Depression as a lead factor to tobacco consumption: A case of South Africa

Nomusa Yolanda Nkomo (University of Zululand)
Beatrice Desiree Simo-Kengne (University of Johannesburg)
Mduduzi Biyase (University of Johannesburg)

Abstract

The relationship between smoking and depression is unclear, and the factors that account for their comorbidity have received only limited attention. This study aims to clarify such associations in South Africa using the collated five waves of the National Income Dynamic Survey dataset. Different statistical approaches are used to account for selection bias and endogeneity. Empirical results reveal that, apart from socio-economic factors, depression significantly influence the decision to smoke, but not the smoking intensity. Furthermore, tobacco consumption has been found to be significantly higher among men, though the gap seems to be narrowing partly because women are being targeted in marketing. Interestingly, smoking proves to be more prevalent in the younger generation and tends to decline with age, while married individuals have a lower likelihood of smoking and, when they do smoke, their smoking frequency is lower compared to unmarried individuals.

Keywords: Depression, tobacco consumption, addiction, selection bias, endogeneity
JEL Classification. C23, I18, L66

Introduction

Tobacco is the single largest preventable cause of death and chronic disease in the world today. While most adults globally are conscious of the physical health risks of tobacco consumption, research indicates that smoking also affects people's mental health. Though smoking provides short-term benefits such as an improved mood, concentration, reduced anger and stress, and relaxation of the muscles, these instances are outweighed by the higher rates of smoking-related physical health problems like lung cancer, which are common among people with mental health problems. The motives for smoking vary individually, and an understanding of the reasons for smoking might assist those who wish to stop.

Verma et al. (2023) and Buckley (2007) found that social and psychological factors contribute towards smokers continuing to smoke. For instance, many teenagers point out that they started smoking to experiment. Yet, other factors encourage people to become regular smokers, including having friends or family who smoke and parents' attitude towards smoking (Xaba et al., 2021; Bashirian et al., 2021; Ritt-Olson et al., 2005). Moreover, as teenagers grow into adulthood, they are more likely to become regular smokers if they drink alcohol or abuse substances; such issues trigger the likelihood of an individual encountering stress (Nkomo and Adanlawo, 2024; Munafo et al., 2008) and eventually starting to smoke. Some adults claim that smoking has become a habit as it helps them relax and minimise stress (Musa et al., 2023; Fithria et al., 2021; Khantzian, 1997). This hypothesis, that individuals smoke to relieve stress, is known as 'self-medication'. Thus, stress is one of the contributors to some people finding it difficult to cope with work or pressure, while others suffer from headaches, breathlessness, irritability or anxiety. Continuous feelings of stress are also associated with depression, which often causes individuals to self-medicate by smoking and/or drinking more alcohol than usual.

Although smoking is thought to help people relax, Buckley (2007) reported that smoking escalates feelings of anxiety and tension. Nicotine causes an instant sense of relaxation, making people believe that it reduces stress and anxiety. However, this feeling of relaxation is short-lived, soon giving way to withdrawal symptoms and continuously increased cravings. Escobedo et al. (1998) found that the rate of smoking in adults with depression is approximately twice as high as in adults without depression. The same study also pointed out that individuals suffering from depression experience difficulties when they try to quit smoking. In addition, several studies concluded that

most individuals start smoking before showing any signs of depression, so it is unclear whether smoking causes depression or depression inspires people to smoke. However, a higher rate of depression has been reported amongst smoking individuals than non-smokers (Chaka and Adanlawo, 2022; Kluwe-Schiavon et al., 2020; Breslau et al., 1998). Thus, increased symptoms of depression are likely to be linked to the progressive use of tobacco.

The theory of normal consumer choice assumes that every person is likely to be a consumer of all goods, provided they can afford it (Mncwango et al., 2024; Adanlawo and Nkomo, 2023). However, this might not be applicable for some commodities like tobacco, which some people cannot be persuaded to consume, regardless of the price or their disposable income levels. For these goods, non-consumption is not necessarily explained by economic affordability; non-zero consumption may also be attributed to other behavioural factors, such as mental conditions. Thus, tobacco consumption is based on discrete random preference systems, where smokers may have a different preference structure to non-smokers (Kenkel et al., 2024; Reynolds et al., 2024; Adanlawo, 2017; Pudney, 1989), including addiction, self-medication, an imitation of the habits of the people around them and enjoyment, among others.

Oppression leads to depression, which aggravates mental disorders, and South Africa has historically experienced oppression due to apartheid. Although critics have admonished against continuing to blame apartheid for the deteriorated mental health of South Africans (Dawes et al., 1989; Dawes, 1990), mental health illnesses are still reflected as one of the main public health problems in the country (McGinty et al., 2024; Tomlinson et al., 2009). Although depressed individuals are likely to initiate regular smoking to self-medicate against depressive moods, it is surprising that existing studies on tobacco use in South Africa have focused solely on socio-economic and demographic reasons for smoking. This study investigates the plausibility of mental health conditions acting as stimuli for smoking in South Africa.

Literature review

Smoking behaviour can be understood as an intertemporal decision causing instantaneous satisfaction, yet long-term health risks (Xaba et al., 2024; Miura, 2019). As determined, tobacco smoking can lead to poor mental and physical health. In the empirical literature, tobacco use has been established as the main cause of deaths from cancer, heart disease, stroke, lung diseases, diabetes and respiratory diseases (Nyeleti and Maphela, 2024; West, 2017). In addition, mental

disorders can also arise as both a cause and a consequence of smoking, although most existing evidence focuses on smoking as the driver of mental illnesses.

