The Demand for Cigarettes: New Evidence from South Africa

Alfred K. Mukong & Ernest Ngeh Tingum

ERSA working paper 745

May 2018

The Demand for Cigarettes: New Evidence from South Africa*

Alfred Kechia Mukong† & Ernest Ngeh Tingum‡

May 15, 2018

Abstract

This paper provides an extensive analysis of the demand for cigarettes based on longitudinal data drawn from the South Africa National Income and Dynamic Study (NIDS: 2008 - 2014). We compare the results of the pooled OLS (POLS), the standard two-part model, the random and the fixed effect (RE, FE) panel regression. Like previous evidence into cigarette prices, we obtain negative price elasticity of demand for cigarettes, with the conditional elasticity (POLS and RE estimates) significantly smaller than the total price elasticity (two-part model estimates). We find that over the same period, estimates from the fixed effect model are insignificant, properly due to the limited within variation in prices. However, it should be noted that estimates from the between variation models (RE and POLS) could be biased as they do not control for unobserved heterogeneity. Thus, with between variation models, increased tobacco taxes can, in the presence of the changing market structure, remain a desirable policy tool for reducing cigarette consumption.

Keywords: Cigarettes; Price; Elasticity; Demand; South Africa

JEL Code: C23, D12, I12, P22

*The authors gratefully acknowledge funding support from the International Development Research Centre (IDRC), the Medical Research Councils (MRC) of South Africa (Project number: 108442-002), from Cancer Research UK (CRUK) and the Carnegie foundation New York. We would also like to thank Corne Van Walbeek and Hana Ross for their constructive criticisms. We are also grateful for the valuable comments and suggestions received from the anonymous reviewer of the Economic Research Southern Africa (ERSA) working paper series. All views expressed in this article are those of the authors and do not necessarily represent the views of, and should not be attributed to the funders.

†Corresponding author: Post-doctoral Research Fellow, Economics of Tobacco Control (ETCP), South African Labour and Development Research Unit (SALDRU), School of Economics, University of Cape Town, Rondebosch 7701, Cape Town, South Africa. E-mail: alfred.mukong@uct.ac.za

‡Post-doctoral Research Fellow, Economics of Tobacco Control (ETCP), South African Labour and Development Research Unit (SALDRU), School of Economics, University of Cape Town. E-mail: ngehernest@gmail.com

1 Introduction

Cigarette taxes are known to be the most effective control policy in reducing cigarette consumption. However, smokers have no incentive to know the amount of tax on cigarettes and the effectiveness of tobacco taxes in reducing cigarette use is contingent on the degree of excise tax pass-through (Linegar and van Walbeek 2017). The demand for cigarettes has long been of interest not just in its own right, but because cigarette use has important and significant societal costs through among others, harmful effects on health, cost of medical care and productivity loss (Ezzati and Lopez 2003, Vallejo-Torres and Morris 2010, Mukong et al. 2017). Smoking is the leading cause of lung disease, with smokers 25 times more likely to develop lung cancer and 12 times more likely to die from chronic obstructive pulmonary disease (COPD) than non-smokers (USDOHS et al. 2014). There is overwhelming evidence that higher cigarette prices reduce tobacco use, with greater reductions among young people, women and those from more socioeconomically disadvantages groups (IARC 2011, Vellios and van Walbeek 2016). Much of the evidence suggests that the price elasticity of demand for cigarettes vary widely, with developing countries having higher price elasticity (Warner 1990, Chaloupka et al. 2000, Abedian and Jacobs 2001)[1].

A recent comprehensive review identified disturbingly few studies that used individual-level data to estimate the demand for cigarettes in developing countries (IARC 2011, NCI 2016). Studies that use aggregate time-series data are unable to examine the effect of price change on smoking participation and ignores the underlying effects of many important factors, particularly individual level characteristics. These variables are treated as constant and are assumed to have no bearing on the demand for cigarettes. There are reasons to believe that these factors, including the smoking addiction hypothesis, for example could play a strong role in determining the demand for cigarette (Becker and Murphy 1988). The exclusion of these variables simplifies the analysis of the demand for cigarettes, but potentially hides important information that may affect the estimates of price. Thus, the inclusion of these variables does not only ensure more accurate estimates of price elasticity but also provide evidence of their effect on cigarette use.

