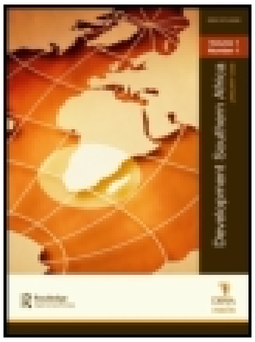

Development Southern Africa

ISSN: 0376-835X (Print) 1470-3637 (Online) Journal homepage: https://www.tandfonline.com/loi/cdsa20

# Lifestyle risk factors, non-communicable diseases and labour force participation in South Africa

**Nozuko Lawana, Frederik Booysen, Asrat Tsegaye, Forget Mingiri Kapingura & Charles Hongoro**

DEVELOPMENT SOUTHERN AFRICA
https://doi.org/10.1080/0376835X.2019.1678459

Routledge
Taylor & Francis Group

Check for updates

# Lifestyle risk factors, non-communicable diseases and labour force participation in South Africa

Nozuko Lawana [a,b], Frederik Booysen [c], Asrat Tsegaye[b], Forget Mingiri Kapingura [b] and Charles Hongoro [a]

[a]Research use and Impact Assessment, Human Sciences Research Council, Pretoria, South Africa; [b]Department of Economics, University of Fort Hare, East London, South Africa; [c]School of Economics and Business Sciences, University of Witwatersrand, Johannesburg, South Africa

**ABSTRACT**

This paper investigates the indirect effects of lifestyle risk factors associated with non-communicable diseases on labour force participation in South Africa utilising data from the National Income Dynamics Study. Endogenous multivariate probit models with a recursive simultaneous structure were employed in the study as a method of analysis. Findings showed a negative effect of non-communicable diseases on labour force participation. When the analysis was disaggregated by gender, the results showed that the effect of stroke and heart diseases were only significant for men, while diabetes and hypertension were only significant for women. The results also emphasised the significant indirect impact of obesity, physical activity and alcohol consumption on labour force participation through non-communicable diseases, especially for men. The policy implications of this study are thus gender-specific. These results can be used to inform the South African National Department of Health to strengthen current health strategies with the aim of reducing lifestyle risk factors and thus promoting sustained labour force participation rates in South Africa.

**KEYWORDS**

Labour force participation; lifestyle risk factors; non-communicable diseases; multivariate probit

**JEL CLASSIFICATION**

I1; J01

## 1. Background

Non-communicable diseases (NCDs) are the leading cause of death worldwide, accounting for 70% of deaths every year (World Health Organisation, 2015). About 70–80% of these deaths occur in low and middle income countries (Abegunde et al., 2007; Miranda et al., 2008; World Health Organisation, 2015). The 2010 report on the global status of non-communicable diseases indicates that the prevalence of NCDs is rapidly rising in the African region and it is estimated that NCDs will be the most common cause of death by 2030 (World Health Organisation, 2010). South Africa's Strategic Plan for the Prevention and Control of Non-Communicable Diseases (2013–2017) reported that the country is facing a multiple burden of disease, including both communicable and non-communicable diseases. The major NCDs include diabetes,

**CONTACT** Nozuko Lawana nlawana@hsrc.ac.za Human Science Research Council, 134 Pretorious Street, Pretoria, South Africa

cardiovascular diseases, cancer, asthma, and mental disorders. Much of previous literature has documented that lifestyle risk factors are associated with the risk of developing NCDs (Zhang et al., 2009; Khuwaja et al., 2011; van Zyl et al., 2012). Major risk factors that are attributed to the development of NCDs include smoking, unhealthy diet, obesity, physical inactivity and the harmful use of alcohol (van Zyl et al., 2012). Most NCDs are interrelated through common risk factors. Obesity, for example, is associated with hypertension and increased cardiovascular disease (Narkiewicz, 2006). The burden of NCD risk factors is high in South Africa. According to the 2013–17 Strategic Plan for the Prevention of Non-Communicable Diseases, 51.1% of individuals were physically inactive, 31.1% were obese, and 14.1% smoked tobacco.

The prevalence of NCDs impacts negatively on economic development through loss of productivity and income. Labour force participation is an important factor that determines the productivity and economic growth of a country (Shahid, 2014). A large share of the South African population derives income from the labour force (salaries and wages). The country's three largest industries that have high employment rates and contributions to projected economic growth, which includes finance, general government and wholesale and retail trade. Therefore, the possible negative impact of non-communicable diseases on the labour market has implications for poverty. South Africa is among the richest countries in Africa, however, it has been experiencing a decline in labour force participation rates (Nwosu & Woolard, 2017).

Despite the burden of NCDs and its impact on productivity, the study of the factors associated with the negative impact of NCDs on labour force participation is surprisingly neglected in developing countries and in South Africa. Numerous international studies conducted in developed countries have found that chronic diseases impact negatively on labour force participation (Harris, 2009; Zhang et al., 2009; Renna & Thakur, 2010; Schofield et al., 2015, 2008).

Zhang et al. (2009) employed a multivariate probit model to investigate the impact of various NCDs (diabetes, cardiovascular disease, mental illnesses and other diseases) on the probability of labour force participation, using data from the 2005 Australian National Health Surveys. The estimations were separated by gender and age group and the results showed a significant negative effect of diabetes on labour force participation for men only and a significant effect of cardiovascular disease for men aged 50- 64 years. Harris (2009) estimated simultaneous equations from the multivariate probit model to analyse the influence of two major NCDs (diabetes and cardiovascular disease) on labour supply. The data used was extracted from the AusDiab dataset on the residential population of Australia aged ≥ 25 and the author found that diabetes and cardiovascular disease have a strong negative impact on labour force participation. Moreover, the estimations showed a negative and significant indirect effect on labour force participation of lifestyle risk factors like obesity, insufficient exercise, lipid abnormality and smoking.

These studies have been conducted in developed countries. Therefore, replication of a similar study in a developing country context is important, particularly insofar as burdens of NCDs in the developing world is on the increase. This study adds to the literature by estimating the effects of cardiovascular disease (stroke, heart diseases, and hypertension) and diabetes on labour force participation.

Moreover, this study contributes to the literature on the determinants of labour force participation in various respects. Firstly, the study considers the joint effect of non-communicable diseases (stroke, diabetes, heart disease and hypertension) on labour force participation to see whether the four diseases affect labour force participation differently, accounting for their correlation through unobservable factors such as allergies and genetics. Secondly, this study determines the indirect effects of lifestyle risk factors on labour force participation through non-communicable diseases. Evidence of this nature can provide useful information into how health-harming behaviours impact on non-communicable diseases and thereby influences the decision to participate in the labour force. Given the different gender-specific dynamics underlying health behaviours and outcomes as well as labour force participation, the analysis is disaggregated by gender to test for gender-based differences in the effect of NCDs on labour force participation in South Africa.

## 2. Method

### 2.1. Data

The dataset used for the purposes of this research is the National Income Dynamics Study (NIDS). It makes it possible to do research on NCDs and labour market outcomes as it is designed to elicit information on individual and household levels covering various factors influencing labour market decisions. This includes the self-reported prevalence of non-communicable diseases and lifestyle risk factors as well as demographics and other economic characteristics of study participants.