Although, the causal direction between smoking and mental disorders is inconclusive, numerous studies affirm a positive link between smoking and mental illnesses. However, some prior studies did report a negative or zero relationship between cigarette smoking and depression. Audrain-Mcgovern et al. (2006) reported the absence of an association between depression and tobacco consumption. Similarly, Park and Romer (2007) found no link between smoking and mental illness. These findings were consistent with those of White et al. (2007), who also found no association between depression and successive smoking among African and Hispanic Americans.

Contrary to those studies, Munafo et al. (2008) emphasised tobacco use as a cause of depression and concluded that the relationship between smoking and MHBs might be clarified by causal inference. In support, Kang and Lee (2010) conducted a study – using logistic regression – on the causal connection between tobacco consumption and depression, concluding that smoking instigated depression. Another causal study by Boden et al. (2010) used longitudinal data on adolescents and proved that smoking is a cause of mental illness, like depression. Apart from these studies showing that cigarette smoking is a predictor of depression, some research has established that smoking escalates the dangers of depression (Maphela, 2021; Pasco et al., 2008; Breslau et al., 1998). Fergusson et al. (2003); Chaka and Adanlawo (2024), and Choi et al. (1997) further concluded that peer tobacco consumption leads to depression among adolescents.

Other studies separated this relationship along gender lines. Kandel et al. (2007) discovered that smoking trends in men are related to depression in their future. There was also reported support that the connection between tobacco use and depression is stronger among women than men (Repetto et al., 2005). Upon investigating the risk of depression among smoking women, Milic et al. (2010) concluded that tobacco use is linked to an increased rate of depression in women and has the potential to cause harmful consequences to their mental health. In support of this study, Pasco et al. (2008) explored the role of tobacco smoking as a cause of major depression among women. Their study concluded that depressive disorders in women result from smoking, and smoking continues to be a predictor of depression after controlling for family and hereditary background.

Methodology and data

This study used the NIDS dataset to investigate the impact of MHBs on tobacco consumption. The dataset is a longitudinal survey of South African households. This survey is biannual and is conducted by the Southern Africa Labour and Development Research Unit (SALDRU). This study used all five waves that had been released at the time of data collection to conduct this investigation. The variables incorporated in this study included the key explanatory variable, depression, the education variable, a categorical variable that is categorised into five divisions, namely no schooling, primary education, secondary education, matric and tertiary education; urbanisation; race; household income; gender; marital status; employment status; and age. All variables are either binary or categorical, except for household income, total consumption, and age. The dependent variables are regular smokers and the intensity of smoking. The regular smoking variable was drawn from the survey questions, where individuals were asked whether they smoked or not. This variable was categorised into binary variables (0 if they do not smoke and 1 if they smoke). However, individuals who do not smoke are indicated by 0, while those who smoke, but cannot afford to smoke at a certain time for any reason, display as 0 as well.

The study uses different approaches in addressing the zero observation concerns, namely Heckman's selection model and the double-hurdle model proposed by Cragg (1971). Like the Heckman model, the double-hurdle model is a parametrical generalisation of the Tobit model determined by two distinct stochastic processes (Chaka and Adanlawo, 2023). One of the key differences between these suggested models is the fact that two-part models assume only true zeros, while the Tobit Type II (Heckman selection) accounts for both types of zeros in the first selection equation and the outcome equation at y > 0. However, as Cragg (1971) proposed, a double-hurdle model accounts for 'unobserved' zeros during the selection and 'true' zeros during the outcome equation. This leads to two equations set:

The outcome/consumption model:

$$Y_{it}^{*} = X_{it}\beta 1 + \mu_{it} \quad (1)$$

The selection model:

$$S_{it}^{*} = Z_{it}\gamma + \nu_{it} > 0 \quad (2)$$

where Y^*_{it} (smoking frequency) and S^*_{it} (regular smoking) represent the unobserved latent variables, and X_{it} and Z_{it} are explanatory in both models, respectively. Generally, X is presumed to be a subset of Z, in that all elements that predict smoking frequency (Y/outcome) predict regular smoking as well (S/selection). The symbols μ and ν display the normally distributed error terms, and β and ɣ present the parameters to be estimated. The smoking frequency, which is the outcome variable, is observed only if regular smoking or selection is above 0, i.e.:

$$S_{it} = \begin{pmatrix} 1 \; if \; S^*_{it} > 0 \\ 0 \; if \; S^*_{it} \geq 0 \end{pmatrix} \quad (4)$$

The key notion of the Heckman selection model is that it appears theoretically and relatively possible that unobservable factors may influence both smoking frequency and the possibility of smoking regularly (S_{it} or selection). These unobserved elements will be contained in the residuals of the two equations. If selected into the primary sample, the expected value of smoking frequency or the outcome in the key equation are displayed by:

$$E(y|z,\nu) = X\beta + E(\mu|\nu) \quad (5)$$

Since the covariates X and ν conjointly define selection into the sample, $Corr(x, \nu/s) = 1$ is non-zero, and as a result, the β coefficients will be biased and inconsistent if μ and ν are linked or correlated. This correlation is easy to estimate by inserting separate models for Y and S, and estimating covariates between μ and ν. The Heckman selection posits that this bias can be fixed by calculating the expected value of ν on Z in the sample, and also adding the same term in the primary empirical equation. To achieve consistent coefficients, the maximum likelihood technique F is used to estimate the probability of an individual to become a smoker (selection model); the output of which helps to calculate the inverse Mills ratio for every individual. The inverse Mills ratio is finally included in the outcome regression model as an explanatory variable. Similar to the Heckman selection, the double-hurdle selection also uses two separate equations; firstly, the decision to become a regular smoker, and secondly, the frequency of smoking per individual. This model accounts for zeros in participation derived from the decision to smoke or assumes that the two separate hurdles must pass before the positive level of smoking will be observed.