The purpose of this paper is to address this evidence gap using new data from South Africa. The analysis of panel data reduces the potential bias from aggregate data and allows for the estimation of the effect of price on smoking participation, and to test the important effect of individual characteristics on cigarette consumption. Longitudinal data are particularly useful in analysing how individual smoking behaviour changes over time and can answer more questions than aggregate data. With aggregate data, the degree

[1] Consumers from low income countries are more sensitive to price and are likely to have higher price elasticity (Warner 1990).

of correlation between independent variables is often high and may result in unstable estimates (IARC 2011). In addition, the market clearance price could be determined by the interaction of both demand and supply sides, and price may be endogenous (IARC 2011).

Relative to many low- and middle-income countries (LMICs), South Africa is noted for the use of large excise taxes to reduce cigarette consumption by almost half per adult within 15 years (see Jha and Peto 2014). In 1994, the total tax burden on cigarettes was 33%, increased to 50% in 1997 (Malan and Leaver 2003), to 52% in 2006, and remained unchanged until recently (see Linegar and van Walbeek 2017). The dominant producer and distributor of cigarette (British American Tobacco (BAT)) was an unchallenged price leader until 2010. With its pricing power, the real price of cigarette (net-of-tax) doubled and the retail price almost tripled between 1994 and 2010. The rapid increase in cigarette price among other factors reduced the smoking prevalence from a third to less than a fifth between 1994 and 2012[2]. In 2010, the cigarette market changed substantially as small manufacturers and distributors were attracted by the high profits of the multinationals. These small firms were selling at prices substantially lower than the economy price brands of the incumbents (Linegar and van Walbeek 2017). This also substantially increased the illicit market (van Walbeek 2015). While there is evidence of a tax pass-through in this new market structures (Linegar and van Walbeek 2017), there is need to unpack how the pass-through and the changing market structure affect the smoking behaviour of individuals.

The elasticity of demand for cigarettes is generally lower in high than in low income countries. In the United States (US) for example, the elasticity of demand for cigarettes has been between -0.005 to -0.62 (see Huang and Yang 2006). Using aggregate panel data, Coats (1995) estimated the price elasticity in the US to be between -0.005 and -0.016, while Yurekli and Zhang (2000) results suggest that the elasticity values are between -0.48 and -0.62. Estimates of price elasticity after this period remain within the range -0.005 to -0.62 (Baltagi and Griffin 2001, Huang et al. 2004). For developing countries, an analysis of 32 studies from Latin America concluded that on average, the price elasticity of demand for cigarettes is likely to be -0.5 (Guindon et al. 2016). Of these studies, the lower-bound was estimated to be between -0.1 and -0.3, the medium-bound between -0.3 and -0.6 and the higher-bound between -0.6 and -1.5. The higher-bound common among younger population group and in low-income countries. While evidence from the US and other developed countries made use of both time series, panel as well as cross-sectional data, research in many LMICsrely mostly on aggregated data.

[2]Shisana O, Labadarios D, Rehle T. South African national health and nutrition examination survey (SANHANES-1). Available from: http://www.hsrc.ac.za/en/ research-outputs/view/6493 (2013 17 April 2014)

In the context of the South African cigarette market, demand estimates have been inconsistent, with price elasticity found to vary over time. Using data for the period 1970 to 1989, Reekie (1994) estimated the price elasticity of demand to be -0.87. Using annual data from 1970 to 1990, Van Walbeek (1996) estimated long-run prices elasticity to lie between -0.53 to -1.52. Using data from 1970 to 1995, Van der Merwe and Annett (1998) reported the value of -0.69, while Van Walbeek (2000) results from 1970 to 1998 annual data suggest an estimate of -0.6. Most of these estimates are generally considered to be very high as the specifications do not account for many of the demand shift-factors (Abedian and Dorrington 1994, Boshoff 2008). Using quarterly data (1996-2006) and controlling for consumer preferences, Boshoff (2008) estimated the price elasticity to be between -0.5 to -0.7. For recent evidence on cigarette consumption using individual data, Vellios and van Walbeek (2016) find that an increase in cigarette prices reduce smoking initiation while Koch (2018) finds that tobacco taxes are less regressive (the burden for increasing cigarette prices falls less heavily on the poor). In addition, non-price factors have also been shown to play a significant role in determining smoking participation and intensity (Mukong 2017).