This study employs the pooled cross-sectional data from four waves of NIDS. The surveys use a two-stage stratified cluster sampling design. In the first stage, 400 primary sampling units in 53 districts of all 9 South African provinces were selected from Statistics South Africa's 2003 master sample of 3,000 PSUs. Private households were randomly selected within each sampled PSU and individuals from the selected households were then interviewed. The sample of the study is restricted to the working age population between 20 and 64 years at the date of interview. Although the working age population is defined by the International Labour Organisation as individuals aged 15–64 years, this study excludes younger respondents who were mainly students at the time of analysis (Nwosu & Woolard, 2017). Furthermore, Asians and Indians were excluded from the study due to their small sample size.

The binary variable of labour force participation is generated from the employment status information obtained in the surveys. In line with the broad definition of unemployment, labour force participants are defined as individuals, who at the time of the interview, reported working for pay, being unemployed but looking for a job, and discouraged from seeking work. Dummy variables for non-communicable diseases are constructed based on the question of chronic health conditions.

Definitions and descriptive statistics of the variables of interest are presented in Table 1. A larger proportion of males than females reported to be participating in the labour force. Of the female population, only 63.6% reported being in the labour force, compared to 78.4% for the male population.

NCDs are more prevalent amongst females than males, about 4.0% of females aged 20 years and above reported having been diagnosed with diabetes, 2.6% diagnosed with heart diseases, 16.5% with hypertension and 1.0% reported having been diagnosed with stroke.

**Table 1.** Definitions and descriptive statistics.

| Variables | Definition | Total Mean | Total SD | Male Mean | Male SD | Female Mean | Female SD |
|---|---|---|---|---|---|---|---|
| Labour force | Employed or unemployed but not retired or discouraged work seekers | 0.704 | 0.457 | 0.784 | 0.412 | 0.636 | 0.481 |
| **NCDs** | | | | | | | |
| Stroke | Diagnosis of stroke | 0.008 | 0.087 | 0.005 | 0.073 | 0.010 | 0.010 |
| Diabetes | Diagnosis of diabetes | 0.035 | 0.186 | 0.031 | 0.173 | 0.040 | 0.196 |
| Heart disease | Diagnosis of heart disease | 0.020 | 0.140 | 0.013 | 0.112 | 0.026 | 0.159 |
| Hypertension | Diagnosis of high blood pressure | 0.126 | 0.332 | 0.081 | 0.274 | 0.165 | 0.371 |
| Married | Married or living with partner | 0.344 | 0.475 | 0.336 | 0.472 | 0.351 | 0.477 |
| Age | between 20 and 64 years old | 37.058 | 12.141 | 36.685 | 12.014 | 37.389 | 12.242 |
| **Education** | | | | | | | |
| No education | No educational attainment | 0.056 | 0.231 | 0.049 | 0.217 | 0.063 | 0.242 |
| Primary | Primary completed | 0.171 | 0.376 | 0.178 | 0.382 | 0.165 | 0.371 |
| Secondary | not completed secondary school | 0.414 | 0.493 | 0.413 | 0.492 | 0.415 | 0.493 |
| Matric | secondary school completed | 0.194 | 0.396 | 0.202 | 0.402 | 0.187 | 0.390 |
| Tertiary | University diploma or degree completed | 0.165 | 0.371 | 0.158 | 0.365 | 0.170 | 0.376 |
| **Race** | | | | | | | |
| African | African population | 0.806 | 0.395 | 0.812 | 0.391 | 0.802 | 0.399 |
| Coloureds | Coloureds population | 0.097 | 0.296 | 0.094 | 0.292 | 0.099 | 0.299 |
| White | White population | 0.097 | 0.295 | 0.094 | 0.292 | 0.099 | 0.299 |
| **Area** | | | | | | | |
| Rural formal | Living in rural area | 0.066 | 0.248 | 0.071 | 0.256 | 0.062 | 0.241 |
| Traditional authority | Living in farm or traditional area | 0.273 | 0.446 | 0.243 | 0.429 | 0.300 | 0.458 |
| Urban formal | Living in urban formal area | 0.559 | 0.497 | 0.581 | 0.493 | 0.539 | 0.499 |
| Urban informal | Living in urban informal | 0.102 | 0.303 | 0.105 | 0.307 | 0.099 | 0.299 |
| Household size | Number of individuals in the household | 4.635 | 3.301 | 4.050 | 3.222 | 5.152 | 3.284 |
| **Risk factors** | | | | | | | |
| Obese | BMI is more than or equal to 30 | 0.394 | 0.489 | 0.332 | 0.471 | 0.45 | 0.497 |
| Exercise | (regular physical activity) | 0.121 | 0.326 | 0.177 | 0.382 | 0.074 | 0.261 |
| Smoking | smoke regularly | 0.052 | 0.221 | 0.092 | 0.289 | 0.029 | 0.167 |
| Drinking | consume alcohol regularly | 0.036 | 0.186 | 0.063 | 0.243 | 0.013 | 0.114 |

Note: Population weighted mean adjusted for national representation using post stratified weights.

The corresponding figures for males are 3.1%, 1.3%, 8.1% and 0.5% respectively. Hypertension is the most prevalent NCD in the country, with 12.6% of the population reporting hypertension.

The majority of the population are Africans (81.2% and 80.2% for males and females respectively). A large fraction of the population reported not completing secondary education (41.1%) and only an average of 16.5% reported having a bachelor degree or higher degree qualification. The results illustrate a low prevalence of regular alcohol use (3.6%) and smoking (5.2%) in the adult population, whereas the prevalence of obesity is high at 39.4% (33.3% for men and 45.0% for women).

### *2.2. Regression models*

The study began by estimating five separate univariate equations of labour force participation, one each for diabetes, stroke, heart diseases, and hypertension. Based on the theory that health is endogenous and is influenced by lifestyle risk factors, endogeneity bias can arise in estimating the effect of NCDs on labour force participation (Cai & Kalb, 2006). Given the likely presence of endogeneity and unobserved heterogeneity in NCDs, many empirical studies have relied on single equation instrumental variable methods such as 2SLS and GMM to address the potential endogenous bias (Bastida &

Pagán, 2002; Norton & Han, 2008; Cai & Cong, 2009; Minor, 2011). This study estimated five recursive simultaneous equations from the multivariate probit model with simulated maximum likelihood to control for the resultant potential bias of the estimates of the effects of lifestyle risk factors and NCDs on labour force participation.