Descriptive statistics and preliminary analysis

Table 1 presents a summary of the key features of the analysed sample, which is a merged NIDS dataset of the 2008, 2010, 2012, 2014 and 2017 waves. The total sample size is 92 460 observations

(individuals) with an average age of approximately 25 years, and the majority being female (54.6%). Of the sample, 42.5% self-reported that they were depressed, and 13% indicated experiencing signs of sleeplessness or stress. The statistics also suggest that the maximum number of cigarettes that smoking individuals can smoke per day, is five. For all the binary variables, the outcome of 0 is treated as a reference variable, compared to an outcome of 1. The following reference variables are considered: male, rural, African, Black, not married (widowed, separated, divorced, never married), no schooling, employed, non-smoker. All the variables were controlled for time-fixed effects and the issue specifics attached to the different waves.

Table 1: Descriptive statistics

Variable	Description	Mean	Std.Dev	Min	Max
Regular smoker	1 if a regular smoker, 0 otherwise	0.0344	0.1822	0	1
Smoking frequency	Avg number of cigarettes consumed per day	3	1.144	1	5
Sleepless	1 if sleepless, 0 otherwise	0.1305	0.3368	0	1
Urban	1 if urban, 0 otherwise	0.5429	0.4981	0	1
Female	1 if female, 0 otherwise	0.5466	0.4978	0	1
Married	1 if married, 0 otherwise	0.2603	0.4388	0	1
Own house	1 if owning a house, 0 otherwise	0.7902	0.4071	0	1
Depressed	1 if depressed, 0 otherwise	0.423	0.494	0	1
Western Cape	1 if from Western Cape, 0 otherwise	0.1213	0.3265	0	1
Eastern Cape	1 if from Eastern Cape, 0 otherwise	0.0697	0.2547	0	1
Northern Cape	1 if from Northern Cape, 0 otherwise	0.051	0.22	0	1
Free State	1 if from Free State, 0 otherwise	0.2856	0.4517	0	1
KwaZulu-Natal	1 if from KwaZulu-Natal, 0 otherwise	0.065	0.2465	0	1
North West	1 if from North West Prov., 0 otherwise	0.1284	0.3345	0	1
Gauteng	1 if from Gauteng, 0 otherwise	0.0707	0.2563	0	1
Mpumalanga	1 if from Mpumalanga, 0 otherwise	0.0766	0.2659	0	1
Unemployed	1 if unemployed, 0 otherwise	0.3798	0.4853	0	1
Age	Age of each individual	24.83	.19.88	0	107
Age-SQD	Age of each individual squared	101.06	1372.87	0	11449
Coloured	1 if Coloured, 0 otherwise	0.1428	0.3498	0	1
Asian_Indian	1 if Asian or Indian, 0 otherwise	0.0292	0.1683	0	1
White	1 if White, 0 otherwise	0.0806	0.2722	0	1
Primary education	1 if primary, 0 otherwise	0.3049	0.4603	0	1
Secondary education	1 if secondary, 0 otherwise	0.2584	0.4378	0	1
Matric	1 if matric, 0 otherwise	0.1076	0.3033	0	1
Tertiary education	1 if tertiary, 0 otherwise	0.1132	0.3169	0	1
HH income	Household total income	8.5331	1.0778	0.677	14.773

Table 2 shows the results of the test for selection bias, checking whether endogeneity exists due to reverse causality between regular smoking and MHBs (explanatory variable). In the initial stage, a probit model is estimated (model 1), and the inverse Mills ratio (IMR) is derived from it. The IMR derived from the first stage is used to correct selection bias. Thereafter, the IMR variable is included as an explanatory variable in the second equation (model 2), where MHBs show potential endogeneity among the variables. In model 2, MHB is proxied by

depression, which is the dependent variable. In model 3, the coefficient of the derived residual is significantly not equal to zero where endogeneity exists.