This paper considers individual level data to tease out new evidence on the price effect of demand for cigarettes in South Africa. We estimate the conditional and unconditional (total) price elasticity of demand for cigarettes. The conditional elasticities (pooled OLS and random effect estimates) are significantly smaller than the unconditional elasticities (two-part model estimates). However, findings from the fixed effect model suggest that price has no significant effect on cigarette consumption. Overall, models with between variation components (RE and POLS) do not control for unobserved heterogeneity and the effect of price on cigarette consumption are likely to be biased.

2 Data and Empirical Strategy

2.1 Data and Sample

The analysis uses longitudinal data from the South African National Income Dynamic Study (NIDS), which provide information on individual smoking patterns and background characteristics. The NIDS is an ongoing and nationwide representative panel survey, which target South African population aged 15 years and older[3]. The survey tracks the same household members every two years, to examine the livelihoods of individuals and households over time. The current analysis uses all the available four waves

[3]SALDRU (2016d), 'National Income Dynamics Study 2014-2015, Wave 4 [dataset]. Version 1.0', Cape Town: Southern Africa Labour and Development Research Unit [producer], Cape Town: DataFirst [distributor].

(2008 to 2014). The baseline sample 2008 provides information on 7,236 households with 16,781 adult individuals. The data used in our analysis covers the periods between 2008 and 2014 and the pooled data consist of a total of 142,045 observations. Of this sample, a total of 13,437 (these are unweighted sample, see Table 1 for estimates of the weighted sample) of the observations were smokers at the time of the surveys, but 14,934 provided information on their smoking intensity (the excess are former smokers). In the empirical analysis, we exclude former smokers who provided information on their smoking history. The final sample size varies by each specification and could be less than 13,437 after accounting for missing data (for conditional demand estimates only). Concerning the response rate between waves, 78% of individuals interviewed in Wave 1 were successfully interviewed in Wave 4, 84% interviewed in Wave 2 were successfully interviewed in Wave 4 and 92% interviewed in Wave 3 were successfully interviewed in Wave 4 (Chinhema et al. 2016). The attrition rate between Wave 1 and Wave 4 (22%) is consistently low for the possibility of systematic bias from the decline in response rate (Baigrie and Eyal 2014, Chinhema et al. 2016).

2.2 Measures

The NIDS uses the following questions to measure individual smoking behaviour: For current smokers, do you smoke cigarettes? For non-smokers, did you ever smoke cigarettes regularly? Both smokers and ex-smokers were asked the age at which they first smoked cigarettes and only ex-smokers were asked when they last smoked cigarettes regularly. Finally, individuals were asked to indicate the average number of cigarettes they smoked per day. Only individuals who smoke cigarettes are used in estimating smoking intensity (excluding non-smokers and former smokers). A smoker is defined as someone who smokes cigarettes at the time of the interview. Smoking intensity is defined as the average number of cigarettes an individual smokes per day.

Average prices of cigarette provided by Statistics South Africa are used. The cigarette prices are annual and have been deflated by the consumer price index (CPI) so that each cigarette price is in 2010 prices. The profitable cigarette market in South Africa attracted cheap competitors in 2010, causing real cigarette prices of the popular brands to remain fairly constant but budget cigarettes became cheaper in real terms. With this price war, consumers are faced with varying options, and rather than quitting or reducing consumption due to a price rise, they might switch to cheaper brands of cigarette. Our data does not allow for cross-price elasticity of demand for cigarette type and we capture the effects of this exogenous

shock by using an interrupted time-series dummy ("price war") for the market structural change in 2010. Regarding other regulatory policies, the latest tobacco amendment act in South Africa was in 2007/2008, just around our study period. This Act was signed at the beginning of 2009 and came into operation towards the end of 2009. While this is expected to have some bearing on our results, it is important to note that these policies coincided with the price war and the interrupted time-series dummy therefore captures the effects of both price war and regulatory policy change.

2.3 Method

The main purpose of the analysis is to estimate the price elasticity of demand for cigarettes. Let Q_i, be defined as the average quantity smoked per day, Y_i the participation decision for individual i, P the average price of cigarette (either economy or mid price), I_i household per capita income and X_i individual level characteristics that affect their cigarette consumption decision. The participation equation (equation 1) and the consumption equation (equation 2) are written as:

$$Y_i = \beta_0 + \beta_1 lnP + \beta_2 lnI_i + \beta_3 X_i + \varepsilon_i \tag{1}$$