The multivariate probit model with simulated maximum likelihood is an appropriate model for controlling the likely bias due to potential endogeneity associated with unobserved heterogeneity. The empirical specifications for the multivariate regression models was informed by conceptual frameworks and overviews published in the relevant literature (Cappellari & Jenkins, 2003; Jones, 2007; Harris, 2009). These models are estimated as follows:

$$L^* = X_L'\beta_L + \gamma_L' C_s + \gamma_L' C_h + \gamma_L' C_b + \gamma_L' C_d + \varepsilon_L \quad L = 1 \text{ if } L^* > 0\text{ , } 0 \text{ otherwise} \quad (1)$$

$$C_s^* = \beta_s' C_d^* + \beta_s' C_h^* + \beta_s' C_b^* + X_s'\beta_s + \varepsilon_s \quad Cs = 1 \text{ if } C_s^* > 0\text{ , } 0 \text{ otherwise} \quad (2)$$

$$C_h^* = \beta_h' C_d^* + \beta_h' C_b^* + X_h'\beta_h + \varepsilon_h \quad C_h = 1 \text{ if } C_h^* > 0\text{ , } 0 \text{ otherwise} \quad (3)$$

$$C_d^* = \beta_d' C_b^* + X_d'\beta_d + \varepsilon_d \quad C_d = 1 \text{ if } C_d^* > 0\text{ , } 0 \text{ otherwise} \quad (4)$$

$$C_b^* = X_b'\beta_b + \varepsilon_b \quad C_b = 1 \text{ if } C_b^* > 0\text{ , } 0 \text{ otherwise} \quad (5)$$

Where $L^*$ is the latent labour force participation outcome and $X_L$ is the vector of exogenous variables that affect labour force participation, including demographic variables, education, geographical area and other variables. C comprises of four dummy variables ($C_s$, $C_h$, $C_d$,and $C_b$), indicating the diagnoses of stroke, heart diseases, diabetes and hypertension respectively and ε is the error term. The latent variable $L^*$ is associated with the observed binary variable L, which takes on a value of 1 if the individual is in the labour force and 0 otherwise.

Equations 2–5 are estimated to determine the effects of lifestyle risk factors on non-communicable diseases and the indirect effects of risk factors on labour force participation. Diabetes, heart diseases and hypertension are the regressors in the stroke equation and diabetes and hypertension are the regressors in the heart disease equation. Other lifestyle risk factors included in equations 2–4 were physical activity, regular smoking, regular alcohol consumption, and age. These lifestyle variables are denoted by $X\beta_s$ in non-communicable diseases equations. The study used the simultaneous recursive approach to obtain the consistent estimators of the parameters of reduced form equation of non-communicable diseases and the structural labour force participation model. The consistent estimates are achieved by assuming that the error terms in equation 1–5 have a multivariate normal distribution with mean zero and covariance matrix Σ: ($\varepsilon_L$, $\varepsilon_s$, $\varepsilon_h$, $\varepsilon_d$, $\varepsilon_b$,) ~ MVN (0, Σ) with the five variances assumed equal to 1.

To set the scene, the association between NCDs and labour force participation is investigated using a univariate probit model. A univariate approach does not account for the correlation between the error terms and ignoring the potentially non-zero off-diagonal elements in Σ will result in biased coefficient estimates where correlations across the error term exist (Maddala, 1983)

The use of the covariance matrix for identification purposes was suggested in the literature (Chib & Greenberg, 1998). The logical consistency and identification of recursive simultaneous probit models based on the exclusion restrictions (Maddala, 1983; Balia & Jones, 2008). Maddala (1983) claimed the necessity of exclusion variable for identification and argued that for a bivariate probit model with endogenous variables appearing on the right hand side of the equation, a recursive structure is required to satisfy the logical consistency condition.

The identification requires the exclusion restriction that the probit equation without an endogenous dummy variable on the right hand side contains extra exogenous variables that are not in the probit equation with endogenous dummy on the right hand side (Maddala, 1983). In particular, the recursive structural model without exclusion restriction leads to failure of identification. On the other hand, in contrast to Maddala's argument, Wilde argues that identification is achieved even if the same exogenous variables appear in both equation so that theoretical identification does not require the availability of any additional instruments (Wilde, 2000). This study followed the exclusion restriction approach, lifestyle risk factors contained in NCD equations were excluded from the labour force participation equation because it is assumed that they indirectly influence the probability of participating in the labour market. Therefore, the empirical specification of this study satisfies the condition for a multivariate probit model. The variance covariance matrix of the cross-equation error terms and Wald test were estimated in the study to test the null hypothesis of exogeneity. If the null hypothesis from the multivariate probit model cannot be rejected, the univariate model for labour force participation and non-communicable disease equations can be estimated separately. Marginal effects on labour force participation of all endogenous and exogenous variables were estimated.

## 3. Empirical results

### *3.1. Univariate probit regression analysis*

Tables 2 and 3 show the marginal effects from the probit regressions of four types of NCDs (stroke, heart diseases, diabetes, and hypertension) on labour force participation and the marginal effects of lifestyle risk factors on the prevalence of NCDs. All the NCDs have negative associations with labour force participation, however, only stroke is statistically significant. Table 2 shows the results for the total population aged 20–64 years. Having had a stroke has a large negative effect compared to other NCDs and reduces the probability of participating in the labour force by as much as 17.3%.

Education, age, and sex have the expected signs and are statistically significant. In terms of education level, the higher the level of education obtained, the higher the probability of being in the labour force. In other words, the probability of participating in the labour force increases with the level of education. For example, the predicted probability of labour force participation for those with a matric level of education is 16.5% compared to those with no education, and the probability of participating in the labour force for individuals with a tertiary qualification is 31.1%.

Turning to the NCD models (columns II-V), the results as expected indicate that NCDs are interrelated. The probability of having a stroke for an individual with heart disease is 1.4% and the probability of having heart disease for an individual with diabetes is 1.3%.

**Table 2.** Probit model of labour force participation and NCDs.

| | Labour force (I) | Stroke (II) | Heart diseases (III) | Diabetes (IV) | Hypertension (V) |
|---|---|---|---|---|---|
| Stroke | −0.173*** | | | | |
| Heart disease | −0.036 | 0.014*** | | | |
| Diabetes | −0.003 | 0.003 | 0.013*** | | |
| Hypertension | −0.006 | 0.011*** | 0.025*** | 0.065*** | |
| Age | 0.326*** | 0.000*** | 0.001*** | 0.002*** | 0.008*** |
| Age squared | −0.418*** | | | | |
| Female | ref | | | | |
| Men | 0.133*** | | | | |
| Not married | ref | | | | |
| Married | 0.002 | | | | |
| No education | ref | | | | |
| Primary | 0.063*** | | | | |
| Secondary | 0.118*** | | | | |
| Matric | 0.165*** | | | | |
| Tertiary | 0.311*** | | | | |
| Traditional area | ref | | | | |
| Rural | 0.113*** | | | | |
| Urban-formal | 0.097*** | | | | |
| Urban-informal | 0.122*** | | | | |
| White | ref | | | | |
| Coloured | 0.026 | | | | |
| African | 0.019 | | | | |
| Household size | −0.008*** | | | | |
| Wave 1 | ref | | | | |
| Wave2 | −0.129*** | −0.003 | −0.014*** | 0.002 | −0.040*** |
| Wave3 | −0.073*** | −0.002 | −0.012*** | 0.005 | 0.010** |
| Wave4 | −0.055*** | −0.002 | −0.014*** | 0.005 | −0.023*** |
| Obese | | 0.000 | 0.003 | 0.016*** | 0.075*** |
| Exercise | | −0.006** | 0.006* | −0.001 | −0.008 |
| Smoking | | −0.001 | 0.003 | 0.011 | 0.014* |
| Drinking | | −0.014*** | −0.030*** | −0.009 | −0.016 |
| Constant | −1.096*** | −0.055*** | −0.112*** | −0.237*** | −0.568*** |
| Wald $\chi^2$ | 2900.43 | 272.75 | 421.26 | 950.18 | 5070.38 |
| Pseudo $R^2$ | 0.145 | 0.121 | 0.1226 | 0.2301 | 0.1906 |
| *P*-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| N | 46 477 | 40 069 | 40 137 | 40 286 | 40 519 |

Notes: (***), (**) and (*) indicate 1%, 5% and 10% significant, respectively.