Table 2: A test for endogeneity

Variables	Model 1	Model 2	Model 3
Depressed	0.176***		
	(-0.058)		
Sleepless	0.322***	0.822***	-0.707***
	(-0.075)	(-0.028)	(-0.271)
Urban	0.190**	0.0195	0.0362
	(-0.082)	(-0.024)	(-0.043)
Female	-1.052***	0.0285	-0.255*
	(-0.076)	(-0.019)	(-0.141)
Married	0.041	-0.144***	0.138**
	(-0.076)	(-0.025)	(-0.057)
Unemployed	-0.092	0.049**	-0.034
	(-0.067)	(-0.021)	(-0.033)
Age	0.055***	0.038***	0.002
	(-0.010)	(-0.002)	(-0.015)
Age-SQD	0.004***	-0.003***	0.000
	(-0.000)	(0.000)	(-0.000)
Coloured	0.806***	-0.168***	0.377***
	(-0.121)	(-0.041)	(-0.126)
Asian_Indian	0.359*	-0.180**	0.378***
	(-0.211)	(-0.070)	(-0.115)
White	0.794***	-0.231***	0.999***
	(-0.121)	(-0.047)	(-0.139)
Western Cape	-0.341**	0.043	-0.063
	(-0.137)	(-0.047)	(-0.073)
Eastern Cape	-0.115	-0.077	0.000
	(-0.126)	(-0.047)	(-0.057)
Northern Cape	0.113	0.139**	-0.067
	(-0.159)	(-0.058)	(-0.087)
Free State	-0.596***	0.031	0.043
	(-0.136)	(-0.045)	(-0.096)
KwaZulu-Natal	-0.245	0.109**	-0.0246
	(-0.163)	(-0.054)	(-0.087)
North West	-0.108	0.043	0.018
	(-0.123)	(-0.047)	(-0.059)
Gauteng	-0.358**	-0.056	0.145*
	(-0.157)	(-0.053)	(-0.083)
Mpumalanga	-0.492***	-0.101*	0.110

	(-0.169)	(-0.053)	(-0.104)
Primary education	0.065	0.002	0.080
	(-0.151)	(-0.049)	(-0.069)
Secondary education	-0.118	-0.0395	0.155**
	(-0.156)	(-0.049)	(-0.072)
Matric	-0.129	-0.0765	0.177**
	(-0.169)	(-0.053)	(-0.083)
Tertiary education	-0.102	-0.140**	0.227**
	(-0.167)	(-0.054)	(-0.091)
HH income	0.013	-0.057***	0.090***
	(-0.034)	(-0.011)	(-0.024)
House/dwelling	-0.037	-0.039*	0.031
	(-0.071)	(-0.024)	(-0.031)
No exercise		0.084***	
		(-0.025)	
IMR_RESD			0.932***
			(-0.323)
Constant	-3.421***	-0.580***	0.888
	(-0.409)	(-0.124)	(-0.584)

Note: ***, **, * denote statistical significance at 1%, 5% and 10% level of significance, respectively. Standard errors are reported in parentheses.

Empirical results and discussion

Table 3 shows the probit regression model results of the determinants of tobacco consumption. In model 1, only socio-economic factors are controlled for, while model 2 includes all the control variables and accounts for MHBs proxied by depression. In this data set, 17 368 observations were part of the estimations. The likelihood ratio (LR) χ^2 test, probability value and log-likelihood display the overall significance of the models. The output for model 2 shows that the model accounting for MHBs is the best, as it has a log-higher likelihood close to zero. The LR $\chi^2$of 122.79, with a probability value of 0.000 (model 1), and LR $\chi^2$of 120.98 with a probability value of 0.000 (model 2) indicate that both these models are statistically significant. Both sets of results indicate that several socio-economic factors play an important role in explaining the smoking decision. However, when MHBs (depression) are controlled for in model 2, the results also indicate a positive and highly significant relationship between MHBs and tobacco consumption. This suggests MHBs also have a significant influence on an individual who decides to start smoking. The estimates from the probit model might be biased as they take selection bias and endogeneity into consideration.

Table 3: Probit model estimates of the determinants of tobacco consumption

Variables	***Model 1***	***Model 2***
Depressed		0.176***
		(-0.058)
Sleeplessness	0.374***	0.322***
	(-0.073)	(-0.075)
Urban	0.189**	0.190**
	(-0.082)	(-0.082)
Female	-1.047***	-1.052***
	(-0.076)	(-0.076)
Married	0.031	0.041
	(-0.076)	(-0.076)
Own house	-0.040	-0.037
	(-0.072)	(-0.071)
Western Cape	-0.339**	-0.341**
	(-0.137)	(-0.137)
Eastern Cape	-0.119	-0.115
	(-0.126)	(-0.126)
Northern Cape	0.122	0.113
	(-0.16)	(-0.159)
Free State	-0.597***	-0.596***
	(-0.136)	(-0.136)
KwaZulu-Natal	-0.235	-0.245
	(-0.163)	(-0.163)
North West	-0.104	-0.108
	(-0.123)	(-0.123)
Gauteng	-0.367**	-0.358**
	(-0.158)	(-0.157)
Mpumalanga	-0.501***	-0.492***
	(-0.169)	(-0.169)

Socio-demographic factors and smoking

Table 4 displays the results of the Heckman model and double-hurdle model. The log-likelihood information provided indicates the fitness of the models. Specifically, the log-likelihood, Wald test, and the likelihood ratio (LR) tests indicate the overall significance of the models. The results for all the columns show probability values of 0.000, which indicates that all the models meet the goodness of fit criteria and are statistically significant.

When selection bias is controlled for, the Heckman selection (column 1) indicates that the depression and control variables – sleeplessness, being in urban areas, gender, age, age squared, race (Black being the reference category), married, secondary education, matric, tertiary education, household income, own house, and location (provincial) – have significant effects on the selection equation. It indicates that they are important drivers of smoking participation. The estimates for the intensity equations are shown in column 2 of the Heckman selection. The results portray a different conclusion from the selection model. The coefficients become insignificant except for race (White population), secondary education, and household income. This simply shows that the determinants of smoking are not the same as those explaining the intensity of tobacco consumption. The double-hurdle model indicates that sleeplessness, urban location, being female, age, age squared, race and household income are vital factors in determining smoking frequency.