$$lnQ_i = \beta_0 + \beta_1 lnP + \beta_2 lnI_i + \beta_3 X_i + \mu_i \tag{2}$$

The empirical analysis of equation 2 assume a log-log relationship between cigarette consumption and price and β_1 and β_2 are price and income elasticities respectively. lnQ_i is observed only when the individual is a current smoker (equation 1). While the FE panel regression could be the most appropriate (account for unobserved heterogeneity), nature our main variable (price) and the need to estimate the effect of some time-invariant variables such as gender explain why both the RE and the FE models are used. In South Africa, the period of rapid tax and price increases abruptly came to an end in 2010 and real prices have remained largely constant (see Panel B of Table 1). The within variation in price therefore is very low (Table 2), and using the fixed effect model will produce less efficient estimates and will not estimate time-invariant variables (see Plümper and Troeger 2007). Individuals are not large enough to influence the market price and the price variable is expected to be exogenous (IARC 2011).

The decision not to smoke could be an explicit one or absolute (under no circumstances would individuals change their mind to smoke, health reasons for example). Others would only participate; if cigarette

prices go below a certain threshold or their disposable income is increased for example. In this scenario, the error terms of equation 1 and equation 2 could be correlated, and estimates from equation 2 would be inconsistent. In this case, a type II Tobit model can be used to control for the sample selection endogeneity that relies on the functional form assumed for the error term of the selection equation (Heckley et al. 2017). This requires an identification of an exclusion restriction that in itself is a difficult empirical challenge as we cannot say with certainty which factors are associated with smoking participation and not the intensity decision. The Heckman two-step method (Heckman 1979) assumes the error term in equation 1 are normally distributed and the selection equation can be estimated using the probit model. The conditional mean for the intensity equation (lnQ_i), given the selection equation (Y_i) is:

$$E\left[Q|X, Y^* > 0\right] = X^{'}\beta_2 + \delta\lambda\left(X^{'}\beta_2\right) \tag{3}$$

Where $\lambda\left(.\right) = {}^{\phi(.)}/_{\Phi(.)}$ is the inverse Mills Ratio (IMR), Y^* is a latent variable for the smoking participation (Cameron and Trivedi 2005), δ, is the covariance of the error terms of the participation and intensity equation. Equation 2 is therefore estimated using equation 3 providing estimates of covariates conditional on positive outcomes ($Y^* > 0$). The Monte Carlo evidence shows that the two-part model generally produces better estimates than the correctly specified generalised tobit model without an exclusion restriction (Belotti et al. 2015). Since there is no exclusion restriction to identify equation 1, the two-part model is superior on precision ground and is an adequate way to model equation 3. We use the probit model to estimate the first part and the generalised linear model (GLM) to estimate the second part. Unlike the OLS, the GLM allows for greater flexibility and allows for response variables that have error distribution models other than a normal distribution.

3 Empirical Findings

3.1 Descriptive statistics

Table 1 presents descriptive statistics for both smoking behaviour and individual characteristics by sample wave. Smoking participation rate declined from 21% in 2008 to 18% in 2010, but remained stable in 2012 (20%) and 2014 (20%). The average number of cigarettes smoked daily remained stable over all waves with approximately 9 cigarettes per smoker per day. Similar patterns are visible for smoking duration. The average years of smoking per smoker remained stable at 20 over all waves. On average, household

monthly income range from R6,245 in 2008 to R10,351 in 2014. In 2008, approximately 37% of the sample were alcohol users. The proportion reduced to 34% in 2010 and increased to 36% in 2012 and 46% in 2014. There are less males in the sample (48% in 2008 and 49% between 2010 and 2014) than females.

