\beta_0 + \beta_2 + X\beta)$; and

- $P(y = 1|X, Treat = 1, Post = 0) = F(\beta_0 + \beta_1 + \beta_2 + \beta_{12} + X\beta)$.

The parameter β_1 allows the $P(y = 1|x)$ (conditional probability that there exists a food expenditure shock) to vary for all households in the post treatment period (wave 5) compared to the pre-treatment periods (wave 3 and wave 4). β_2 allows the $P(y = 1|x)$ to vary for treatment group compared to control group. β_{12} allows the $P(y = 1|x)$ to vary in the post treatment period, and therefore, the conditional probability that there is a food expenditure shock to vary over and above the difference attributable to the non-linearity of the model for households in the treatment group versus those in control group.

4.3.Model III

Endogeneity exists when the explanatory variable is correlated with the error term, that is, it becomes unclear to what extent the results are driven by reverse causality or omitted variables correlated with health and income changes. For example, Mahmoud and Theile (2013) finds that AIDS is a behavioural disease, because sexual transmissions account for the numerous number of HIV infections in Africa, therefore, the likelihood of endogenous selection into the treatment (prime-age mortality) increases.

Kadiyala et al. (2009) and Kadiyala et al. (2011) use propensity score matching (PSM) with a DID estimator to estimate the impact of prime age adult mortality (PAM) on child survival and growth and the effect of PAM on household composition in rural Ethiopia using the ERHS dataset, respectively. Similarly, Mahmoud and Thiele (2013) use the DID-PSM estimator to solve the problem of endogenous selection. The above DID estimator (Model II) controls for time invariant unobserved characteristics while PSM removes the selection bias due to observed differences between treated and control households.

The ideal experiment compares the outcomes of two identical individuals or households exposed to the same treatment (health shock) and the PSM approximates this kind of an experiment. However, there exists a dimensionality problem, that is; it is not clear how each characteristic should be weighted when matching characteristics between treatment group and control group. Rosenbaum and Rubin (1983) suggest the use of PSM to solve problem of dimensionality.

Model III estimates PSM with DID. However, unlike the original Ashenfelter and Card (1984) model, this method explores the logit Propensity Score Matching with DID that Kadiyala et al. (2011) utilized in their study to address endogeneity of PAM (whereby PAM takes the value 1 if household has an adult aged between 15 and 54 die between 1994 and 1997, 0 otherwise) to investigate the effect of PAM on household's welfare using dependency ratios.

The above Method II (DID) estimator approach may have selection bias in the model. The assignment mechanism (treatment and control) is not random. The basic idea is that the bias is reduced when comparison of outcomes is performed using treated and control subjects who are similar as possible (Rosenbaum and Rubin, 1983).This Propensity Score analysis will ensure that the DID estimator above is validated by analysing performing the sensitivity analysis.

However, it is worth noting that PSM reduces but does not eliminate the bias generated by unobservable confounding factors.

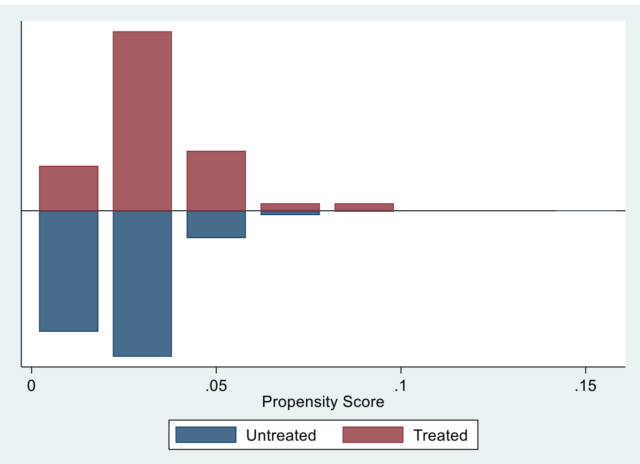
The PSM technique involves a construction of artificial control group by identifying for every treated observation an untreated observation that has the most similar (observable) characteristics. This matching estimates' technique that PSM model results in creating the average treatment effect of health shocks on food expenditure shock. Rosenbaum and Rubin (1983) states that the propensity scores which result in a balanced score whereby the spread of the observed covariates are independent of the binary treatment. The implication of this is that the spread of there exists the same spread of the covariates for the treated (experienced health shock) and control (no health shock) groups, thus implying more robust treatment effect.

A stepwise logistic regression to select variables was estimated as proposed by Rosenbaum and Rubin (1984). This resulted in significantly related variables in the final model. However, the disadvantage of using PSM model simulation is that it can be sensitive to different model specifications (Smith and Todd, 2005). Thus, the study conducted by Smith and Todd (2005) raised some uncertainties on the reliability of the non-parametric evaluation' results. Reichardt et al. (1980) proposes limiting analyses to variables of theoretical importance to treatment selection and those previously demonstrated to predict outcome (strategies to specify a correct model predicting propensity score). The matching algorithm employed in this study use the Nearest Neighbour matching which also known as the traditional pairwise matching. However, one of the disadvantages of this algorithm is that the matching faces the risk of bad matches if the closest neighbour is far away.

4.3.1. Common Support

The common support or overlap condition can be tested by visual inspection of the densities of propensity scores of treat and control groups. An adequate number of observations requires to be present post matching and the exclusion of those observations that are outside the common support. If the assumption of common support is not met, the implications is that there are not enough matched samples in the data. Thus, the treated group and control group are unmatched.

Figure 1: Whole Sample's common support



The final propensity score matching drops the observation that are off support that is the observation will only be in the unmatched sample.

4.3.2. Analysis of Sensitivity

The robustness of the unobservable of the matched estimates to selection bias is explored using the sensitivity analysis in (Becker & Caliendo, 2007). The conditional independence assumption (CIA) assumes that the spread of the propensity scores is identical for those receiving treatment and not receiving treatment due to all factors that generate dependence are truly the same. However, the CIA is always assumed in most cases because it cannot be tested in practice since it involves unobserved potential outcomes. However, the sensitivity analysis can be carried out using the suggested MH non-parametric test contrasts the number of individuals, which were successful in the treatment group with the same expected number with the initial condition that the treatment effect is zero (Becker & Caliendo, 2007). That it tests the hidden bias, whereby gamma (γ) equals one implies that the study is free of hidden selection bias. Similarly, higher values of γ greater than one implies that the odds of receiving treatment because of pre-treatment differences increase.

Therefore, Model III uses the psmatch2 procedure in STATA to obtain propensity scores (Leuven & Sianesi, 2018) which replaces the panel weights used in Model II to solve selection bias that may exist in Model II.

4.4. Limitations

An alternative method is Instrumental Variable (IV) approach that explicitly addresses the reverse causality of illness as employed by Genoni (2012). The author extends Gertler and Gruber (2002) analysis by instrumenting illness with changes in the prices of health inputs. However, Genoni (2012) maintains Gertler and Gruber's use of individual reports about physical functioning abilities as an illness measure; but in addition, Genoni discusses some of the limitations of these variables that should be taken into account when interpreting his results. De Weerdt and Dercon (2006) used medical expenditure and reduced labour hours as measures of health shock to perform their Instrumental Variable regression. However, the NIDS data has a limited medical expenditure information, which makes the undertaking of regressions that involves medical expenditure challenging. The following section 5 outlines in details the data and variables used.

5. DATA AND DESCRIPTIVE STATISTICS

5.1. NIDS Data

Data utilized in the investigation is extracted from the National Income Dynamics Study (NIDS) of South Africa. Supervised by the Southern Africa Labour Development Research Unit (SALDRU), NIDS data is a national representative panel study in South Africa and currently has five waves with approximately 2-year intervals. The panel dataset has approximately 7 300 households and 28 000 individuals captured in the database. NIDS is a panel of individuals, which can be connected to households using their person-identifier (pid). Chinhema et. al. (2016:6) indicated that a sum of 29 733 main sample individuals were followed from 2008 to 2015, taking into account investigation of key changes in groups' encounters of, inequality, employment, wealth, health and the effects of social spending such as education. Thus, the database contains information necessary to understand the changing face of poverty in South Africa.

To avoid misrepresentation of specific characteristics of the population, the NIDS dataset contains panel weights to account for systematic attrition and non-response. This legitimizes SALDRU's execution of a probit model for use as a panel weight to evaluate the probability of an individual staying present through the survey periods. To avoid the potential for wrong outcomes, SALDRU cut qualities at the first and 99th percentiles of the weight dissemination (Chinhema et. al., 2016). To ensure that the outcomes are nationally representative, the panel weights and post-stratified weights are employed throughout the analysis.

The study uses two distinct datasets and a number of control variables and elaboration is given in the following subsections.