The smoking selection equation in the double-hurdle model captures the smoking participation effects. Results indicate that most of these variables have not affected smoking participation, owing to their statistical significance. These include MHBs, sleeplessness, urban/rural location, gender, age, age squared, race (Black being the reference category), and provincial location (Western Cape, Free State, Gauteng, and Mpumalanga). Marital and employment status, residing in the Eastern Cape, KwaZulu-Natal and North West, and educational status, as well as homeownership, do not significantly predict tobacco use (participation). Consistent with the Heckman selection, the first-hurdle estimates suggest that individuals who suffer from depression are more likely to smoke, and that women are less likely to smoke than men. In addition, living in an urban area also leads to a higher probability of an individual being a regular smoker in this model. This might be attributed to exposure to publicity, as well as the image portrayed about the “coolness” of tobacco smoking in cities through television and peers. Furthermore, age and age squared are significant in participation and consumption equations. The age coefficient is positive and highly significant, while age squared is significant with a

negative sign. This means that the younger generation is more likely to be regular smokers than the older cohort, simply because older individuals have smoked more in their lives and, therefore, better understand the health risks involved in smoking, in line with lifecycle patterns in smoking behaviour (Kerr, 2004; Aristei & Pieroni, 2005).

Table 4: Estimation output

	Heckman selection		***Double-hurdle model***	
	Smoking selection	***Smoking intensity***	***Smoking selection***	***Smoking intensity***
Depressed	0.154***	-0.0076	0.175***	-0.0045
	(0.0256)	(0.0336)	(0.0581)	(0.0232)
No exercise	-0.0190			
	(0.0302)			
Sleeplessness	0.107***	0.0596	0.321***	0.085***
	(0.0362)	(0.0379)	(0.0753)	(0.0321)
Urban	0.140***	0.0354	0.192**	0.0643**
	(0.0322)	(0.0380)	(0.0818)	(0.0304)
Female	-1.252***	-0.0772	-1.052***	-0.209***
	(0.0267)	(0.186)	(0.0764)	(0.0286)
Married	-0.387***	0.0735	0.0440	0.0204
	(0.0331)	(0.0632)	(0.0757)	(0.0297)
Unemployed	-0.0304	0.0081	-0.100	0.0060
	(0.0286)	(0.0287)	(0.0642)	(0.0263)
Age	0.103***	0.0232	0.0556***	0.0312***
	(0.0043)	(0.0160)	(0.0099)	(0.0042)
Age SQD	-0.0011***	-0.0001	-0.0004***	-0.0002***
	(5.08e-05)	(0.0001)	(0.0001)	(4.86e-05)
Coloured	1.044***	0.0926	0.810***	0.233***
	(0.0472)	(0.153)	(0.119)	(0.0409)
Asian_Indian	0.720***	0.121	0.370*	0.274***
	(0.0811)	(0.132)	(0.209)	(0.0762)
White	1.130***	0.643***	0.808***	0.788***
	(0.0531)	(0.169)	(0.116)	(0.0469)
Western Cape	-0.166***	-0.0149	-0.341**	-0.0121
	(0.0551)	(0.0565)	(0.137)	(0.0477)
Eastern Cape	-0.114**	-0.0337	-0.117	-0.0053
	(0.0509)	(0.0468)	(0.126)	(0.0406)
Northern Cape	0.174**	0.0265	0.113	0.0827
	(0.0675)	(0.0690)	(0.159)	(0.0588)
Free State	-0.364***	0.0917	-0.595***	0.0719
	(0.0539)	(0.0739)	(0.135)	(0.0482)

	Heckman selection		*Double-hurdle model*	
	Smoking selection	***Smoking intensity***	***Smoking selection***	***Smoking intensity***
KwaZulu-Natal	-0.124*	0.0877	-0.245	0.119*
	(0.0667)	(0.0692)	(0.163)	(0.0625)
North West	-0.0861	0.0371	-0.104	0.0439
	(0.0539)	(0.0517)	(0.123)	(0.0465)
Gauteng	-0.146**	0.104	-0.356**	0.0885
	(0.0638)	(0.0650)	(0.157)	(0.0570)
Mpumalanga	-0.229***	0.0341	-0.491***	0.0134
	(0.0667)	(0.0748)	(0.169)	(0.0624)
Primary education	0.0262	0.0679	0.0676	0.0672
	(0.0659)	(0.0634)	(0.153)	(0.0592)
Secondary education	-0.179***	0.118*	-0.113	0.0572
	(0.0665)	(0.0679)	(0.156)	(0.0598)
Matric	-0.426***	0.126	-0.121	0.0315
	(0.0723)	(0.0929)	(0.167)	(0.0662)
Tertiary education	-0.441***	0.141	-0.0893	0.0687
	(0.0727)	(0.0935)	(0.164)	(0.0658)
Own house	-0.138***	-0.0042	-0.0330	-0.0292
	(0.0299)	(0.0356)	(0.0710)	(0.0267)
lHH income	-0.0381***	0.0518***		0.062***
	(0.0147)	(0.0157)		(0.0138)
Residual				
Residual_depresd				
Constant	-1.876***	0.707	-3.313***	0.322**
	(0.166)	(0.490)	(0.295)	(0.155)
Observations	20,247	20,247	17,371	3,357
Wald chi2	(26)182.77	(25)102.12	(26)277.54	(26) 277.54
P>Chi2	0.000	0.000	0.000	0.000
Log-likelihood	-	-	-696.529	-1104.741

Note: ***, **, * denote statistical significance at 1%, 5% and 10% level of significance, respectively. Standard errors are reported in parentheses.