Hypertension has a positive effect for all the NCDs and the effects are statistically significant. Particularly, the marginal effect of being diagnosed with hypertension on having a stroke is 1.1%, heart disease is 2.5% and diabetes is 6.5%.

Being physically active reduces the risk of a stroke by 0.6%. Obesity is strongly statistically associated with diabetes and hypertension whereas regular alcohol consumption is strongly associated with stroke and heart disease. Albeit, the results show unexpected signs for the effect of the regular consumption of alcohol on NCDs. The signs are negative, suggesting that regular alcohol consumption reduces the risk of stroke and heart disease by 1.4% and 3.0% respectively.

Table 3 shows the disaggregated results by gender. The results of the models I and VI indicate that there is a significant negative association between stroke and labour force participation and the effect of stroke is much greater for men compared to women. The association between heart diseases and labour force participation is not significant for males. In terms of the comparative effect of the NCDs on other diseases, one can see that the effect of diabetes on stroke and heart diseases for men is not significant while

**Table 3.** Marginal effects from probit models of labour force and NCD by gender.

| | Males | | | | | Females | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Labour force (I) | Stroke (II) | Heart diseases (III) | Diabetes (IV) | Hypertension (V) | Labour force (VI) | Stroke (VII) | Heart diseases (VIII) | Diabetes (IX) | Hypertension (X) |
| Stroke | −0.256*** | | | | | −0.130*** | | | | |
| Heart disease | −0.113** | 0.012*** | | | | −0.004 | 0.014*** | | | |
| Diabetes | 0.040 | −0.010 | −0.007 | | | −0.031 | 0.005* | 0.024*** | | |
| Hypertension | −0.001 | 0.007*** | 0.018*** | 0.059*** | 0.005*** | 0.003 | 0.012*** | 0.026*** | 0.069*** | |
| Age | 0.253*** | 0.000 | 0.001*** | 0.002*** | 0.008*** | 0.378*** | 0.001*** | 0.001*** | 0.002*** | 0.008*** |
| Age squared | −0.539*** | | | | | −0.351*** | | | | |
| Not married | ref | | | | | ref | | | | |
| Married | 0.095*** | | | | | −0.054*** | | | | |
| No education | ref | | | | | ref | | | | |
| Primary | 0.063*** | | | | | 0.057*** | | | | |
| Secondary | 0.088*** | | | | | 0.137*** | | | | |
| Matric | 0.113*** | | | | | 0.200*** | | | | |
| Tertiary | 0.205*** | | | | | 0.380*** | | | | |
| Traditional | ref | | | | | ref | | | | |
| Rural | 0.149*** | | | | | 0.074*** | | | | |
| Urban-formal | 0.088*** | | | | | 0.094*** | | | | |
| urban-informal | 0.095*** | | | | | 0.139*** | | | | |
| White | ref | | | | | ref | | | | |
| Coloured | −0.083** | | | | | 0.076** | | | | |
| African | −0.090*** | | | | | 0.062** | | | | |
| Household size | −0.012*** | | | | | −0.006*** | | | | |
| Wave 1 | ref | | | | | ref | | | | |
| Wave2 | −0.120*** | −0.007** | −0.008* | −0.001 | −0.013 | −0.134*** | 0.000 | −0.017*** | 0.003 | −0.039*** |
| Wave3 | −0.065*** | −0.005* | −0.003 | 0.001 | 0.024** | −0.079*** | 0.002 | −0.016*** | 0.006 | 0.009 |
| Wave4 | −0.045*** | −0.007*** | −0.013*** | 0.002 | 0.010 | −0.062*** | 0.002 | −0.014*** | 0.005 | −0.029*** |
| Obese | | 0.001 | 0.003 | 0.015** | 0.053*** | | 0.000 | 0.000 | 0.018*** | 0.080*** |
| Exercise | | −0.004** | 0.012*** | −0.008 | −0.019* | | −0.002 | 0.000 | 0.004 | 0.000 |
| Smoking | | −0.001 | 0.002 | 0.003 | 0.025** | | −0.001 | 0.008 | 0.012 | 0.037 |
| Drinking | | −0.033** | −0.078 | −0.008 | 0.024 | | | −0.030** | −0.025 | 0.072 |
| Constant | −0.746*** | −0.035*** | −0.065*** | −0.219*** | −0.418*** | −1.183*** | −0.066*** | −0.128*** | −0.235*** | −0.610*** |
| Wald $\chi^2$ | 1382.64 | 151.46 | 107.71 | 385.61 | 415.15 | 1676.12 | 192.64 | 330.42 | 661.29 | 1272.45 |
| Pseudo $R^2$ | 0.1524 | 0.1852 | 0.1722 | 0.2584 | 0.2001 | 0.1285 | 0.1078 | 0.1153 | 0.2219 | 0.1651 |
| *P*-value | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| Sample size | 18 825 | 12 526 | 12 542 | 12 579 | 12 652 | 27 652 | 27 352 | 27 592 | 27 704 | 27 864 |

Notes: (***), (**) and (*) indicate 1%, 5% and 10% significant, respectively.

the effect is positive and significant for women. The size of the effect of heart disease on stroke is approximately the same for men (1.20%) and women (1.40%).

### 3.2. Multivariate probit results

Table 4 presents the estimated correlation coefficients from the multivariate probit models between the errors terms of five equations (labour force participation, stroke, heart disease, diabetes, and hypertension).

The correlation coefficients of NCDs with the labour force participation equation, reported in column I of Table 4, are all individually statistically significant. The correlation coefficients of NCDs are statistically significant for all equations of NCDs, which implies that non-communicable diseases influenced each other. The correlations between labour force and NCDs for the subsample of men and women are also significant except for diabetes and BP for men and stroke for women. These results suggest the joint estimation of labour force participation and NCDs equations simultaneously. The likelihood ratio test for the hypothesis that non communicable diseases and labour force participation are independent is rejected at the 1% confidence level (see Appendix A1).

Table 5 presents the aggregated direct, indirect and total marginal effects on labour force participation of all exogenous variables, including non-communicable diseases, estimated using recursive multivariate probit models.