5.1.1. First dataset

The first dataset allowed both health shock and food expenditure shock in all three waves. Table 1 shows that households experienced lower percentages of health shocks which seem to be lower than 4 percent in all the three waves. This table shows that 1.96 percent, 3.50 percent and 2.73 percent of households experienced a health shock between wave 2 and wave 3, wave 3 and wave 4, and wave 4 and wave 5, respectively. Similarly, 6.89 percent, 9.67 percent and 10.7 percent of the households experienced food expenditure shock between wave 2 and wave 3, wave 3 and wave 4, and wave 4 and wave 5, respectively

Table 1: Percentages of the presence of Health Shocks and Food Expenditure shocks

Variables	Wave 3	Wave 4	Wave 5
Health shock	1,96	3,50	2,73
Food Expenditure Shock	6,89	9,67	10,70
Number of Households	1 684	1 686	1 683

5.1.2. Second dataset (Quasi-experimental dataset)

The second dataset restricted the sample such that health shock is observed only between 4th wave and 5th wave. This second dataset is the primary dataset of the study, which allowed the division between treatment group and control group. The treatment group consists of those households whose heads experienced a decline in change in BMI that was more than 2.9018 in wave 5. While Wagstaff (2007) used a BMI as a dummy variable which take the number 1 if less than 18.5 in a consecutive wave or a drop in BMI between waves exceeding one standard deviation of the BMI-change distribution (0.088), and 0 otherwise. Unlike the first dataset, those households, which experienced health shock in first period and second period were dropped when constructing the second dataset. Primarily a measure of nutritional status, BMI has been found to be a good predictor of mortality, with mortality risk higher at both the bottom and top ends of the scale (Calle et al., 1999). The study only focuses on the changes that leads to the bottom of the scale, which in turn signify adverse weight loss.

5.1.3. Variables Discussion

This study uses three waves (Wave 3 to Wave 5) in the regressions analysis, the health shock dummy which is the changes in BMI was created and takes the value of 1 if a decrease in BMI is more than the standard deviation and the individual's BMI in previous wave is less than 30 which is the threshold for obese. For the simplification of analysis, the investigation only considered the household heads as representatives of the households. This method of considering only the heads of household health shock is similar to the one used by Asfaw & Braun (2004) which only considered heads of household illness. However, the consideration of heads of household only in the estimations is restrictive.

Control variables are household characteristics, which includes location, household size and characteristics of household's head, which consists of age, gender, education, medical aid coverage and employment status. The dependent variable used is the food expenditure shock that is a dummy variable that takes the number 1 for growth in food expenditure per capita that decreased more than the standard deviation and the number 0, otherwise. Several variables' description is on Table A.1 of the Appendix.

5.2. Descriptive Statistics

Table 2 contains the average weighted means and standard errors for all variables used in the primary sample (quasi experimental dataset). This table illustrates observable differences in variables between treatment and control groups for the whole three periods. The following has been divided between those who experienced health shock as referred to as "Treated" because they are part of the treatment group. Those who did not experience a health shock in the third period are under the "Control" group. The dataset contains 4 894 observations, and the households data shows a high households' average income as R4 019.62. The below descriptive statistics show the households' characteristics.

Table 2: Summary Statistics of Households' Characteristics of DID sample (Average Weighted)

Variable	Sample	N	Mean	Std. Error
	FULL SUB-SAMPLE			
Food Expenditure Shock	Whole	4 771	0,0793	0,2702
HH Income	Treated	123	4019,62	4383,3630
	Control	4 689	3927,71	8016,4520
HH Medical Expenditure	Treated	123	0,98	9,9149
	Control	4 689	46,86	536,9118
HH Food Expenditure	Treated	123	608,88	651,2580
	Control	4 689	521,69	605,8528
Logged Income	Treated	123	7,6532	1,2140
	Control	4 689	7,4360	1,2295
Logged Food Expenditure	Treated	122	6,0272	0,8918

	Control	4 669	5,8348	0,9141
Household Size	Whole	4 812	2,9867	2,4574
1-3	Treated	123	0,7959	0,4047
	Control	4 689	0,6659	0,4717
4-6	Treated	123	0,1383	0,3466
	Control	4 689	0,2418	0,4282
7-9	Treated	123	0,0504	0,2196
	Control	4 689	0,0714	0,2576
10-24	Treated	123	0,0154	0,1238
	Control	4 689	0,0208	0,1429
Urban	Treated	123	0,5905	0,4938
	Control	4 689	0,6897	0,4627
Piped Water	Treated	123	0,8855	0,3198
	Control	4 689	0,9241	0,2648
toilet	Treated	123	0,6011	0,4917
	Control	4 689	0,6638	0,4724

On average, 7.93 percent of the households experienced food expenditure shock and 2.56 percent of households experienced a health shock in wave 5 (this is because there are 41 treatment observations per period). On average, a typical household in South Africa consists of three individuals, this is supported by observing that 80 percent and 67 percent of the treatment group and control group have household sized between 1 and 3, respectively. StatsSA (2016) reported that in 2016 the average household size in South Africa was 3.3. This corresponds with the above whole sample' average household size of 3 individuals. Household's income per capita and food expenditure per capita have large deviation and applying logs on the values result in decreasing the variation.

Furthermore, comparative differences between treatment and control groups are observed for individual level (household's head) variables shown in Table 3 below.

Table 3: Summary Statistics of Household and Locality Characteristics (Average Weighted)

Variable	Sample	N	Mean	Std. Error
	FULL SUB-SAMPLE			
Medical Aid Coverage	Treated	123	0,0764	0,2668
	Control	4 689	0,1598	0,3664
Female	Treated	123	0,2856	0,4535
	Control	4 689	0,5239	0,4995
Married	Treated	123	0,3794	0,4872
	Control	4 689	0,1393	0,3463
Age	Whole	4 812	46,3230	13,4817
17-29	Treated	123	0,0017	0,0418
	Control	4 689	0,0992	0,2990
30-49	Treated	123	0,4658	0,5009
	Control	4 689	0,5268	0,4993

50-64	Treated	123	0,3484	0,4784
	Control	4 689	0,2706	0,4443
65-91	Treated	123	0,1841	0,3891
	Control	4 689	0,1033	0,3044
Education attainment	Whole	4 811	10,3557	5,8050
None	Treated	123	0,0848	0,2797
	Control	4 689	0,0765	0,2659
Primary	Treated	123	0,3655	0,4835
	Control	4 689	0,2301	0,4209
Secondary	Treated	123	0,5024	0,5020
	Control	4 689	0,4412	0,4966
Certificate	Treated	123	0,0473	0,2131
	Control	4 689	0,1325	0,3391
Undergraduate	Treated	123	0,0000	0,0000
	Control	4 689	0,0905	0,2869
Postgraduate	Treated	123	0,0000	0,0000
	Control	4 689	0,0290	0,1678
Pension recipient	Treated	123	0,2257	0,4197
	Control	4 689	0,1367	0,3436
Employed	Treated	123	0,4238	0,4962
	Control	4 689	0,5560	0,4969

From above, on average only 8 percent of those belonging to the treatment group are covered by a medical aid while 16 percent of those in the control group are covered by a medical aid. The control group average corresponds with StatsSA (2017a) which stated that the percentage of individuals covered by medical aid increased marginally from 15.9 percent to 16.9 percent when evaluating the General Household Survey between 2002 and 2017. Most of the households' heads on average are aged between 30 and 49 years, with a secondary education as highest education attainment. Over 40 percent of the sample heads of households seem to be unemployed. However, according to StatsSA (2018), the unemployment rate was reported to be on the rise because it increased from 21.5 percent to approximately 28 percent from 2008 to 2018. The difference in the unemployment rates is likely due to the fact that this sample contains households' heads who are generally older and therefore, less likely to be employed. In addition, many household heads receive old age pensions.

The following Table 4 combines the household and individual characteristics per quintile. The whole sample is divided into two quintiles[1] where by Quintile 1 consists of households with logged income per capita between R3.93 and R7.34 while Quintile 2 consists of households with logged income per capita between R7.35 and R11.94. The logged income per capita of defining the quintiles is preferred because the logged income appears to have less variation. The below results show the differences between lower income households (Quintile 1) and higher income households (Quintile 2).

[1] Attempts to divide this data sample into more quintiles (5 and 10) failed due data constraints.

Table 4: Summary Statistics of Household and Individual Characteristics (Average Weighted, Quintile)