Conclusion and policy recommendations

This study investigated the relationship between MHBs and tobacco consumption using the Heckman and double-hurdle models. Firstly, the main explanatory variable MHBs (proxied by depression) showed positive and highly significant coefficients, implying that people tend to smoke when they are depressed. MHBs appear to be a significant driver of the smoking

decision according to the Heckman selection model. In terms of residential area, results illustrate that living in urban areas contributes significantly to the decision to smoke (participation equation), but does not indicate the intensity of the smoking or the amount of tobacco smoked. Age was also a significant factor in determining tobacco use, with results showing that, as people grow older, there is a reduction in their smoking, which could largely be explained by the fact that our human health status deteriorates as we grow older.

In terms of race, Black people were least likely to be regular smokers, possibly signifying a protective factor(s) against tobacco use in the Black community, or connotations associated with tobacco smoking within the Black culture. High levels of unemployment in this part of the population also make it near impossible to support the habit. The estimates for Coloured and White individuals are positive with highly statistically significant coefficients compared to their counterparts (the Black race). These findings support prior studies (Rodriguez & Ruigomez, 1999), with evidence showing the relationship between being unemployed, the use of substances, and dissimilarities between different population groups in terms of socio-demographic factors.

Results also indicate that the proportion of women who are regular smokers is smaller than that of men. This implies women are less likely to smoke than their male counterparts. This is consistent with the analyses conducted by Boden et al. (2010) and McGue et al. (2000), mirroring the stereotypical norms of males concerning smoking. This study only reported a negative and significant relationship with regard to secondary, matric and tertiary education in the Heckman selection model (column 1), but a positive relationship in all other models. The results from these variables show that people who attain these levels of education are more likely to smoke. Consequently, even when individuals further their studies and obtain more knowledge on the risks of smoking, they continue to smoke. Many studies have shown contradictory findings between smoking and education. Some researchers reported that different levels of education are positively related to tobacco use (Hu & Tsai, 2000; Yu & Abler, 2008), while others report the opposite negative relationship (Decker & Schwartz, 2000; Yen, 2005).

Over the years, several policies have been implemented concerning tobacco consumption. In 2019 alone, 136 countries initiated at least one policy that aims to reduce smoking (WHO, 2019). These policies were applied on a national level, rather than targeting the affected group. This study provided positive insights into the smoking behaviour of people with MHBs, such

as depression, in South Africa. The use of the Heckman selection and the double-hurdle model allowed us to investigate both the decision to smoke and the intensity of smoking per individual. An analysis of smoking intensity provides better guidelines to examine the smoking behaviour of a particular community or group. Since the data used (NIDS) were from a national survey and representative of the whole population, the outcome of this study will inform policymakers as well as healthcare professionals about the smoking decision and intensity thereof among depressed individuals. The findings show that people suffering from MHBs smoke more than those who are not affected. These findings advocate that socio-demographic and mental health features influence smoking behaviours.

The outcomes of this study raise important concerns on the trends of tobacco use in South Africa and the accompanying factors leading to its use. Policies and behavioural interventions are recommended to reduce smoking in order to improve the lives of those South Africans suffering from MHBs. The available tobacco control policies directed at the general population may not have worked as effectively for people with mental illness. The main focus of the smoke-free policy and tobacco ban should thus be tailored to individuals suffering from MHBs.

Conflict of interest.

The authors declare that they have no conflict of interest.

Informed consent

All procedures followed were in accordance with the ethical standards of the University of Cape Town's School of Economics who granted ethical approval to the Southern Africa Labour and Development Research Unit (SALDRU) to conduct the National Income Dynamics Study (NIDS), a longitudinal household study in South Africa.

References

Adanlawo, E. F. (2017). *Community development: The use of corporate social responsibility initiatives by shopping centre Landlords* (Doctoral dissertation, University of Zululand).

Adanlawo, E. F., & Nkomo, N. Y. (2023). The implications of population aging on local health care expenditure: A 22-year panel data analysis. *International Journal of Innovative Technologies in Social Science*, 3 (39).

Aristei, D. & Pieroni, L. (2005). A double-hurdle approach to modeling tobacco consumption in Italy. *Applied Economics*, *40*(19):2463-2476.

Bashirian, S., Barati, M., Karami, M., Hamzeh, B., Afshari, M., & Ezati, E. (2021). Determinants of waterpipe smoking among women: A systematic review. *International Journal of Preventive Medicine*, *12*.

Baye, F.M. & Fambon, S. (2010). Parental Literacy and Child Health Production in Cameroon. *African Journal of Economic Policy*, *17*(2):99-130.

Boden, J.M., Fergusson, D.M. & Horwood, L.J. (2010). Cigarette smoking and depression: tests of causal linkage using a longitudinal birth cohort. *British Journal of Psychiatry*, *196*: 440–446.