Table 1: Mean characteristics by panel waves

Variable	2008		2010		2012		2014	
Panel A: Mean and standard deviation for individual attributes								
	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.
Smoking intensity	8.73	(7.41)	8.94	(7.69)	8.67	(7.37)	8.70	(7.80)
Smoking prevalence	0.21	(0.41)	0.18	(0.38)	0.20	(0.40)	0.20	(0.40)
Smoking duration (in years)	20.21	(14.20)	19.62	(13.72)	20.35	(14.02)	20.30	(14.43)
Household average monthly income	6,245	(11,005)	7,905	(21,561)	8,368	(14,845)	10,351	(27,201)
Individual ever drank alcohol	0.37	(0.48)	0.34	(0.47)	0.36	(0.48)	0.46	(0.50)
Individual is male	0.48	(0.50)	0.49	(0.50)	0.49	(0.50)	0.49	(0.50)
No formal education	0.16	(0.36)	0.18	(0.38)	0.19	(0.39)	0.17	(0.37)
Primary education	0.31	(0.46)	0.28	(0.45)	0.26	(0.44)	0.25	(0.43)
Secondary education	0.45	(0.50)	0.45	(0.50)	0.45	(0.50)	0.46	(0.50)
Tertiary education	0.08	(0.27)	0.09	(0.29)	0.10	(0.30)	0.12	(0.32)
Individual is age 15 - 24	0.33	(0.47)	0.33	(0.47)	0.31	(0.46)	0.29	(0.46)
Individual is age 25 - 34	0.25	(0.43)	0.25	(0.43)	0.26	(0.44)	0.27	(0.44)
Individual is age 35 - 44	0.19	(0.39)	0.19	(0.39)	0.19	(0.40)	0.20	(0.40)
Individual is age 45 - 54	0.14	(0.35)	0.14	(0.35)	0.14	(0.35)	0.14	(0.35)
Individual is age 55 - 64	0.09	(0.29)	0.09	(0.29)	0.10	(0.30)	0.10	(0.30)
Religion important to individual	0.88	(0.32)	0.90	(0.30)	0.90	(0.30)	0.92	(0.28)
Panel B: Cigarette prices by province	Economy prices				Mid price			
Variable	2008	2010	2012	2014	2008	2010	2012	2014
Western Cape	15,24	19,71	17,84	16,36	21,40	24,98	24,87	24,55
Eastern Cape	19,39	19,99	20,03	21,50	21,15	24,66	24,73	24,76
Northern Cape	17,73	20,01	20,59	23,37	21,17	24,44	24,53	24,73
Free State	17,02	19,86	19,71	19,55	21,41	24,85	24,97	24,68
KwaZulu Natal	18,15	19,72	19,62	19,94	21,27	25,56	25,85	26,03
North West	15,39	19,72	20,76	24,57	21,42	24,98	25,06	25,75
Gauteng	19.67	20,29	20,25	23,13	21,61	25,09	25,26	24,87
Mpumalanga	12,43	18,98	17,61	19,80	21,13	25,06	25,03	24,71
Limpopo	16,72	19,72	18,69	22.13	20,99	24,69	24,73	25,71
Average price	16,86	19,92	19,65	21,15	21,31	24,99	25,13	25,03

Table 2 presents summary statistics of the between and within variation. In viewing and interpreting the estimates of the regression analysis, it is important to note that the variation in both smoking behaviour and price is driven by between variance. The within variance is however large enough to allow for the estimation of the FE model (Table 2).

Table 2: Variance composition for smoking behaviour and price

Variables		Range	Mean	Standard deviation.
Smoking intensity	overall	1 - 60	7.65	6.56
	between			5.98
	within			3.29
Smoking prevalence	overall	0 - 1	0.18	0.39
	between			0.36
	within			0.17
Economy price	overall	12.43 - 24.57	19.47	2.01
	between			1.53
	within			1.51
Mid price	overall	20.99 - 26.03	24.38	1.63
	between			1.48
	within			0.87

3.2 Multiple regression results

Table 3 shows the random effect regression results. As expected, there is a negative and significant association between price and cigarette demand, and a positive and significant association with income. The relationship is consistent across all models (see Table 5 and 6 for comparison). The results are estimated separately for the economy price and the mid price and the estimates are consistent with the different prices. In general, the regression estimates indicate that between 2008 and 2014, the price estimates have been sensitive to other observable factors including, smoking duration and alcohol use (see column (2) to (6) of Table 3, 6 and 8). Turning to price and policy change, the abrupt end of the rapid tax and price increase in 2010 and the introduction of new tobacco policies in 2009, as assessed by the interrupted time-series, had a negative and significant association with smoking intensity (see Panel A of Table 8), with the exception of mid price specifications, which consistently showed insignificant effects.