Variable	Sample	N	Mean	Std. Error	Min	Max
	FULL SUB-SAMPLE					
Total Medical Expenditure	Quintile 1	3 094	23,34	107,54	0,00	3067,49
	Quintile 2	1 718	101,38	775,50	0,00	10224,95
HH Food Expenditure	Quintile 1	3 094	251,18	235,85	0,00	4089,98
	Quintile 2	1 718	796,38	730,09	0,00	8179,96
HH Income	Quintile 1	3 094	747,19	387,86	51,12	1547,38
	Quintile 2	1 718	7116,90	10307,57	1549,03	152777,80
Logged Income	Quintile 1	3 094	6,4484	0,6276	3,93	7,34
	Quintile 2	1 718	8,4344	0,8107	7,35	11,94
Urban	Quintile 1	3 094	0,5740	0,4946	0	1
	Quintile 2	1 718	0,8012	0,3992	0	1
Household Size	Quintile 1	3 094	4,1252	2,6993	1	24
	Quintile 2	1 718	1,8466	1,4807	1	22
1-3	Quintile 1	3 094	0,4715	0,4993	0	1
	Quintile 2	1 718	0,8662	0,3405	0	1
4-6	Quintile 1	3 094	0,3584	0,4796	0	1
	Quintile 2	1 718	0,1206	0,3258	0	1
7-9	Quintile 1	3 094	0,1311	0,3376	0	1
	Quintile 2	1 718	0,0108	0,1034	0	1
10-24	Quintile 1	3 094	0,0390	0,1937	0	1
	Quintile 2	1 718	0,0024	0,0488	0	1
Age	Quintile 1	3 094	48,3844	13,9490	17	91
	Quintile 2	1 718	44,2585	12,6682	17	90
17-29	Quintile 1	3 094	0,0840	0,2774	0	1
	Quintile 2	1 718	0,1103	0,3133	0	1
30-49	Quintile 1	3 094	0,4663	0,4989	0	1
	Quintile 2	1 718	0,5848	0,4929	0	1
50-64	Quintile 1	3 094	0,3110	0,4630	0	1
	Quintile 2	1 718	0,2335	0,4232	0	1
65-91	Quintile 1	3 094	0,1387	0,3457	0	1
	Quintile 2	1 718	0,0714	0,2576	0	1
Head Education attainment	Quintile 1	3 094	7,6603	4,7068	0	24
	Quintile 2	1 717	13,0562	5,5409	0	24
None	Quintile 1	3 094	0,1296	0,3359	0	1
	Quintile 2	1 718	0,0238	0,1523	0	1
Primary	Quintile 1	3 094	0,3398	0,4737	0	1
	Quintile 2	1 718	0,1261	0,3321	0	1
Secondary	Quintile 1	3 094	0,4467	0,4972	0	1
	Quintile 2	1 718	0,4382	0,4963	0	1
Certificate	Quintile 1	3 094	0,0716	0,2579	0	1
	Quintile 2	1 718	0,1898	0,3922	0	1
Undergraduate	Quintile 1	3 094	0,0109	0,1039	0	1
	Quintile 2	1 718	0,1663	0,3724	0	1

Postgraduate	Quintile 1	3 094	0,0014	0,0368	0	1
	Quintile 2	1 718	0,0554	0,2288	0	1
Employed	Quintile 1	3 094	0,3336	0,4716	0	1
	Quintile 2	1 718	0,7729	0,4191	0	1
Food Expenditure Shock	Quintile 1	3 070	0,1116	0,3149	0	1
	Quintile 2	1 701	0,0470	0,2116	0	1
Health Shock	Quintile 1	3 094	0,0190	0,1365	0	1
	Quintile 2	1 718	0,0243	0,1541	0	1
Medical Aid Coverage	Quintile 1	3 094	0,0147	0,1204	0	1
	Quintile 2	1 718	0,3015	0,4590	0	1
Pension Recipient	Quintile 1	3 094	0,2182	0,4131	0	1
	Quintile 2	1 718	0,0590	0,2356	0	1

The total medical expenditure for higher income households on average is approximately four times higher than that of lower income households. However, on average 1.47 percent of lower income households' heads are covered by medical insurance while 30.15 percent of higher income households' heads are covered by medical insurance. Approximately 2 percent of the households on average experienced a health shock. Given the medical aid coverage, this implies that most households' face out of pocket medical spending if they experience a health shock. Nevertheless, lower income households on average experienced more food expenditure shocks (11.16%) than the higher income households (4.70%). This suggests that the Department of Health has an additional task to prioritise those who may belong to higher income quintile while they do not have medical aid. Therefore, the implication on the suggested NHI by DoH (2017) should consider those households that may fall under higher income quintile, so that the NHI does not only benefit the poor households as stipulated in the white paper.

The lower income households' size is on average higher than the higher income households' size, the higher categories of household size are all higher for the lower income households compared to the baseline category (1 to 3 members household size) while 86 percent of higher income households' size lie in the baseline income household. The higher income households' heads have on average higher education attainment of at least a certificate qualification than lower income households' heads with a grade 8. Therefore, it is not surprising that on average there is a higher percentage of employed heads of households in higher income quintile than those in lower income households. According to South African law, the government's pension grant recipient is only for those below a certain salary income. Thus, it is not surprising that on average, only 6% of households' heads in higher income quintile while 22% of lower income households receive the old age government grant.

Furthermore, the descriptive statistics on household and individual characteristics divided by employment status are presented in Table A.2 of the Appendix. The table shows that among the households' heads, there were on average 1.66 percent and 2.79 percent who experienced a health shock for those employed and unemployed, respectively. This shows that health shocks can significantly reduce labour supply; this is also confirmed by Beegle (2005), Lindelow and

Wagstaff (2005), Rocco et al. (2011) and Bales (2013) investigations. Similarly, households headed by those employed and unemployed, experienced food expenditure shock where by approximately 6 percent and 10% of households headed by employed and unemployed person experienced food expenditure shock.

Therefore, Section 6 will show results of models proposed in Section 4 whereby restrictions on the whole sample's datasets will be applied to demonstrate the treatment effect on different households' characteristics. The restrictions will include dividing the whole sample by quintiles, employment status and combination of quintiles and employment statuses.

6. RESULTS

6.1. Model I Results

Model I results were obtained from running a regression using the first dataset that allowed health shocks from wave 3 to wave 5. Thus, the results that will follow show the importance of dividing the data into two distinct logged income quintiles, where quintile 1 and quintile 2 are lower income households and higher income households respectively.

6.1.1. Model I insignificant models

Table 5 shows the results of running Model I for the whole sample, quintile 1, unemployed and quintile 1 and employed, respectively. The models are all insignificant, showing that lower income households do not adjust their food expenditure adversely in response to health shocks.

Table 5: Panel Logit Model Outputs

	Model I-I	Model I-II	Model I-III	Model I-III
	Whole sample	Quintile 1	Unemployed	Quintile 1 and Employed
Health shock	0.2977	-0.2141	-0.3345	0.7082
	(0.3812)	(0.4983)	(0.5715)	(0.9873)
Medical Aid coverage	0.0602	-1.3793	-1.2909	-0.6744
	(0.2838)	(0.8456)	(0.7997)	(0.9866)
Female	-0.0611	-0.2816	-0.0831	-0.3483
	(0.1892)	(0.2441)	(0.2694)	(0.4273)
Married	-0.3558*	-0.2903	-0.3019	0.0603
	(0.1968)	(0.2389)	(0.2763)	(0.4475)
Age				
30-49	-0.0065	0.1725	0.2335	0.0078
	(0.2903)	(0.3969)	(0.4592)	(0.6844)
50-64	-0.4262	-0.3328	-0.0926	-0.5176
	(0.3260)	(0.4397)	(0.4907)	(0.7641)
65-91	-0.1104	-0.1137	0.1680	-0.3127
	(0.4026)	(0.5306)	(0.5504)	(2.0880)
Education Attainment				
Primary	-0.0950	-0.1574	-0.2006	-0.1121

	(0.2498)	(0.2819)	(0.2926)	(0.6625)
Secondary	-0.7806***	-0.8000**	-0.8162**	-0.7074
	(0.2843)	(0.3364)	(0.3569)	(0.6976)
Certificate	-0.6652*	-0.3515	-0.6515	-0.9908
	(0.3650)	(0.4753)	(0.5578)	(0.8766)
Undergraduate	-0.0930	0.7380	0.7080	0.4705
	(0.4018)	(0.7427)	(0.7383)	(1.1673)
Postgraduate	-0.3338	-	-	-
	(0.7391)			
Household size				
4-6	0.7463***	0.7176***	0.8726***	0.5301
	(0.1720)	(0.2156)	(0.2386)	(0.4110)
7-9	0.9801***	0.9020***	0.9718***	1.0148*
	(0.2416)	(0.2811)	(0.3119)	(0.5602)
10-24	1.2936***	1.1886***	1.2777***	1.7016**
	(0.3625)	(0.4028)	(0.4547)	(0.8655)
Pension recipient	-0.5299**	-0.4015	-0.5685**	-0.8340
	(0.2567)	(0.3009)	(0.2896)	(1.2858)
Employed	-0.2855*	-0.4396**	-	-
	(0.1697)	(0.2119)		
Urban	0.1444	0.1296	0.0201	0.2517
	(0.2398)	(0.3045)	(0.3533)	(0.5302)
Toilet	-0.0792	0.1125	0.1131	-0.1229
	(0.2272)	(0.2868)	(0.3413)	(0.4870)
Piped Water	-0.0565	-0.1198	-0.1175	-0.0831
	(0.2247)	(0.2520)	(0.2704)	(0.5816)
Constant	-3.0914***	-3.0589***	-3.4250***	-3.4408***
	(0.4507)	(0.5629)	(0.6423)	(1.1085)
Observations	5053	3248	2822	947

Note: *** Significant at 1% level, ** significant at the 5% level, * Significant at the 10* level

The above results show that households living on the margin do not have room to further bring down food expenditure (Islam & Maitra, 2012). There exists different reasoning for this in literature. Mohanan (2013) has argued that the society usually intervene when the ability of a household's food consumption smoothing is limited by a shock particularly from poorer subpopulations. The community intervention can be in the form of neighbours and relatives of the household. This is further justified by Udry (1994), Townsend (1994, 1995) and Morduch (1995) who also confirmed that informal coping mechanisms (borrowing from landowners, relatives' transfers or other support network) are very popular in developing countries because of the absence or malfunctions of formal credit and insurance markets. Therefore, the former reasoning explains the possibility of negative coefficients, which shows that in some cases, there may even be an existence of a negative relationship. The other combinations of quintile 1 are insignificant too because they are all associated with lower income households. It is worth noting that this study did not test the coping mechanisms. However, the poor households cannot use food expenditure as coping mechanism, as it is already low. The poor may forego healthcare, as they potentially cannot trade-off their already constrained food consumption further for medical expenses.