Breslau, N., Peterson, E., Schultz, L., Chilcoat, H. & Andreski, P. (1998). Major depression and stages of smoking: A longitudinal investigation. *Archives of General Psychiatry*, *55*:165-166.

Buckley, P.F. (2007). Dual diagnosis of substance abuse and severe mental illness: the scope of the problem. *Journal of Dual Diagnosis, 3*(2):59-62.

Chaka, M., & Adanlawo, E. F. (2022). Role of public relations (PR) in nation-building: A case study of South Africa. *Journal of Nation-Building and Policy Studies*, *6*(3), 5.

Chaka, M., & Adanlawo, E. F. (2023). The Impact of Ethnicity on South Africa's National Unity. *African Renaissance*, *20*(2), 315-330.

Chaka, M., & Adanlawo, E. F. (2024). The role of communication in nation-building: A theoretical framework for South African national unity. *Studies in Media and Communication*, *12*(3), 325-334.

Cragg, J. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica, 39*:829-844.

Dawes, A. (1990). The effects of political violence on children: A consideration of South African and related studies. *International journal of psychology*, *25*(1):13-31.

Dawes, A., Tredoux, C. & Feinstein, A. (1989). Political violence in South Africa: Some effects on children of the violent destruction of their community. *International journal of mental health*, *18*(2):16-43.

Decker, S. & Schwartz, A.E. (2000). Cigarettes and alcohol: substitutes or complements?.

Dierker, L.C., Avenevoli, S., Stolar, M. & Merikangas, K.R. (2002). Smoking and depression: an examination of mechanisms of comorbidity. *American Journal of Psychiatry*, *159*:947–953.

Escobedo, L.G., Reddy, M. & Giovino, G.A. (1998). The relationship between depressive symptoms and cigarette smoking in US adolescents. *Addiction, 93*:433–440.

Fithria, F., Adlim, M., Jannah, S. R., & Tahlil, T. (2021). Indonesian adolescents' perspectives on smoking habits: a qualitative study. *BMC Public Health*, *21*, 1-8.

Frijters, P., Johnston, D.W. & Shields, M.A. (2010). Mental health and labor market participation: Evidence from IV panel data models. IZA Discussion Paper No. 4883.

Garen, J. (1984). The Returns to Schooling: A Selectivity Bias Approach with a Continuous Choice Variable. *Econometrica, 52*(5):1199-1218.

Hamilton, B.H. & Nickerson, J.A. (2003). Correcting for endogeneity in strategic management research. *Strategic Organization, 1*(1):51–78.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica, 47*:153-161.

Heckman, J.J. & R. Robb (1985). Alternative Methods for Evaluating the Impacts of Interventions: An Overview. *Journal of Econometrics*, 30, 239-267.

Hu, T.W. & Tsai, Y.W. (2000). Cigarette consumption in rural China: survey results from 3 provinces. *American Journal of Public Health, 90*(11):1785.

Jones, A.M. (1989). A double-hurdle model of cigarette consumption. *Journal of Applied Econometrics*, *4*:23-39.

Kendler, K.S., Neale, M.C. & Maclean, C.J. (1993). Smoking and major depression. A causal analysis. *Archives of General Psychiatry, 50*:36–43.

Kenkel, D. S., Mathios, A. D., Phillips, G. N., Suryanarayana, R., Wang, H., & Zeng, S. (2024). *Consumer Preferences for Cigarettes and Heated Tobacco Products in Japan: Evidence from a Discrete Choice Experiment* (No. w33301). National Bureau of Economic Research.

Kerr, D. (2004). Family transformations and the well-being of children: Recent evidence from Canadian longitudinal data. *Journal of Comparative family studies*, *35*(1):73-90.

Khantzian, E.J. (1997). The self-medication hypothesis of substance use disorders: a reconsideration and recent applications. *Harvard Review of Psychiatry, 4*:231–244.

Kluwe-Schiavon, B., Viola, T. W., Sanvicente-Vieira, B., Lumertz, F. S., Salum, G. A., Grassi-Oliveira, R., & Quednow, B. B. (2020). Substance related disorders are associated with impaired valuation of delayed gratification and feedback processing: a multilevel meta-analysis and meta-regression. *Neuroscience & Biobehavioral Reviews, 108*, 295-307.

Lewis, H.G. (1974). Comments on selectivity biases in wage comparisons. *Journal of Political Economy, 82*(6):1145-1155.

Liu, Y., Lan, Q., Siegfried, J.M., Luketich, J.D. and Keohavong, P. (2006). Aberrant promoter methylation of p16 and MGMT genes in lung tumors from smoking and never-smoking lung cancer patients. Neoplasia, 8(1),.46-51.

Madden, D. (2008). Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of health economics*, *27*(2):300-307.

Maphela, B. (2021). The Role of Local Economic Development and Environmental Health in the Emerging Funeral Home Sector in Mpumalanga, South Africa. *African Journal of Development Studies*, *11*(4).

McGinty, E. E., Alegria, M., Beidas, R. S., Braithwaite, J., Kola, L., Leslie, D. L., ... & Eisenberg, M. D. (2024). The Lancet Psychiatry Commission: transforming mental health implementation research. *The Lancet Psychiatry*, *11*(5), 368-396.

McGue, M., Elkins, L. & Lacono, W.G. (2000). Genetic and environmental influences on adolescent substance use and abuse. *American Journal of Medical Genetics*, *96*:671-677.