Table 3: Regression results: Random effect (RE) models

Panel A: Economy Price						
	(1)	(2)	(3)	(4)	(5)	(6)
Log of average price of cigarette	-0.178***	-0.147***	-0.266***	-0.230***	-0.239***	-0.195***
	(0.050)	(0.051)	(0.049)	(0.051)	(0.051)	(0.063)
Individual smoking duration (in years)		0.008***		0.007***	0.007***	0.018***
		(0.001)		(0.001)	(0.001)	(0.001)
Log of household per capita income			0.112***	0.107***	0.105***	0.092***
			(0.006)	(0.006)	(0.006)	(0.008)
Individual drinks alcohol					0.071***	0.052***
					(0.016)	(0.018)
Constant	2.260***	2.028***	1.756***	1.555***	1.536***	1.410***
	(0.147)	(0.151)	(0.148)	(0.153)	(0.153)	(0.193)
Observations	12,700	11,534	12,694	11,529	11,513	8,535
Panel B: Mid Price						
	(1)	(2)	(3)	(4)	(5)	(6)
Log of average price of cigarette	-0.145*	-0.138*	-0.384***	-0.362***	-0.372***	-0.233**
	(0.080)	(0.083)	(0.081)	(0.084)	(0.084)	(0.092)
Individual smoking duration (in years)		0.008***		0.007***	0.007***	0.019***
		(0.001)		(0.001)	(0.001)	(0.001)
Log of household per capita income			0.113***	0.108***	0.107***	0.094***
			(0.006)	(0.007)	(0.007)	(0.008)
Individual drinks alcohol					0.069***	0.051***
					(0.016)	(0.018)
Constant	2.197***	2.031***	2.183***	2.016***	2.002***	1.566***
	(0.256)	(0.265)	(0.254)	(0.263)	(0.263)	(0.290)
Observations	12,700	11,534	12,694	11,529	11,513	8,535

Results presented in this table are from the Random effect estimation procedure. In column (6) other possible individual observable characteristics are included (see column (4) for Panel A and column (5) for Panel B in Table 7). Testing the null of no effect $***p \prec 0.01; **p \prec 0.05; *p \prec 0.10$.

Interestingly, price is less elastic after controlling for smoking addiction (see column (1) and (2) of Table 3 for comparison) and more elastic after controlling for income, alcohol consumption and individual factors (see column (1), (3), (5) and (6) of Table 3 for comparison). The POLS is considered as between unit estimation because coefficients only capture differences between units. The RE transformation process includes the between-unit and within-unit variation (Wooldridge 2010), and is placed between the POLS and FE approach. The RE model considers the panel structure but the estimates are not consistent if unobserved heterogeneity is present. The corresponding Hausman test does not reject the null hypothesis that the differences between the RE and FE estimates are not systematic $\left[Chi^2(15) = 63.63, p = 0.000\right]$. Therefore, FE is the consistent model as it controls for unobserved heterogeneity.

Table 4: Regression results: Fixed effect (FE) models.

	(1)	(2)	(3)	(4)	(5)	(6)
Log of average price of cigarette	-0.055	-0.040	-0.066	-0.054	-0.053	-0.100
	(0.103)	(0.108)	(0.103)	(0.108)	(0.108)	(0.144)
Individual smoking duration (in years)		0.009***		0.009***	0.009***	0.008***
		(0.002)		(0.002)	(0.002)	(0.002)
Log of household per capita income			0.035***	0.037***	0.036***	0.021
			(0.012)	(0.013)	(0.013)	(0.016)
Individual drinks alcohol					0.052**	0.054*
					(0.023)	(0.028)
Time trend						
2008						
2010	0.082***	0.059**	0.077***	0.055**	0.056**	0.088**
	(0.026)	(0.028)	(0.026)	(0.028)	(0.028)	(0.035)
2012	0.041*	0.003	0.025	-0.014	-0.014	0.036
	(0.023)	(0.026)	(0.024)	(0.027)	(0.027)	(0.035)
2014	0.084***	0.030	0.058**	0.003	-0.000	0.048
	(0.026)	(0.030)	(0.027)	(0.032)	(0.032)	(0.047)
Constant	1.861***	1.670***	1.669***	1.471***	1.434***	1.801***
	(0.291)	(0.308)	(0.298)	(0.315)	(0.315)	(0.446)
Observations	12,700	11,534	12,694	11,529	11,513	8,535
R-squared	0.004	0.010	0.005	0.012	0.013	0.015

Results presented in this table are from the Fixed effect estimation procedure. In column (6) other possible individual observable characteristics such as gender, education, age, marital status and religion are included. Testing the null of no effect $***p \prec 0.01; **p \prec 0.05; *p \prec 0.10$.