6.1.2. Model I significant models

The following table shows the Odds Ratio (OR) version of the original Model I whose panel logit results are presented in Table A.3 in the appendix.

Table 6: Panel Logit Model Outputs (Odds ratios)

	Model I-I Quintile 2	Model I-II Employed	Model I-III Quintile 2 and Employed
Health shock	3.132*	2.751*	4.705**
	[0.82;12.02]	[0.95;7.96]	[1.10;20.14]
Medical Aid coverage	1.699	1.38	1.920*
	[0.88;3.27]	[0.75;2.53]	[0.92;4.00]
Female	1.168	0.995	1.154
	[0.64;2.16]	[0.61;1.63]	[0.57;2.36]
Married	0.592	0.698	0.405*
	[0.27;2.15]	[0.39;1,25]	[0.16;1.04]
Age			
30-49	0.728	0.835	0.733
	[0.31;1.72]	[0.41;1.70]	[0.29;1.84]
50-64	0.616	0.485*	0.324*
	[0.23;1.69]	[0.21;1.13]	[0.10;1.06]
65-91	1.760	1.224	1.303*
	[0.45;6.94]	[0.20;7.61]	[0.14;11.85]
Education Attainment			
Primary	0.993	1.422	3.404
	[0.28;3.49]	[0.53;3.85	[0.43;27.20]
Secondary	0.299*	0.617	0.870
	[0.08;1.08]	[0.224;1.696]	[0.11;6.69]
Certificate	0.298*	0.584	1.021
	[0.07;1.22]	[0.19;1.81]	[0.12;8.44]
Undergraduate	0.576	0.941	1.457
	[0.14;2.30]	[0.30;3.00]	[0.18;12.00]
Postgraduate	0.736	1.138	2.597
	[0.11;4.81]	[0.21;6.21]	[0.21;31.39]
Household size			
4-6	1.688	1.817**	1.793
	[0.84;3.39]	[1.10;3.01]	[0.78;4.11]
7-9	3.014	2.950**	6.061*
	[0.68;13.37]	[1.328;6.55]	[0.90;40.76]
10-24	7.644	5.676**	24.057*
	[0.33;174.57]	[1.52;21.25]	[0.67;864.65]
Pension recipient	0.207**	0.520	0.912

	[0.05;0.78]	[0.10;2.69]	[0.07;11.31]
Employed	1.565	-	-
	[0.73;3.37]		
Urban	1.213	1.406	1.816
	[0.54;2.71]	[0.74;2.68]	[0.70;4.75]
Toilet	0.7257	0.742	0.673
	[0.33;1.60]	[0.41;1.36]	[0.27;1.70]
Piped Water	1.609	0.945	1.200
	[0.43;6.00]	[0.41;2.16]	[0.25;5.83]
Constant	0.021***	0.034***	0.012***
	[0.00;0.17]	[0.01;0.14]	[0.00;0.18]
Observations	1801	2225	1275

Note: *** Significant at 1 percent level, ** significant at the 5 percent level, * Significant at the 10* level. The confidence intervals are in the parenthesis. The presented odds ratios are compared to the base lines for example: Education Attainment base is no schooling with OR=1 and Household size base is 1-3 members with OR=1

The above Table 6 shows that the health shock is one of the factors that are associated with household's food expenditure shock (poverty incidence) because the OR for all 3 models are above 1 and statistically significant at 1 percent and 5 percent level for both Model I-I and Model I-II, and Model I-III, respectively. For example, consider the Model I-III, which only considers sample of households with high logged-income and employed household heads. The OR for health shock is 4.705, which means that if a head of a household switches from being healthy to being ill, the household's odds of being in poverty are multiplied by 4.705. Falling ill does not help households to stay above the poverty line. There exists a shortage of literature that explores the reasons why households might want to increase non-medical consumption in clean fuel, blankets, food, blankets etc (Wagstaff, 2007). Results lead towards the conclusion that households respond by reallocating households' spending towards renovating their houses in order to make them suitable for a member who is hospitalised and allocating away from food expenditure (Wagstaff, 2007).

Furthermore, Model I-III shows that a number of other factors were associated with households' food expenditure shock: higher number of household members (household size between 10-24 and household size between 7-9 members) versus lower household size of 1-3 members; household head being covered by a medical aid with OR=1.920 and household head aged over 65 with OR=1.303. These results are in line with the finding of Atake (2018) who concluded that households with higher household's size are more vulnerable to poverty.

Various factors were associated with decreasing the chances of food expenditure shock (poverty incidence): household heads aged between 50 and 64 (OR=0.324) and household heads with spouses (OR=0.405). Showing that getting married, and older than 50 and before 65 years of age help household to stay above the poverty line. Model I-I showed that having a secondary or higher education certificate qualification decreases the chances of food expenditure shock as shown by Odds ratios of 0.299 and 0.298 respectively at 10 percent significance level.

The results outlined in Table 6 above corresponds to Table A.3 in the appendix. However, unlike the above, Table A.3 shows the direction of the effect of the explanatory variables on the food expenditure shock. For instance, household heads aged between 50-64 and households head who are married are negatively associated with the food expenditure shock. While households sized 7-9 and 10-24, and households' heads covered by medical aid are positively associated with households' food expenditure shock.

The expectations were that piped water (safe drinking water) and access to toilets that flush (sanitation) would result in a decrease in food expenditure shock. Satterthwaite (2003) stated that the reason for this is that insufficient provision of both descent sanitation and piped water results in increased vulnerability. However, the coefficients are insignificant. This may be because higher proportions of the households in the study have access to pipe water and descent sanitation (as shown in Table 4). This resulted in removal of the covariates in the next model.

6.2. Model II Results

The previous model cannot compute the interaction effect for a panel data logit with random effects because the model assumes homogeneity (Karaca-Mandic et al., 2012). To resolve the exclusion elimination of the interaction effect, a DID estimator can be used. However, the DID estimator assumes that there exists a parallel trend between pre-treatment period and post treatment period. Therefore, since the investigation is based on three periods, the parallel trend test used period 1 and period 3. Dummy variables for each time period was created and each period allowed to interact with the treatment group. A logit regression was estimated with two interactions and covariates. The results are presented in Table A.4 of the Appendix, which shows that period 1 interaction is insignificant. This approach of parallel trend test followed the Autor (2003) methodology. This implies that the following results obtain for Model II non-linear DID estimation are accurate which uses the quasi-experimental dataset.

Table 7: Difference in Differences Model output

	Model II-I	Model II-II	Model II-III	Model II-IV	Model II-V	Model II-VI	Model II-VII	Model II-VIII	Model II-IX
	whole sample	Quintile 1	Quintile 2	Employed	Unemployed	Quintile 1 and employed	Quintile 2 and employed	Quintile 1 and unemployed	Quintile 2 and unemployed
Post	0.4467**	0.6489***	-0.0034	0.2349	0.6165***	0.4017	0.0451	0.7800***	-0.1098
	(0.1797)	(0.2000)	(0.3598)	(0.2934)	(0.2129)	(0.3960)	(0.4259)	(0.2199)	(0.8430)
Treat	-1.5646***	-1.1333*	-2.7073**	-2.3761***	-1.1726*	-2.0563*	-2.4585*	-0.8195	-
	(0.5522)	(0.5988)	(1.2104)	(0.9205)	(0.6711)	(1.2488)	(1.3388)	(0.6396)	
Treat*Post	2.3725***	1.3821	5.7218***	5.2347***	1.4637	-	6.9960***	1.2077	-
	(0.8638)	(0.9582)	(1.6597)	(1.4717)	(1.0057)		(2.1305)	(1.0091)	
Medical Aid coverage	-0.4324	-1.8649*	0.1521	-0.1923	-1.6165*	-1.5192	0.3173	-	-1.3355
	(0.3151)	(1.0191)	(0.4207)	(0.3516)	(0.8306)	(1.0190)	(0.4511)		(1.0053)
Female	-0.3529**	-0.6428***	0.1123	-0.4688*	-0.2031	-1.1088***	0.0353	-0.3356	0.3133
	(0.1786)	(0.2009)	(0.3070)	(0.2522)	(0.2652)	(0.3494)	(0.3485)	(0.2705)	(0.5738)
Married	0.1271	0.3771	-0.7280	0.0036	0.2283	0.4407	-0.7482	0.3069	-0.9602