The total marginal effects are the sum of the direct and indirect partial effects. The equations are statistically significant overall on Wald's chi-squared test ($\rho = 0.000$). These results suggest that estimating the effect of NCDs on labour force participation using a univariate or standard probit equation could produce biased results. Once unobserved heterogeneity is accounted for in multivariate probit analysis, the estimation results

**Table 4.** Correlation coefficients from multivariate probit model.

| | Total (I) | Males (II) | Females (III) |
|---|---|---|---|
| ρlabour force, stroke | 0.247** | 0.771*** | 0.212 |
| | (0.102) | (0.147) | (0.208) |
| Plabour force, heart | 0.291*** | 0.365** | 0.306** |
| | (0.087) | (0.145) | (0.140) |
| ρlabour force, diabetes | 0.220** | −0.027 | 0.264*** |
| | (0.091) | (0.142) | (0.074) |
| ρlabour force, hypertension | 0.219*** | 0.161 | 0.335*** |
| | (0.079) | (0.102) | (0.084) |
| ρstroke, heart | 0.254*** | 0.306** | 0.402*** |
| | (0.045) | (0.133) | (0.060) |
| ρstroke, diabetes | 0.099** | −0.027 | 0.263*** |
| | (0.049) | (0.127) | (0.062) |
| ρstroke, hypertension | 0.138*** | 0.167 | 0.323*** |
| | (0.036) | (0.108) | (0.043) |
| ρheart, diabetes | 0.152*** | 0.126* | 0.308*** |
| | (0.042) | (0.076) | (0.045) |
| ρheart, hypertension | 0.191*** | 0.344*** | 0.318*** |
| | (0.031) | (0.080) | (0.030) |
| ρdiabetes, hypertension | 0.370*** | 0.558*** | 0.507*** |
| | (0.024) | (0.050) | (0.029) |

Notes: standard errors in brackets. (***), (**) and (*) indicate 1%, 5% and 10% significant, respectively.

**Table 5.** Multivariate probit model of labour force participation.

| | | Indirect effects from NCDs | | | | |
|---|---|---|---|---|---|---|
| | Direct effects (I) | Stroke (II) | heart disease (III) | Diabetes (IV) | Hypertension (V) | Total effect (VI) |
| Stroke | | | | | | −0.319*** |
| Heart disease | | | | | −0.240*** | |
| Diabetes | | | | | | −0.117** |
| Hypertension | | | | | | −0.116*** |
| Female | ref | | | | | ref |
| Male | 0.139*** | | | | | 0.139*** |
| Not married | ref | | | | | ref |
| Married | −0.006 | | | | | −0.006 |
| Wave 1 | ref | ref | ref | ref | ref | ref |
| Wave 2 | −0.141*** | 0.005** | 0.017*** | 0.003 | 0.033*** | −0.088*** |
| Wave 3 | −0.066*** | 0.002 | 0.012*** | −0.009 | −0.012* | −0.073** |
| Wave 4 | −0.066*** | 0.000 | 0.016*** | 0.003 | 0.023** | −0.024** |
| Age | 0.334*** | −0.001*** | −0.001*** | −0.003*** | −0.008*** | 0.321*** |
| Age squared | −0.416*** | | | | | −0.416*** |
| Not educated | ref | | | | | ref |
| Primary | 0.096*** | | | | | 0.096*** |
| Secondary | 0.158*** | | | | | 0.158*** |
| Matric | 0.226*** | | | | | 0.226*** |
| Tertiary | 0.400*** | | | | | 0.400*** |
| Traditional area | ref | | | | | ref |
| Urban-informal | 0.158*** | | | | | 0.158*** |
| Urban-formal | 0.115*** | | | | | 0.115*** |
| Rural | 0.118*** | | | | | 0.118*** |
| White | ref | | | | | ref |
| African | 0.017 | | | | | 0.017 |
| Coloured | −0.012 | | | | | −0.012 |
| Household size | −0.007*** | | | | | −0.007*** |
| Obese | | −0.003* | −0.007*** | −0.021*** | −0.076*** | −0.107*** |
| Exercise | 0.006** | −0.008** | 0.005 | 0.020* | 0.024* | |
| Smoking | | −0.001 | −0.001 | −0.011 | −0.029** | −0.041 |
| Drinking | | 0.018*** | 0.029*** | 0.010 | 0.007 | 0.063** |
| Constant | −1.135*** | 0.070*** | 0.127*** | 0.261*** | 0.554*** | −0.123*** |

Notes: (***), (**) and (*) indicate 1%, 5% and 10%significant, respectively.

change compared to those reported in Table 2. The marginal effects of heart diseases, diabetes and hypertension on labour force participation are not significant in the probit model but they became statistically significant in the multivariate probit model.

The negative effect of NCDs on labour force participation are even larger than before. For instance, having had a stroke reduces the probability of labour force participation by 17.3% in the univariate model, whereas in the multivariate probit, the effect of stroke reduces the probability of labour force participation by as much as 31.9%, thus nearly doubling the relative size of the effect. Table 5 also shows the marginal indirect effect of lifestyle risk factors on labour force participation. The total effect of being obese increases the risk of stroke, heart disease, diabetes and hypertension and thereby indirectly reduces the probability of being in the labour force by 10.70%. On aggregate, physical activity indirectly increases the probability of labour force participation by 2.40% through non-communicable diseases. Being a regular drinker increases the probability of working by 6.30% with a strong marginal effect of 2.9% from heart diseases.

Disaggregated multivariate probit results by gender for marginal effects of NCDs and other variables of interest are given in Table 6. For men, the total effect of stroke reduces the labour force participation by 78.0% and heart diseases by 37.4%, while for women the