Mncwango, S. I., Adanlawo, E. F., & Vezi-Magigaba, M. F. (2024). Training interventions that enhance entrepreneurship self-efficacy among unemployed youth: A South Africa perspective. *International Journal of Business Ecosystem & Strategy (2687-2293)*, *6*(4), 113-121.

Moshoeshoe, R. (2012). Consumer demand for alcoholic beverages and tobacco in Lesotho: A double-hurdle approach. Economic Research Southern Africa (ERSA), Working paper, (315).

Munafo, M.R., Hitsman, B., Rende, R., Metcalfe, C. & Niaura, R. (2008). Effects of progression to cigarette smoking on depressed mood in adolescents: evidence from the National Longitudinal Study of Adolescent Health. *Addiction*, *103*:162–171.

Musa, S., Dergaa, I., Bachiller, V., & Saad, H. B. (2023). Global implications of COVID-19 pandemic on Adults' lifestyle behavior: the invisible pandemic of noncommunicable disease. *International Journal of Preventive Medicine*, *14*(1), 15.

Mwabu, G. (2009). The Production of Child Health in Kenya: A Structural Model of Birth Weight. *Journal of African Economies*, *18*(2):212-260.

Newman, C., Henchion, M. & Matthews, A. (2003). A double-hurdle model of Irish household expenditure on prepared meals. *Applied Economics*, *35*:1053-1061.

Nkomo, N. Y., & Adanlawo, E. F. (2023). The implications of population ageing on savings rates. *Management and Entrepreneurship: Trends of Development*, *2*(24), 8-16.

Nkomo, N. Y., & Adanlawo, E. F. (2024). The effects of education on tobacco consumption: A panel data analysis. *Journal of Ecohumanism*, *3*(3), 308-316.

Nyeleti, G. C., & Maphela, B. (2024). The Socio-Economic Effects of Delayed Rural Road Construction Projects: The Case of Malamulele. *Pakistan Journal of Life and Social Sciences, 22(2), 17970-17979.*

Pfeifer, C. (2012). A note on smoking behavior and health risks taking. *Nordic Journal of Health Economics, 1(2):135-151.*

Reynolds, R. M., Popova, L., Ashley, D. L., Henderson, K. C., Ntansah, C. A., Yang, B., ... & Thrasher, J. (2024). Messaging about very low nicotine cigarettes (VLNCs) to influence policy attitudes, harm perceptions and smoking motivations: a discrete choice experiment. *Tobacco control*, *33*(3), 325-332.

Ritt-Olson, A., Unger, J. & Valente, T. (2005). Exploring peers as a mediator of the association between depression and smoking in young adolescents. *Substance Use & Misuse*:77–98.

Rivers, D. & Vuong, Q.H. (1988). Limited Information Estimators and Exogeneity Tests for Simultaneous Probit. *Journal of Econometrics, 39*:347-366.

Rodriguez, L.A.G. & Ruigomez, A. (1999). Increased risk of irritable bowel syndrome after bacteria gastroenteritis: cohort study. *The BMJ, 318*(7183):565-566.

Rohde, P., Kahler, C.W., Lewinsohn, P.M. & Brown, R.A. (2004). Psychiatric disorders, familial factors, and cigarette smoking: II. Associations with progression to daily smoking. *Nicotine & Tobacco Research, 6*:119–132.

Semykina, A. & Wooldridge, J.M. (2010). Estimating Panel Data Models in the Presence of Endogeneity and Selection. *Journal of Econometrics, 157*:375-380.

Sreeramareddy, C.T. & Pradhan, P.M.S. (2015). Prevalence and social determinants of smoking in 15 countries from North Africa, Central and Western Asia, Latin America and Caribbean: secondary data analyses of demographic and health surveys. *PLoS One, 10*(7):e0130104.

Tomlinson, M., Grimsrud, A.T., Stein, D.J., Williams, D.R. & Myer, L. (2009). The epidemiology of major depression in South Africa: results from the South African Stress and Health study: mental health. *South African Medical Journal, 99*(5):368-373.

Verma, M., Rana, K., Bhatt, G., Sharma, N., & Lal, P. (2023). Trends and determinants of tobacco use initiation in India: analysis of two rounds of the Global Adult Tobacco Survey. *BMJ open*, *13*(9), e074389.

Wiesner, M. & Ittel, A. (2002). Relations of pubertal timing and depressive symptoms to substance use in early adolescence. *The Journal of Early Adolescence*, *22*(1):5-23.

Wodjao, T.B. (2007). A Double-Hurdle Model of Computer and Internet Use in American Households.

Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Xaba, F., Adanlawo, E. F., & Nkomo, N. Y. (2024). Are local communities participating in ecotourism projects?: a study of four local municipalities. *International Journal of Business Ecosystem & Strategy (2687-2293), 6*(3), 266-275.

Xaba, F., Adanlawo, E. F., & Chaka, M. (2025). Exploring higher education policies regulating breastfeeding on university campuses: implications for the wellbeing of women academics. *International Journal of Research in Business & Social Science, 14*(2).

Yen, S.T. (2005). Zero observations and gender differences in cigarette consumption. *Applied Economics, 37*:1839–1849

Yu, X. & Abler, D. (2008). *Family structure, education and cigarette smoking of the adults in China: A double-hurdle model.* University Park, PA: Penn State University.