	(0.2350)	(0.2590)	(0.5114)	(0.3515)	(0.3108)	(0.4656)	(0.6087)	(0.3181)	(0.8299)
Age									
30-49	0.0535	0.1389	-0.3026	-0.1965	0.3435	-0.3880	-0.1401	0.5459	-1.9024
	(0.3383)	(0.4514)	(0.4852)	(0.4968)	(0.4100)	(0.7937)	(0.5574)	(0.4718)	(1.1856)
50-64	-0.1950	-0.2339	-0.1898	-0.4610	0.0689	-0.8925	-0.2823	0.2090	-0.3240
	(0.3848)	(0.4951)	(0.6246)	(0.6134)	(0.4348)	(0.9457)	(0.7398)	(0.4938)	(0.9720)
65-91	0.4317	0.1172	1.3573	0.8368	0.8077	-1.2810	1.2573	0.6545	1.1477
	(0.5817)	(0.5855)	(0.9484)	(1.2306)	(0.6148)	(1.5900)	(1.4572)	(0.6015)	(1.2141)
Education Attainment									
Primary	0.4556*	0.4894*	0.1119	0.3488	0.4928**	0.4551	0.3046	0.5152**	-0.5387
	(0.2557)	(0.2723)	(0.8645)	(0.6269)	(0.2427)	(0.8315)	(1.1159)	(0.2517)	(1.0107)
Secondary	-0.1041	-0.1329	-0.2871	-0.3787	0.0419	-0.5029	-0.3121	0.0476	0.2319
	(0.3054)	(0.3476)	(0.7648)	(0.6247)	(0.3309)	(0.8977)	(0.9703)	(0.3525)	(0.9315)
Certificate	-0.2325	0.1413	-0.5652	-0.7093	0.4822	-0.6735	-0.5895	0.7788*	-0.1215
	(0.4082)	(0.4645)	(0.8959)	(0.7525)	(0.4224)	(1.0726)	(1.1005)	(0.4642)	(1.1570)
Undergraduate	0.4743	1.1142	0.3199	-0.0072	1.3603**	-0.4223	0.2022	1.8802*	1.7467
	(0.4234)	(0.7448)	(0.8083)	(0.7128)	(0.6127)	(1.5739)	(1.0096)	(0.9647)	(1.1201)
Postgraduate	-1.2023	-	-1.4044	-1.4381	-	-	-1.3664	-	-
	(0.8319)		(1.1544)	(1.0639)			(1.3357)		
Household size									
4-6	0.6904***	0.3840*	0.7931*	0.8415***	0.5508**	0.3670	0.8044*	0.3438	0.6891
	(0.1910)	(0.2211)	(0.4239)	(0.2866)	(0.2538)	(0.4023)	(0.4812)	(0.2702)	(0.6875)
7-9	0.6020**	0.2663	1.0036	0.5625	0.5642**	0.0866	0.5235	0.3018	0.8960
	(0.2396)	(0.2582)	(0.6124)	(0.4888)	(0.2777)	(0.5328)	(0.9766)	(0.2958)	(1.1163)
10-24	0.8836***	0.5037	2.7845**	1.2564**	0.6870*	0.8480	1.8755	0.3212	3.6079**
	(0.3133)	(0.3395)	(1.1118)	(0.5626)	(0.3867)	(0.6518)	(1.2631)	(0.4050)	(1.4701)
Pension recipient	-0.6926*	-0.4501	-0.8109	0.1845	-0.8227**	0.3215	0.9588	-0.5504	-1.8676**
	(0.3762)	(0.3139)	(0.8391)	(0.8054)	(0.3877)	(0.8664)	(1.2221)	(0.3353)	(0.9303)
Employed	-0.2983	-0.2326	0.8722*	-	-	-	-	-	-
	(0.2011)	(0.2290)	(0.5026)						
Urban	0.1345	0.1702	0.3577	0.2056	0.1091	0.2029	0.4867	0.2034	-0.4871
	(0.1688)	(0.1922)	(0.3161)	(0.2868)	(0.2134)	(0.4172)	(0.3575)	(0.2193)	(0.5971)
Constant	-2.3943***	-2.0273***	-3.7639***	-2.2263***	-2.8511***	-1.1273	-3.1458***	-2.8101***	-2.5993**
	(0.4143)	(0.5053)	(1.0350)	(0.6957)	(0.4951)	(0.9510)	(1.1237)	(0.5562)	(1.1027)
Observations	4771	3067	1701	2093	2673	890	1196	2161	491
Margins Treat*Post (P-value)	0.0693	0.2272	0.0319	0.0134	0.2253	-	0.0002	0.2827	-

Note: *** Significant at 1% level, ** significant at the 5% level, * Significant at the 10* level

Puhani (2012) showed that the coefficient on the interaction term post x treatment represents the treatment effect in the non-linear 'Difference-in-Differences' model. The above results shows that there exists a positive treatment effect, showing that the health shock increases the likelihood of food expenditure shock by more than 2 times in four models (whole sample, quintile 2, employed, and quintile 2 and employed) at 1 percent significant level. There exists a positive relationship between the presence of a health shock and food expenditure shock which is shown by the coefficient of treatment effect (Treat*Post) especially for quintile 2 and the combinations of quintile 2 except with unemployment sample. This model shows similar results to the panel logit Model I presented in subsection 6.1 because those quintile 1, unemployed and combinations of quintile 1 are all still insignificant. However, unlike Model I, the whole sample's non-linear DID treatment estimator is significant. The reasoning for insignificant health shock effect for those in quintile 1 and significant health shock effect for those in upper quintile is the same as above subsection 6.1.

Other covariates, consider the results for Model II-I which shows that the likelihood of health shock increases as household's size increases more than the 1 to 3 members sized households. It goes further and shows that that households with a recipient of old age grant are less likely to experience the food expenditure shock.

6.3. Model III Results

The Propensity Scores were obtained from psmatch2 which allowed the estimation of the following PSM-DID model.

Table 6.3: PSM-DID results (Propensity Score Weighted)

	Model III-I	Model III-II	Model III-III	Model III-IV
	whole sample	Quintile 1	Quintile 2	Employed
Post	0.3262***	0.0630	1.3604**	0.6925**
	(0.1191)	(0.1488)	(0.5706)	(0.3522)
Treat	-0.2585	-0.5202	-13.3631***	0.3963
	(0.5047)	(0.5430)	(0.6153)	(1.0840)
Treat*Post	0.3089	-1.0599	17.1028***	1.5651
	(0.7156)	(1.1705)	(1.2857)	(1.7790)
Medical Aid coverage	-0.4790**	-1.0962	-0.3836	-0.2437
	(0.2175)	(1.0160)	(0.5286)	(0.4401)
Female	-0.3007**	-0.2629	0.8439**	-0.3236
	(0.1299)	(0.1643)	(0.4294)	(0.3388)
Married	-0.0083	0.0391	-0.4356	0.0759
	(0.1704)	(0.2030)	(0.5216)	(0.3913)
Age				
30-49	-0.1317	-0.1542	-0.0672	-0.3657
	(0.2079)	(0.3046)	(0.7032)	(0.6075)
50-64	-0.3690	-0.3635	-1.3979*	-0.5624
	(0.2308)	(0.3333)	(0.7671)	(0.5855)
65-91	-0.3405	-0.1187	-4.1163***	0.8377
	(0.2848)	(0.3960)	(1.2383)	(0.9211)
Education Attainment				
Primary	0.1152	0.1670	-0.5827	-0.0308
	(0.1718)	(0.1884)	(0.8306)	(0.7561)

Secondary	-0.3041	-0.0475	-1.1580	0.1900
	(0.1972)	(0.2271)	(0.9666)	(0.6984)
Certificate	-0.1611	0.0700	-2.2624**	-0.7184
	(0.2649)	(0.3524)	(0.8786)	(0.8179)
Undergraduate	0.2542	0.2708	-0.7904	0.4101
	(0.2862)	(0.8044)	(0.8221)	(0.7829)
Postgraduate	0.1079	-	1.1501	0.7148
	(0.6424)		(1.2393)	(1.1035)
Household size				
4-6	0.7147***	0.5060***	0.4346	0.6103*
	(0.1319)	(0.1661)	(0.4997)	(0.3367)
7-9	0.7716***	0.6369***	-1.0352	1.1927***
	(0.1777)	(0.2009)	(1.0420)	(0.4508)
10-24	0.6637**	0.4996*	3.1548**	1.2051
	(0.2637)	(0.2939)	(1.3187)	(0.9059)
Pension recipient	-0.2098	-0.2105	1.0420	-0.2637
	(0.1895)	(0.2304)	(1.5078)	(0.8623)
Employed	-0.3673***	-0.2649	0.3902	
	(0.1412)	(0.1716)	(0.9611)	
Urban	0.1581	0.2168	0.2089	0.2746
	(0.1240)	(0.1440)	(0.5505)	(0.3382)
Constant	-2.0684***	-2.1681***	-2.1403	-2.4363***
	(0.2761)	(0.3803)	(1.5954)	(0.7744)
Observations	4,771	2,678	1,122	1,572
Margins Treat*Post (P-value)	0.6681	0.3142	0.0000	0.3689

*Note: *** Significant at 1% level, ** significant at the 5% level, * Significant at the 10% level*

The above PSM-DID results shows that the quintile 2 Model III-III is the only model, which passes the interaction between treatment and post treatment period test because the p value of the margins for treatment effect is significant at 1% significant level.

Asfaw and Braun (2004), Wagstaff (2005), De Weerdt and Dercon (2006), Khan (2010), Wagstaff and Lindelow (2010) and Genoni (2012) found that the health shocks decrease food consumption by 1.80%, 17.30%, 4.80%, 15.30%, 18.80% and 4.20% respectively. This is in line with the above DID treatment effect finding that the household which experience a health shock are surprisingly 17 times likely to experience food expenditure shock. This figure suggests that the suggested treatment effect by Model I (Quintile 2) of 5.72 underestimates the relationship between the health shock and food expenditure shock of households. These results contradict the study conducted by Gertler and Gruber (2002) which concluded that households which higher income seem to be better insured against negative effects of illness shocks.