**Table 6.** Multivariate probit results by gender.

| | Males | | | | | | Females | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Direct effects | Indirect effects | | | | Total marginal effect | Direct effects | Indirect effects | | | | Total marginal effect |
| | | Stroke | heart disease | Diabetes | Hypertension | | | Stroke | heart disease | Diabetes | Hypertension | |
| | (I) | (II) | (III) | (IV) | (V) | (VI) | (VII) | (VIII) | (IX) | (X) | (XI) | (XII) |
| Stroke | | | | | | −0.780*** | | | | | | −0.218 |
| Heart disease | | | | | | −0.374*** | | | | | | −0.181 |
| Diabetes | | | | | | 0.086 | | | | | | −0.138** |
| Hypertension | | | | | | −0.069 | | | | | | −0.170*** |
| Age | 0.272*** | 0.000 | −0.001*** | −0.001*** | −0.005*** | 0.265*** | 0.382*** | −0.001*** | −0.001*** | −0.003*** | −0.009*** | 0.368*** |
| Age squared | −0.538*** | | | | | −0.538*** | −0.351*** | | | | | −0.351*** |
| Not married | ref | | | | | ref | ref | | | | | ref |
| Married | 0.100*** | | | | | 0.100*** | −0.049*** | | | | | −0.049*** |
| Not educated | ref | | | | | ref | ref | | | | | ref |
| Primary | 0.117*** | | | | | 0.117*** | 0.063*** | | | | | 0.063*** |
| Secondary | 0.115*** | | | | | 0.115*** | 0.137*** | | | | | 0.137*** |
| Matric | 0.156*** | | | | | 0.156*** | 0.205*** | | | | | 0.205*** |
| Tertiary | 0.255*** | | | | | 0.255*** | 0.375*** | | | | | 0.375*** |
| Traditional area | ref | | | | | ref | ref | | | | | ref |
| Urban-informal | 0.113*** | | | | | 0.113*** | 0.136*** | | | | | 0.136*** |
| Urban-formal | 0.085*** | | | | | 0.085*** | 0.093*** | | | | | 0.093*** |
| Rural | 0.152*** | | | | | 0.152*** | 0.064*** | | | | | 0.064*** |
| White | ref | | | | | ref | ref | | | | | ref |
| African | −0.066** | | | | | −0.066** | 0.037 | | | | | 0.037 |
| Coloured | −0.096** | | | | | −0.096** | 0.018 | | | | | 0.018 |
| Household size | −0.012*** | | | | | −0.012*** | −0.005*** | | | | | −0.005*** |
| Wave 1 | ref | ref | ref | ref | ref | ref | ref | ref | ref | ref | ref | ref |
| Wave 2 | −0.132*** | 0.006*** | 0.009** | 0.002 | 0.012 | −0.103** | −0.149 | 0.001 | 0.020*** | 0.003 | 0.037*** | −0.088** |
| Wave 3 | −0.049*** | 0.005** | 0.003 | −0.006 | −0.022** | −0.069** | −0.079 | −0.002 | 0.015*** | −0.007 | −0.011 | −0.084* |
| Wave 4 | −0.054*** | 0.008*** | 0.010** | 0.013 | −0.002 | −0.025** | −0.075 | −0.006* | 0.018*** | 0.005 | 0.036*** | 0.022** |
| Obese | | −0.003* | −0.007*** | −0.017*** | −0.051*** | −0.078*** | | −0.002 | −0.004 | −0.022*** | −0.072*** | −0.100** |
| Exercise | | 0.008** | −0.013*** | 0.013* | 0.015*** | 0.023** | | 0.003 | 0.000 | −0.003 | −0.001 | −0.001 |
| Smoking | | 0.001 | −0.004 | −0.013 | −0.039 | −0.037 | | −0.002 | −0.001 | −0.005 | −0.045* | −0.053 |
| Drinking | | 0.007** | −0.019*** | 0.000 | −0.017 | 0.029** | | 0.118*** | 0.026* | 0.048* | −0.002 | 0.190* |
| Constant | −0.917*** | 0.035*** | 0.090*** | 0.232*** | 0.409*** | −0.151*** | −1.121*** | 0.089*** | 0.143*** | 0.273*** | 0.612*** | −0.004*** |

Notes: (***), (**) and (*) indicate 1%, 5% and 10% significant, respectively effect is insignificant and small. Moreover, the effect of diabetes and hypertension on labour force participation for men is insignificant while for women it is negative and significant. The marginal effect of NCDs estimated using the multivariate probit model for both men and women show that being obese increases the probability of having stroke, heart diseases, diabetes and blood pressure and thereby indirectly reduces the probability of participating in the labour force by 7.8% for men and 10.0% for women, with a stronger effect through hypertension than other diseases. The effect on men of physical activity is to decrease the risk of stroke, diabetes and hypertension and thereby increase the probability of labour force participation. The total effect of physical activity on labour force participation for women is negative but insignificant.

## 4. Conclusion

This study investigates whether labour force participation is adversely affected by non-communicable diseases (stroke, heart diseases, diabetes, and hypertension) and indirectly influenced by lifestyle risk factors (smoking, drinking, obesity, and physical activity) through non-communicable diseases. The study finds that the standard econometric approach (univariate probit model) underestimates the marginal effect of non-communicable diseases on labour force participation and of the indirect effect of lifestyle risk factors on labour force participation through non-communicable diseases. The estimated correlations between error terms (unobserved factors) of labour force participation equations from the multivariate probit regressions are all statistically significant, suggesting that the multivariate probit model represents a better model than single probit equations.

The results suggest that the increasing prevalence of non-communicable diseases is likely to reduce labour force participation. The results of this study are in line with the findings of previous studies in developed countries that looked at the effects of lifestyle risk factors and non-communicable diseases on labour force participation using the recursive simultaneous probit method (Harris, 2009; Zhang et al., 2009).

Comparing the multivariate probit results of males and females, the impact of lifestyle risk factors and non-communicable diseases on labour force participation in South Africa differ by gender. The findings show that stroke and heart diseases are negatively and significantly associated with labour force participation of males, whereas the effect is not significant for females. Harris (2009) found a significant negative effect for both males and females. In addition, consistent with Zhang et al. (2009), this study found no significant effect of cardiovascular diseases on the labour force participation of women.

In terms of the effect of diabetes and hypertension on labour force participation by gender, the study found that diabetes and hypertension are negatively associated with labour force participation but the effects are only statistically significant for females. These results conform with the findings of previous studies that investigated the effect of diabetes on female labour force participation (Minor, 2011; Schofield et al., 2014). This contrasts with the findings of Zhang et al. (2009) where the significant effect of diabetes was found only in the case of males.

This study has also demonstrated a lower participation rate associated with lifestyle risk factors. For example, obesity has a negative indirect effect on labour force participation through stroke, diabetes, heart diseases and hypertension. Findings relating to the negative effect of obesity on labour force participation have been explored in numerous studies (Paraponaris et al., 2005; Tunceli et al., 2006; Morris, 2007; Huffman & Rizov, 2014). When estimates were disaggregated by gender, only a significant effect of obesity for men was found. Regular exercise is reported as the best mechanism to reduce the NCDs and indirectly increase labour force participation, though it is only significant for men. The study shows unexpected findings of a positive association between regular alcohol consumption and labour force participation of males and females. The positive impact of alcohol consumption on labour force participation could be explained by the variable used. Alcohol consumption has both a beneficial and detrimental effect on NCDs depending on the volume of alcohol consumed. Light to moderate drinking tends to lower the risk of certain diseases, whilst heavy drinking tends to increase such risks (Zhou et al., 2016). Regular consumption can have different meanings, individuals can regularly drink a small amount of alcohol or a huge amount of alcohol. The NIDS dataset does not have information on the intensity of drinking.

The limitation of this study is that the data on lifestyle factors and NCDs are based on self-report. It has been suggested in literature that self-reported data could be inaccurate as some of the respondents may be untruthful about their diseases due to stigma whilst others may exaggerate their NCDs status to rationalise labour market decisions (Zhang et al., 2009). Therefore, this may result in biased estimates of the role that lifestyle factors and NCDs play in explaining observed differences in labour force participation. Further research should also look into extending the econometric methodology employed in this paper to a panel data application.