Education shows to be a strong protection against poverty vulnerability (Dhanaraj, 2016). The results show that when the head of household has at least a technical certificate, it turns to decrease the likelihood of food expenditure shock. This is consistent with the findings based on Nigeria data that concluded that in general households with heads who have no schooling are vulnerable (Adepoju & Okunmadewa, 2010). This may be due to the belief that the more education attainment one has, the higher the chances of decent job description.

A reduction of the well-being of the household members occurs due to the large household size that always in turn increases vulnerability (Jalan & Ravallion, 1998; Meenakshi & Ray, 2002; Atake, 2015). An increase in household's size turns to increase vulnerability to poverty in sub Saharan countries (Novignon et al., 2012; Adepoju & Okunmadewa, 2010). The results are consistent with expectations because the higher household size turn be associated with an increase in the likelihood of food expenditure shock.

The Model III (PSM-DID) robustness tests are captured in Table A.5 and Table A.6 in the Appendix. Table A.5 shows that the whole sample's matched pscores result in reduction of selection bias. The MHbounds results in Table A.6 show that the study is sensitive to a bias that will double the odds of experiencing a health shock. However, from Gamma equals to six, the significant level starts to improve. Implying that the PSM-DID estimator for quintile 2 results are likely not sensitive to selection bias.

7. CONCLUSION

This study investigated the relationship between household heads' health shocks and poverty incidence as measured by food expenditure shock. The health shock is defined as a significant decrease in change in BMI that is more than one standard deviation. This approach of BMI change follows the Wagstaff (2007) because BMI has been found to be a good predictor of mortality, with mortality risk higher at both the bottom and top ends of the scale (Calle et al., 1999). The investigation then applied the former approach to detect the presence of food expenditure shock in households. Food expenditure shock specifically indicates whether the household experience a significant decrease in growth rate of logged real food expenditure that is more than one standard deviation.

Therefore, non-linear models were applied because the dependent variable (food expenditure shock) is a binary variable. The models included, Panel Logit model, non-linear DID model and PSM with DID model. However, the study only addresses the selection bias and not the reverse causality because there are limitations of the instrumental variables to use for health shocks. Nevertheless, the study describes treatment effects as association and not causal effect.

The three models' results are consistent and show that higher income households (quintile 2) are most likely to experience food expenditure shock in the presence of health shock. This is due to reasoning that households respond to health shocks by reallocating the households' funds towards renovating the house to accommodate the ill patient and the funds away from food expenditure (Wagstaff, 2007). Generally, the lower income households (quintile 1) has insignificant health shock effect on food expenditure shock. The informal coping mechanisms as suggested by Udry (1994), Townsend (1994, 1995) and Morduch (1995) have not been explored in South African context. However, the poor cannot use food expenditure as coping mechanism, as it is already low. How the poor cope in South Africa in the event of a health shock is therefore open for future research.

Furthermore, higher household sizes show to increase the vulnerability of households to poverty. Whereby the presence of old age grant recipient results in reducing the likelihood of households falling into poverty.

Overall, the poor in South Africa may forego healthcare as they potentially cannot trade-off their already constrained food consumption further for medical expenses. The households in the upper quintile, however, sacrifice food consumption to access healthcare highlighting their vulnerability to health shocks in the absence of medical insurance. This indicates that in South Africa healthcare is a privilege, which is only available to the minority with medical insurance. The vast majority of population is therefore vulnerable due to lack of healthcare insurance underlining the need for a National Health Insurance that would enable universal access to health care in the country.

8. REFERENCE LIST

Abegunde, D. O., & Stanciole, A. E. (2008). The economic impact of chronic diseases: how do households respond to shocks? Evidence from Russia. *Social science & medicine*, *66*(11), 2296-2307.

Adepoju, A. O., & Okunmadewa, F. Y. (2010). Households'vulnerability to poverty in Ibadan Metropolis, Oyo State, Nigeria. *Journal of Rural Economics and Development*, *20*(1623-2016-134908), 44.

Ainsworth, M., & Semali, I. (2000). *The impact of adult deaths on children's health in northwestern Tanzania*. The World Bank.

Alam, K., & Mahal, A. (2014). Economic impacts of health shocks on households in low and middle income countries: a review of the literature. *Globalization and health*, *10*(1), 21.

Asfaw, A., & Braun, J. V. (2004). Is consumption insured against illness? Evidence on vulnerability of households to health shocks in rural Ethiopia. *Economic Development and Cultural Change*, *53*(1), 115-129.

Ashenfelter, O. C., & Card, D. (1984). Using the longitudinal structure of earnings to estimate the effect of training programs.

Atake, E. H. (2015). Technical efficiency of public hospitals in Togo: a directional distance function approach. *Econ Bull*, *35*, 1752-64.

Atake, E. H. (2018). Health shocks in sub-Saharan Africa: are the poor and uninsured households more vulnerable?. *Health economics review*, *8*(1), 26.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of labor economics*, *21*(1), 1-42.

Bales, S. (2013). Impact of health shocks on household welfare in Vietnam—Estimates using fixed effects estimation. *Institute of Health Policy & Management (HEFPA) Working Paper*, *18*.

Banerjee, A. V., Banerjee, A., & Duflo, E. (2011). *Poor economics: A radical rethinking of the way to fight global poverty*. Public Affairs.

Bartels, B. (2008). Beyond" fixed versus random effects": a framework for improving substantive and statistical analysis of panel, time-series cross-sectional, and multilevel data. *The Society for Political Methodology*, *9*, 1-43.

Becker, S. O., & Caliendo, M. (2007). Sensitivity analysis for average treatment effects. *The Stata Journal*, *7*(1), 71-83.

Beegle K. (2005). Labor effects of adult mortality in Tanzanian households. *Economic Development and Cultural Change*, 53:655–683.

Bridges, S., & Lawson, D. (2008). *Health and labour market participation in Uganda* (No. 2008/07). WIDER Discussion Papers, World Institute for Development Economics (UNU-WIDER).

Borooah, V. K. (2002). *Logit and probit: Ordered and multinomial models* (No. 138). Sage.

Calle, E. E., Thun, M. J., Petrelli, J. M., Rodriguez, C., & Heath Jr, C. W. (1999). Body-mass index and mortality in a prospective cohort of US adults. *New England Journal of Medicine*, *341*(15), 1097-1105.

Chaudhuri, S., Jalan, J., & Suryahadi, A. (2002). *Assessing household vulnerability to poverty from cross-sectional data: A methodology and estimates from Indonesia* (Vol. 102, p. 52). Discussion paper.

Chinhema, M., Brophy, T., Brown, M., Leibbrandt, L., Mlatsheni, C. & Woolard, I., eds. (2016). *National Income Dynamics Study Panel User Manual*. Southern Africa Labour and Development Research Unit: Cape Town.

Deininger, K., Garcia, M., & Subbarao, K. (2003). AIDS-induced orphanhood as a systemic shock: Magnitude, impact, and program interventions in Africa. *World development*, *31*(7), 1201-1220.

Department of Health (DoH). (2017). National Health Insurance for South Africa. Towards Universal Health Coverage.

Dercon, S., & Krishnan, P. (2000). In sickness and in health: Risk sharing within households in rural Ethiopia. *Journal of political Economy*, *108*(4), 688-727.

De Weerdt, J., & Dercon, S. (2006). Risk-sharing networks and insurance against illness. *Journal of development Economics*, *81*(2), 337-356.

Dhanaraj, S. (2014). *Health shocks and coping strategies: State health insurance scheme of Andhra Pradesh, India* (No. 2014/003). WIDER Working Paper.

Dhanaraj, S. (2016). Economic vulnerability to health shocks and coping strategies: evidence from Andhra Pradesh, India. *Health policy and planning*, *31*(6), 749-758.

Genoni, M. E. (2012). Health shocks and consumption smoothing: Evidence from Indonesia. *Economic Development and Cultural Change*, *60*(3), 475-506.

Gertler, P., & Gruber, J. (2002). Insuring consumption against illness. *American economic review*, *92*(1), 51-70.

Ghosh, S. (2011). Catastrophic payments and impoverishment due to out-of-pocket health spending. *Economic and Political Weekly*, 63-70.

Hoddinott, J., & Quisumbing, A. (2010). Methods for microeconometric risk and vulnerability assessment. In *Risk, Shocks, and Human Development* (pp. 62-100). Palgrave Macmillan, London.

Islam, A., & Maitra, P. (2012). Health shocks and consumption smoothing in rural households: Does microcredit have a role to play?. *Journal of development economics*, *97*(2), 232-243.

Jalan, J., & Ravallion, M. (1998). Transient poverty in postreform rural China. *Journal of Comparative Economics*, *26*(2), 338-357.

James, W. P. T., Ferro-Luzzi, A., & Waterlow, J. C. (1988). Report of a Working Party of the International Dietary Energy Consultative. *European journal of clinical nutrition*, *42*, 981.

Kadiyala, S., Quisumbing, A., Rogers, B., & Webb, P. (2009). The impact of prime age adult mortality on child survival and growth in rural Ethiopia. *World Development*, *37*(6), 1116-1128.

Kadiyala, S., Rogers, B., Quisumbing, A., & Webb, P. (2011). The effect of prime age adult mortality on household composition and consumption in rural Ethiopia. *Food policy*, *36*(5), 647-655.

Karaca-Mandic, P., Norton, E. C., & Dowd, B. (2012). Interaction terms in nonlinear models. *Health services research*, *47*(1pt1), 255-274.

Khan, F. U. (2010). *Economic consequences of health shocks and coping strategies: Evidence from urban poor households in Bangladesh*. Erasmus University.

Kieny, M. P., Bekedam, H., Dovlo, D., Fitzgerald, J., Habicht, J., Harrison, G., ... & Siddiqi, S. (2017). Strengthening health systems for universal health coverage and sustainable development. *Bulletin of the World Health Organization*, *95*(7), 537.

Leive, A., & Xu, K. (2008). Coping with out-of-pocket health payments: empirical evidence from 15 African countries. *Bulletin of the World Health Organization*, *86*, 849-856C.

Leuven, E., & Sianesi, B. (2018). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.

Ligon, E., & Schechter, L. (2003). Measuring vulnerability. *The Economic Journal*, *113*(486), C95-C102.

Lindelow, M., & Wagstaff, A. (2005). *Health shocks in China: are the poor and uninsured less protected?*. The World Bank.

Mahal, A., Canning, D., Odumosu, K., & Okonkwo, P. (2008). Assessing the economic impact of HIV/AIDS on Nigerian households: a propensity score matching approach. *Aids*, *22*, S95-S101.

Mahmoud, T. O., & Thiele, R. (2013). Does prime-age mortality reduce per-capita household income? evidence from rural Zambia. *World Development*, *45*, 51-62.

Meenakshi, J. V., & Ray, R. (2002). Impact of household size and family composition on poverty in rural India. *Journal of Policy Modeling*, *24*(6), 539-559.

Mohanan, M. (2013). Causal effects of health shocks on consumption and debt: quasi-experimental evidence from bus accident injuries. *Review of Economics and Statistics*, *95*(2), 673-681.

Morduch, J. (1995). Income smoothing and consumption smoothing. *Journal of economic perspectives*, *9*(3), 103-114.

Murphy, A., Mahal, A., Richardson, E., & Moran, A. E. (2013). The economic burden of chronic disease care faced by households in Ukraine: a cross-sectional matching study of angina patients. *International journal for equity in health*, *12*(1), 38.

Novignon, J., Nonvignon, J., Mussa, R., & Chiwaula, L. S. (2012). Health and vulnerability to poverty in Ghana: evidence from the Ghana Living Standards Survey Round 5. *Health economics review*, *2*(1), 11.

Puhani, P. A. (2012). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. *Economics Letters*, *115*(1), 85-87.

Ramaraj, R., & Alpert, J. S. (2008). Indian poverty and cardiovascular disease. *The American journal of cardiology*, *102*(1), 102-106.

Ravallion, M., & Chen, S. (2007). *Absolute poverty measures for the developing world, 1981-2004*. The World Bank.

Reichardt, C. S., Minton, B. A., & Schellenger, J. D. (1980). The analysis of covariance (ANCOVA) and the assessment of treatment effects. *Prevention Evaluation Research Monograph II-Outcome (Section D4, chap. VIII). Denver: University of Denver, Department of Psychology*.

Rocco, L., Tanabe, K., Suhrcke, M., & Fumagalli, E. (2011). *Chronic diseases and labor market outcomes in Egypt*. The World Bank.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American statistical Association*, *79*(387), 516-524.

Satterthwaite, D. (2003). The links between poverty and the environment in urban areas of Africa, Asia, and Latin America. *The Annals of the American Academy of Political and Social Science*, *590*(1), 73-92.

Smith, J. A., & Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators?. *Journal of econometrics*, *125*(1-2), 305-353.

Somi, M. F., Butler, J. R., Vahid, F., Njau, J. D., & Abdulla, S. (2009). Household responses to health risks and shocks: a study from rural Tanzania raises some methodological issues. *Journal of International Development: The Journal of the Development Studies Association*, *21*(2), 200-211.

Statistics South Africa (StatsSA). (2016). Community survey 2016 in brief, Report 03-01-06. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2017a). General Household Survey 2017, Statistical release P0318. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2017b). Poverty trends in South Africa: An examination of absolute poverty between 2006 and 2015. Pretoria, South Africa: Statistics South Africa.

Statistics South Africa. (2018). Quartely Labour Force Survey, Statistical release P0211. Pretoria, South Africa: Statistics South Africa.

Tesliuc, E., & Lindert, K. (2002). Vulnerability: A quantitative and qualitative assessment. *Guatemala Poverty Assessment Program*, 1-91.

Townsend, R. M. (1994). Risk and insurance in village India. *ECONOMETRICA-EVANSTON ILL-*, *62*, 539-539.

Townsend, R. M. (1995). Consumption insurance: An evaluation of risk-bearing systems in low-income economies. *Journal of Economic perspectives*, *9*(3), 83-102.

Udry, C. (1994). Risk and insurance in a rural credit market: An empirical investigation in northern Nigeria. *The Review of Economic Studies*, *61*(3), 495-526.

Van Minh, H., Phuong, N. T. K., Saksena, P., James, C. D., & Xu, K. (2013). Financial burden of household out-of pocket health expenditure in Viet Nam: findings from the National Living Standard Survey 2002–2010. *Social science & medicine*, *96*, 258-263.

Waaler, H. T. (1984). Height. Weight and mortality the Norwegian experience. *Acta medica scandinavica*, *215*(S679), 1-56.

Wagstaff, A. (2005). *The economic consequences of health shocks*. The World Bank.

Wagstaff, A. (2007). The economic consequences of health shocks: evidence from Vietnam. *Journal of health economics*, *26*(1), 82-100.

Wagstaff, A., & Lindelow, M. (2010). *Are health shocks different? Evidence from a multi-shock survey in Laos*. The World Bank.

Whitehead, M., Dahlgren, G., & Evans, T. (2001). Equity and health sector reforms: can low-income countries escape the medical poverty trap?. *The Lancet*, *358*(9284), 833-836.

World Bank. (2012). Health expenditure and finance in South Africa. Public expenditure review (PER). Washington, DC: World Bank.

Yamano, T., & Jayne, T. S. (2004). Measuring the impacts of working-age adult mortality on small-scale farm households in Kenya. *World Development*, *32*(1), 91-119.

9. APPENDIX

Table A.1: Description of Variables

Variable	Description	Type
Food expenditure shock	Identifier of whether the household experienced a significant decrease in growth rate of logged real total food expenditure or not. Y= 1 if the decrease in growth rate of food expenditure was more than the standard deviation of growth rate of logged real food expenditure Y=0 if otherwise	Dummy
Health shock	Identifier of whether the head of the household experienced a significant decrease in BMI change or not X=1 if the decrease in BMI change was more than one standard deviation of the BMI-change distribution X=0 if otherwise	Dummy
HH Income	Real monthly household total income per capita in thousands. Deflated with StatsSA consumer price index data using December 2012 CPI as a base. In rands.	Continuous
HH Medical Expenditure	Real monthly household total medical expenditure per capita in thousands. Deflated with StatsSA consumer price index data using December 2012 CPI as a base. In rands.	Continuous
HH food exp	Real monthly household total food expenditure per capita in thousands. Deflated with StatsSA consumer price index data using December 2012 CPI as a base. In rands.	Continuous
Logged Income	Logged real monthly household total income per capita	
Logged Food exp	Logged real monthly household total food expenditure	Continuous
Urban	Identifier of whether the household' dwelling is located in urban area or in traditional dwelling and farms X=1 if household is in urban location X=0 if otherwise	Dummy
Employed	Identifier of whether self-employed or employee X=1 if self-employed or employee X=0 if otherwise	Dummy
Piped water	Identifier of whether the household has access to piped water X=1 if the source of water is from within dwelling, in the yard or public tap X=0 if otherwise	Dummy
Toilet	Identifier of whether the household has access to a toilet that flushes X=1 if the toilet is in the dwelling or in yard X=0 if otherwise	Dummy

Table A.2: Summary Statistics of Household and Individual Characteristics (Average Weighted, Employment)