In regards to policy implications, the South African National Department of Health has been taking a number of measures to manage and control the increasing burden of lifestyle risk factors and non-communicable diseases. Examples of specific strategies outlined in the Department's Strategic Plan for the Prevention and Control of Non-Communicable Disease include regulations regarding the display of tobacco products at point of sale; smoke free public areas; bans on alcohol advertising and community and social mobilisation to prevent alcohol abuse; a sugar tax; regulations regarding the salt content of food products; and health messaging and warnings on alcohol, tobacco and specific food products. Other relevant policy frameworks include the Strategy for the Prevention and Control of Obesity in South Africa (2015–2020) and the National Health Promotion Policy and Strategy (2015–2019). While general awareness campaigns on how to prevent and control NCDs remains important, a community- rather than facility-based approach is required in promoting and educating communities about healthy lifestyles (Puoane et al., 2017). Men's health programmes addressing healthy lifestyle behaviours are particularly important. This study hence calls for gender-responsive health approaches that take into account gender-specific needs. Gender-specific priorities should be promoted as compared to a blanket approach.

## Acknowledgement

This research was presented at the 5th AfHEA Biennial Scientific Conference, Gold Coast City, Ghana from the 11–14 March 2019. The authors are particularly grateful to participants at the conference for their comments.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

*Nozuko Lawana* http://orcid.org/0000-0003-0027-4725
*Frederik Booysen* http://orcid.org/0000-0001-7449-3221
*Forget Mingiri Kapingura* http://orcid.org/0000-0002-5808-5612
*Charles Hongoro* http://orcid.org/0000-0002-9715-4940

## References

Abegunde, DO, Mathers, CD, Adam, T & Mathers, C, 2007. The burden and costs of chronic diseases in low-income and middle-income countries. The Lancet 370, 1929–38.

Balia, S & Jones, AM, 2008. Mortality, lifestyle and socio-economic status. Journal of Health Economics 27, 1–26.

Bastida, E & Pagán, JA, 2002. The impact of diabetes on adult employment and earnings of Mexican Americans: Findings from a community based study. Health Economics 11, 403–413.

Cai, L & Cong, C, 2009. Effects of health and chronic diseases on labour force participation of older working- Age Australians. Australian Economic Papers 48, 166–182.

Cai, L & Kalb, G, 2006. Health status and labour force participation: Evidence from Australia. Health Economics 15, 241–261.

Cappellari, L & Jenkins, SP, 2003. Multivariate probit regression using simulated maximum likelihood. The Stata Journal: Promoting Communications on Statistics and Stata 3, 278–294.

Chib, S & Greenberg, E, 1998. Analysis of multivariate probit models. Biometrika 85, 347–361.

Harris, A, 2009. Diabetes, cardiovascular disease and labour force participation in Australia: An endogenous multivariate probit analysis of clinical prevalence data. Economic Record 85, 472–484.

Huffman, S & Rizov, M, 2014. Body weight and labour market outcomes in post-Soviet Russia. International Journal of Manpower 35, 671–687.

Jones, AM, 2007. Applied econometrics for health economists: A practical guide. CRC Press, London.

Khuwaja, AK, Khawaja, S, Motwani, K, Khoja, AA, Azam, IS, Fatmi, Z, Ali, BS & Kadir, MM, 2011. Preventable lifestyle risk factors for non-communicable diseases in the Pakistan adolescents schools study 1 (PASS-1). Journal of Preventive Medicine and Public Health 44, 210–214.

Maddala, GS, 1983. Limited dependent and qualitative variables in economics. Cambridge University Press, New York.

Minor, T, 2011. The effect of diabetes on female labor force decisions: New evidence from the National health interview survey. Health Economics 20, 1468–1486.

Miranda, JJ, Kinra, S, Casas, JP, Davey Smith, G & Ebrahim, S, 2008. Non-communicable diseases in low- and middle-income countries: Context, determinants and health policy. Tropical Medicine & International Health: TM & IH 13, 1225–34.

Morris, S, 2007. The impact of obesity on employment. Labour Economics 14, 413–433.

Narkiewicz, K, 2006. Obesity and hypertension—the issue is more complex than we thought. Nephrology Dialysis Transplantation 21, 264–267.

Norton, EC & Han, E, 2008. Genetic information, obesity, and labour market outcomes. Health Economics 17, 1089–1104.

Nwosu, CO & Woolard, I, 2017. The impact of health on labour force participation in South Africa. South African Journal of Economics 85, 481–490.

Paraponaris, A, Saliba, B & Ventelou, B, 2005. Obesity, weight status and employability: Empirical evidence from a French national survey. Economics & Human Biology 3, 241–258.

Puoane, TR, Tsolekile, LP, Egbujie, BA, Lewy, M & Sanders, D, 2017. Advancing the agenda on non-communicable diseases: Prevention and management at community level. SAHR 2017, 171–179.
Renna, F & Thakur, N, 2010. Direct and indirect effects of obesity on U.S. labor market outcomes of older working age adults. Social Science & Medicine (1982) 71, 405–13.
Schofield, DJ, Shrestha, RN, Passey, ME, Earnest, A & Fletcher, SL, 2008. Chronic disease and labour force participation among older Australians. Medical Journal of Australia 189, 447–50.
Schofield, DJ, Cunich, M, Shrestha, R, Passey, M, Kelly, S, Tanton, R & Veerman, L, 2014. The impact of chronic conditions of care recipients on the labour force participation of informal carers in Australia: Which conditions are associated with higher rates of non-participation in the labour force? BMC Public Health 14, 1–9.
Schofield, DJ, Cunich, M, Kelly, S, Passey, ME, Shrestha, R, Callander, E, Tanton, R & Veerman, L, 2015. The impact of diabetes on the labour force participation, savings and retirement income of workers aged 45–64 years in Australia 1–10. PLOS One 9, e89360.
Shahid, M, 2014. Impact of labour force participation on economic growth in Pakistan. Journal of Economics and Sustainable Development 5, 2222–1700.
Tunceli, K, Li, K & Williams, LK, 2006. Long-term effects of obesity on employment and work limitations among U.S. adults, 1986 to 1999. Obesity 14, 1637–1646.
Wilde, J, 2000. Identification of multiple equation probit models with endogenous dummy regressors. Economics Letters 69, 309–312.
World Health Organisation, (WHO), 2010. Global status report on noncommunicable diseases 2010 Global status report.
World Health Organisation, (WHO), 2015. Progress Monitor 2015.
Zhang, X, Zhao, X & Harris, A, 2009. Chronic diseases and labour force participation in Australia. Journal of Health Economics 28, 91–108.
Zhou, Y, Zheng, J, Li, S, Zhou, T, Zhang, P & Li, H-B, 2016. Alcoholic beverage consumption and chronic diseases. International Journal of Environmental Research and Public Health 13, 1–27.
van Zyl, S, van der Merwe, LJ, Walsh, CM, Groenewald, AJ, & van Rooyen, FC, 2012. Risk-factor profiles for chronic diseases of lifestyle and metabolic syndrome in an urban and rural setting in South Africa. African Journal of Primary Health Care & Family Medicine 4, 1–10.