Variable	Sample	N	Mean	Std. Error	Min	Max
	FULL SUB-SAMPLE					
Total Medical Expenditure	Employed	2 101	91,95	734,63	0	10224,95
	Unemployed	2 711	25,66	135,16	0	3565,06
HH Food Expenditure	Employed	2 101	663,66	689,85	0	8179,96
	Unemployed	2 711	350,20	425,60	0	4089,98
HH Income	Employed	2 101	5759,22	10021,17	91,67	152777,80
	Unemployed	2 711	1665,24	2842,23	51,12	33026,59
Logged Income	Employed	2 101	7,9574	1,1357	4,52	11,94
	Unemployed	2 711	6,8010	1,0230	3,93	10,41
Urban	Employed	2 101	0,7683	0,4220	0	1
	Unemployed	2 711	0,5876	0,4924	0	1
Household Size	Employed	2 101	2,4622	2,1047	1	24
	Unemployed	2 711	3,6360	2,6964	1	19
1-3	Employed	2 101	0,7589	0,4279	0	1
	Unemployed	2 711	0,5571	0,4968	0	1
4-6	Employed	2 101	0,1868	0,3898	0	1
	Unemployed	2 711	0,3049	0,4605	0	1
7-9	Employed	2 101	0,0427	0,2023	0	1
	Unemployed	2 711	0,1060	0,3079	0	1
10-24	Employed	2 101	0,0116	0,1072	0	1
	Unemployed	2 711	0,0320	0,1761	0	1
Age	Employed	2 101	42,1346	10,0079	18	84
	Unemployed	2 711	51,5072	15,3045	17	91
17-29	Employed	2 101	0,1069	0,3091	0	1
	Unemployed	2 711	0,0850	0,2789	0	1
30-49	Employed	2 101	0,6490	0,4774	0	1
	Unemployed	2 711	0,3727	0,4836	0	1
50-64	Employed	2 101	0,2343	0,4237	0	1
	Unemployed	2 711	0,3193	0,4663	0	1
65-91	Employed	2 101	0,0098	0,0984	0	1
	Unemployed	2 711	0,2231	0,4164	0	1
Head Education attainment	Employed	2 100	12,3423	5,5864	0	24
	Unemployed	2 711	7,8977	5,0858	0	24
None	Employed	2 101	0,0333	0,1796	0	1
	Unemployed	2 711	0,1304	0,3368	0	1
Primary	Employed	2 101	0,1508	0,3580	0	1
	Unemployed	2 711	0,3348	0,4720	0	1
Secondary	Employed	2 101	0,4530	0,4979	0	1
	Unemployed	2 711	0,4295	0,4951	0	1
Certificate	Employed	2 101	0,1820	0,3860	0	1
	Unemployed	2 711	0,0671	0,2502	0	1
Undergraduate	Employed	2 101	0,1346	0,3413	0	1
	Unemployed	2 711	0,0316	0,1749	0	1

Postgraduate	Employed	2 101	0,0459	0,2093	0	1
	Unemployed	2 711	0,0067	0,0814	0	1
Food Expenditure Shock	Employed	2 093	0,0651	0,2468	0	1
	Unemployed	2 678	0,0971	0,2961	0	1
Health Shock	Employed	2 101	0,0166	0,1277	0	1
	Unemployed	2 711	0,0279	0,1648	0	1
Medical Aid Coverage	Employed	2 101	0,2406	0,4275	0	1
	Unemployed	2 711	0,0557	0,2294	0	1
Pension Recipient	Employed	2 101	0,0157	0,1245	0	1
	Unemployed	2 711	0,2907	0,4542	0	1

Table A.3: Panel Logit Model Outputs

	Model I-I Quintile 2	Model I-II Employed	Model I-III Quintile 2 and Employed
Health shock	1.1417* (0.6860)	1.0121* (0.5418)	1.5486** (0.7418)
Medical Aid coverage	0.5298 (0.3344)	0.3186 (0.3110)	0.6524* (0.3747)
Female	0.1554 (0.3102)	-0.0054 (0.2524)	0.1433 (0.3640)
Married	-0.5241 (0.3981)	-0.3588 (0.2951)	-0.9037* (0.4830)
Age			
30-49	-0.3172 (0.4387)	-0.1798 (0.3613)	-0.3112 (0.4696)
50-64	-0.4841 (0.5140)	-0.7227* (0.4299)	-1.1263* (0.6045)
65-91	0.5655 (0.6997)	0.2024 (0.9325)	0.2647 (1.1263)
Education Attainment			
Primary	-0.0072 (0.6410)	0.3519 (0.5080)	1.2250 (1.0604)
Secondary	-1.2086* (0.6558)	-0.4834 (0.5162)	-0.1397 (1.0413)
Certificate	-1.2101* (0.7175)	-0.5367 (0.5758)	0.0210 (1.0776)
Undergraduate	-0.5511 (0.7062)	-0.0611 (0.5914)	0.3762 (1.0757)
Postgraduate	-0.3062 (0.9579)	0.1289 (0.8662)	0.9544 (1.2715)
Household size			
4-6	0.5237 (0.3562)	0.5970** (0.2575)	0.5838 (0.4234)
7-9	1.1033	1.0817***	1.8019*

	(0.7600)	(0.4071)	(0.9723)
10-24	2.0339	1.7362***	3.1804*
	(1.5962)	(0.6737)	(1.8275)
Pension recipient	-1.5749**	-0.6546	-0.0922
	(0.6772)	(0.8389)	(1.2845)
Employed	0.4477	-	-
	(0.3916)		
Urban	0.1932	0.3410	0.5969
	(0.4099)	(0.3291)	(0.4901)
Toilet	-0.3206	-0.2980	-0.3966
	(0.4018)	(0.3080)	(0.4716)
Piped Water	0.4760	-0.0561	0.1822
	(0.6715)	(0.4209)	(0.8067)
Constant	-3.8602***	-3.3886***	-4.4405***
	(1.0711)	(0.7163)	(1.3978)
Observations	1 801	2 225	1 275

Note: *** Significant at 1 percent level, ** significant at the 5 percent level, * Significant at the 10* level

Table A.4: Parallel Trend Assumption test

	Whole sample	Whole sample with categories
Post	0.4647**	0.4475**
	(0.1824)	(0.1797)
Period 1*Treat	-1.1620	-1.2879
	(1.0554)	(1.0731)
Period 3*Treat	1.9818**	1.8711**
	(0.9432)	(0.9469)
Treat	-1.0316	-1.0633
	(0.6696)	(0.6763)
Medical Aid coverage	-0.2727	-0.4333
	(0.3212)	(0.3152)
Female	-0.2823	-0.3524**
	(0.1868)	(0.1785)
Married	0.1584	0.1315
	(0.2336)	(0.2350)
Age	-0.0059	-
	(0.0104)	
30-49		0.0529
		(0.3383)
50-64		-0.1956
		(0.3848)
65-91		0.4343
		(0.5811)
Education Attainment	-0.0138	-
	(0.0213)	

Primary		0.4557*
		(0.2557)
Secondary		-0.1051
		(0.3054)
Certificate		-0.2333
		(0.4082)
Undergraduate		0.4736
		(0.4234)
Postgraduate		-1.2028
		(0.8319)
Household size	0.1082***	-
	(0.0259)	
4-6		0.6898***
		(0.1909)
7-9		0.6027**
		(0.2397)
10-24		0.8829***
		(0.3131)
Pension recipient	-0.2513	-0.6967*
	(0.2818)	(0.3759)
Urban	0.1656	0.1353
	(0.1685)	(0.1688)
Employed	-0.3202	-0.2981
	(0.1972)	(0.2010)
Constant	-2.0806***	-2.3946***
	(0.5351)	(0.4143)
Observations	4770	4771

Note: *** Significant at 1 percent level, ** significant at the 5 percent level, * Significant at the 10* level

Table A.5: PS-Matching T-Test Results - Logit

Variable	sample	Mean Treated	Control	% Bias	|%| Bias Reduction	t-test t	p>|t|
HH logged Income	Unmatched	6,8815	6,9697	-8,5		-0,36	0,965
	Matched	6,8612	6,8556	0,5	94.1	0,04	0,965
Logged Food Exp	Unmatched	5,4468	5,4687	-2,6		-0,11	1
	Matched	5,4102	5,4102	0	100	0	1
Medical Aid Coverage	Unmatched	0,05	0,05	0		0	0,584
	Matched	0,06612	0,04959	5,9	-	0,55	0,584
African	Unmatched	0,9	0,9	0		0	0,352
	Matched	0,90083	0,93388	-10,2	-	-0,93	0,352
Age 2	Unmatched	3246,8	3096,3	9,6		0,43	0,482

	Matched	3222,6	3069,1	9,7	1.0	0,7	0,482
Illness	Unmatched	0,25	0,2	11,5		0,53	0,671
	Matched	0,29752	0,27273	5,6	51.3	0,43	0,671
HH Food Expenditure	Unmatched	362,28	388,52	-5,6		-0,19	0,833
	Matched	323,61	313,46	2,4	57.1	0,21	0,833
Female	Unmatched	0,65	0,575	15,7		0,68	0,429
	Matched	0,64463	0,59504	10,4	33.8	0,79	0,429
Mean Bias	Unmatched	6.7					
	Matched	5.6					
*B>25%	Unmatched	24.2					
	Matched	23.3					
* R outside [0.5; 2]	Unmatched	0.93					
	Matched	0.58					

Table A.6: Mhbounds output for Sensitivity Analysis for Average Treatment Effects

Gamma	Qmh+	Qmh-	pmh+	pmh-
1	-0,421815	-0,421815	0,66342	0,66342
2	0,384674	0,384674	0,350239	0,350239
3	0,870726	0,870726	0,191952	0,191952
4	1,23334	1,23334	0,108725	0,108725
5	1,52997	1,52997	0,063012	0,063012
6	1,78513	1,78513	0,03712	0,03712
7	2,01158	2,01158	0,022132	0,022132
8	2,21681	2,21681	0,013318	0,013318

Gamma : odds of differential assignment due to unobserved factors
Qmh+ : Mantel-Haenszel statistic (assumption: overestimation of treatment effect)
Qmh- : Mantel-Haenszel statistic (assumption: underestimation of treatment effect)
pmh+ : significance level (assumption: overestimation of treatment effect)
pmh- : significance level (assumption: underestimation of treatment effect)