## Appendix

**Table A1.** Coefficient of multivariate probit regression.

| | Total | | Males | | Females | |
|---|---|---|---|---|---|---|
| | Coefficient | Robust Std. Err. | Coefficient | Robust Std. Err. | Coefficient | Robust Std. Err. |
| Labour force | | | | | | |
| stroke | −1.02*** | 0.266 | −2.807*** | 0.448 | −0.662 | 0.606 |
| diabetes | −0.373** | 0.16 | 0.309 | 0.28 | −0.420** | 0.172 |
| Heart diseases | −0.766** | 0.201 | −1.347*** | 0.459 | −0.549 | 0.347 |
| Hypertension | −0.371*** | 0.136 | −0.249 | 0.201 | −0.516*** | 0.155 |
| Age | 0.202*** | 0.008 | 0.222*** | 0.012 | 0.190*** | 0.009 |
| Age2 | −0.002*** | 0.000 | −0.003*** | 0.000 | −0.002*** | 0.000 |
| Males | 0.373*** | 0.025 | | | | |
| Married | −0.015 | 0.028 | 0.359*** | 0.062 | −0.148*** | 0.032 |
| Primary | 0.257*** | 0.044 | 0.421*** | 0.091 | 0.191*** | 0.049 |
| Secondary | 0.421*** | 0.044 | 0.415*** | 0.089 | 0.415*** | 0.05 |
| Matric | 0.604*** | 0.05 | 0.562*** | 0.097 | 0.623*** | 0.058 |
| Tertiary | 1.070*** | 0.058 | 0.918*** | 0.113 | 1.139*** | 0.065 |
| Urban formal | 0.309*** | 0.025 | 0.305*** | 0.047 | 0.283*** | 0.03 |
| Rural | 0.316*** | 0.041 | 0.546*** | 0.086 | 0.196** | 0.047 |
| Urban informal | 0.425*** | 0.041 | 0.407*** | 0.077 | 0.413** | 0.048 |

(*Continued*)

**Table A1.** Continued.

| | Total | | Males | | Females | |
|---|---|---|---|---|---|---|
| | Coefficient | Robust Std. Err. | Coefficient | Robust Std. Err. | Coefficient | Robust Std. Err. |
| Coloured | −0.038 | 0.084 | −0.344** | 0.149 | 0.054 | 0.104 |
| African | 0.054 | 0.071 | −0.238** | 0.121 | 0.113 | 0.089 |
| Household size | −0.023 | 0.003 | −0.042*** | 0.006 | −0.012*** | 0.004 |
| Wave 1 | −0.450*** | 0.032 | −0.474*** | 0.058 | −0.452*** | 0.039 |
| Wave 2 | −0.212*** | 0.031 | −0.178*** | 0.06 | −0.239*** | 0.037 |
| Wave 3 | −0.210*** | 0.036 | −0.195*** | 0.065 | −0.229*** | 0.043 |
| Constant | −3.628*** | 0.163 | −3.301*** | 0.273 | −3.402*** | 0.185 |
| Stroke | | | | | | |
| Wave 1 | −0.216** | 0.109 | −0.538*** | 0.189 | −0.045 | 0.098 |
| Wave 2 | −0.110 | 0.108 | −0.466** | 0.21 | 0.067 | 0.092 |
| Wave 3 | −0.005 | 0.135 | −0.673*** | 0.236 | 0.206*** | 0.121 |
| Exercise | −0.272** | 0.131 | −0.678** | 0.341 | −0.100 | 0.155 |
| Obese | 0.125* | 0.071 | 0.235* | 0.137 | 0.085 | 0.072 |
| Smoke | 0.026 | 0.150 | −0.058 | 0.207 | 0.080 | 0.240 |
| Drinking | −0.796*** | 0.249 | −0.573** | 0.284 | −4.370*** | 0.123 |
| Age | 0.020*** | 0.003 | 0.018*** | 0.006 | 0.020*** | 0.003 |
| Constant | −3.185*** | 0.201 | −2.980*** | 0.46 | −3.276*** | 0.141 |
| Heart diseases | | | | | | |
| Wave 1 | −0.398*** | 0.069 | −0.375** | 0.176 | −0.379*** | 0.070 |
| Wave 2 | −0.272*** | 0.060 | −0.130 | 0.149 | −0.285*** | 0.061 |
| Wave 3 | −0.359*** | 0.086 | −0.379** | 0.165 | −0.334*** | 0.097 |
| Exercise | 0.180** | 0.091 | 0.513*** | 0.161 | −0.001 | 0.096 |
| Obese | 0.158*** | 0.048 | 0.292*** | 0.112 | 0.073 | 0.052 |
| Smoke | 0.021 | 0.120 | 0.143 | 0.167 | 0.025 | 0.17 |
| Drinking | −0.655*** | 0.190 | −0.740*** | 0.278 | −0.489* | 0.273 |
| Age | 0.024*** | 0.002 | 0.031*** | 0.006 | 0.023*** | 0.002 |
| Constant | −2.905*** | 0.106 | −3.584*** | 0.272 | −2.715*** | 0.11 |
| Diabetes | | | | | | |
| Wave 1 | −0.038 | 0.068 | −0.028 | 0.114 | −0.036 | 0.08 |
| Wave 2 | 0.091 | 0.059 | 0.094 | 0.113 | 0.09 | 0.066 |
| Wave 3 | −0.119 | 0.083 | −0.206 | 0.187 | −0.068 | 0.09 |
| Exercise | −0.075 | 0.122 | −0.218* | 0.127 | 0.043 | 0.166 |
| Obese | 0.281*** | 0.049 | 0.287*** | 0.096 | 0.278*** | 0.055 |
| Smoke | 0.147 | 0.113 | 0.208 | 0.152 | 0.069 | 0.177 |
| Drinking | −0.135 | 0.199 | −0.007 | 0.213 | −0.611* | 0.324 |
| Age | 0.039*** | 0.002 | 0.044*** | 0.004 | 0.036*** | 0.002 |
| Constant | −3.568*** | 0.106 | −3.816*** | 0.197 | −3.452*** | 0.121 |
| Hypertension | | | | | | |
| Wave 1 | −0.179*** | 0.045 | −0.103 | 0.085 | −0.174*** | 0.05 |
| Wave 2 | 0.066* | 0.04 | 0.183** | 0.083 | 0.052 | 0.044 |
| Wave 3 | −0.125** | 0.052 | 0.015 | 0.117 | −0.171*** | 0.056 |
| Exercise | −0.112* | 0.064 | −0.127 | 0.101 | 0.004 | 0.081 |
| Obese | 0.420*** | 0.032 | 0.432*** | 0.068 | 0.341*** | 0.036 |
| Smoke | 0.157** | 0.077 | 0.333*** | 0.093 | 0.214* | 0.115 |
| Drinking | −0.039 | 0.141 | 0.14 | 0.187 | 0.009 | 0.182 |
| Age | 0.044*** | 0.001 | 0.045*** | 0.003 | 0.045*** | 0.001 |
| Constant | −3.042*** | 0.062 | −3.476*** | 0.123 | −2.911*** | 0.07 |
| Sample size | 35 825 | | 11 097 | | 24 728 | |
| Log likelihood | −6.82E + 07 | | −1.90E + 07 | | −4.8E + 07 | |
| Wald $\chi^2$ | 5142.6 | | 1742.02 | | 9058.9 | |
| LR test, *p*-value | 0.000 | | 0.000 | | 0.000 | |