In contrast to the POLS and the RE estimates, the FE estimates show that changes in cigarette prices does not cause any significant change in cigarette consumption. That is, the corresponding estimates for price are not significant (see Table 3, 4 and 6 for comparison). It should be noted that although the FE model completely eliminates unobserved heterogeneity, in many cases, this advantage also comes at a cost (Plümper and Troeger 2007). The first drawback is that the FE model could be inefficient in estimating the effect of variables with very little within variance. This inefficiency produces higher standard errors and highly unreliable point estimates, leading to wrong inferences (Plümper and Troeger 2007). This could be the case in our study as the price data shows very little within variance, a result of the fact that individuals within the same province are faced with the same price (See Panel B of Table 1). In addition, the FE model uses only within variation, disregarding between variation and limiting the inclusion of time-invariant variables (Baltagi and Griffin 2001, Wooldridge 2010). Although not in the area of tobacco, several studies have found null effect in the FE model (estimates of FE model not significant) when estimates of the POLS and the RE are statistically significant or when there is limited variation in the key variable (Ku et al. 2013, Ates 2017).

Table 5: Regression results: Two-part model estimates including individual characteristics

	Mid price	Economy price
	(1)	(3)
Log of average price of cigarette	-0.691***	-0.429***
	(0.100)	(0.056)
Log of household per capita income	0.115***	0.111***
	(0.007)	(0.007)
Individual drinks alcohol	0.055***	0.058***
	(0.017)	(0.017)
Constant	2.905***	2.235***
	(0.294)	(0.173)
Observations	13,689	13,689

Results presented in this table are estimates from the Two-part model. Individual characteristics such as gender, education, age, marital status and religion are included. Testing the null of no effect $***p \prec 0.01; **p \prec 0.05; *p \prec 0.10$.

Estimates from the Two-part model are presented in Table 5. For the total elasticity, a 10% increase in the price of cigarette reduces cigarette consumption by 4.3% for economy brands and 6.9% for mid price brands. For POLS and the random effect estimates, a 10% increase in the price of cigarette reduced consumption of the economy brands by 2% (column (1) and (4) of Table 7) and mid price brand consumption between 2.3% to 2.9% (column (2) and (5) of Table 7). A 10% increase in household per capita increases cigarette use by 1.2%. Generally, the total price elasticity of demand for cigarettes is more than twice the conditional price elasticity of demand for cigarettes (see Table 5 and 7 for comparison).

In column (3) and (6) of Table 7, we present marginal effect estimates for smoking participation. The probability of smoking participation decreases with cigarette price and increases with income. While it is interesting that the estimates are statistically significant across the different estimation approaches, the estimate from the pooled OLS is larger than the estimate from the random effect model. The results suggest that drinkers are more likely to smoke than non-drinkers. Individual characteristics are also important in explaining the smoking participation probability and smoking intensity. For example, men are between 14% and 17% more likely to smoke than women (Table 7). Individuals with secondary and tertiary education are less likely to participation than those with no formal education. Religious conscious individuals are less likely to smoke than religiously less conscious individuals.

4 Discussion

This paper has broadened the analysis of the demand for cigarette using a large individual level data for South Africa. Longitudinal data are particularly useful in analysing changes in individual smoking

behaviour and can answer more questions than aggregate data (address most of the evidence gap from aggregate data). The time period in the analysis is interesting as it is a period where the rapid tax and price increases abruptly came to an end and with changing market structure. In the post-apartheid South Africa, tobacco excise tax increases has been used as one of the major control tool for reducing cigarette consumption (Blecher 2015). Between this period and up to 2010, cigarette was one of the consumer goods that experienced the largest price increases (see Linegar and van Walbeek 2017). In 2010, the situation changed drastically as the highly profitable net-of-tax price attracted competitors, changing the competitive environment, reducing the power of the incumbent firms to raise retail price and under-shifting the excise tax increases. This could render excise tax increases less effective in reducing cigarette consumption. It is evident from the 2010 developments that the degree of tobacco excise tax pass-through, and discretionary increases in cigarette prices, depends on the competitive nature of the market. It is important to extend this analysis to understand consumer's responsiveness to price changes in this new market structure.

We find negative price elasticity for both smoking prevalence and smoking intensity. The total price elasticity estimates are generally similar to those from previous studies in South Africa and some low- and middle-income countries. This observed similarities suggest that in this new market (from a near monopoly to a more competitive market), price increases remain a significant policy tool for reducing both smoking prevalence and the daily cigarette consumption of continued smokers. As expected, total price elasticity estimates are significantly higher than those of the conditional demand for cigarettes. The fixed effect model did not provide any evidence for a significant relationship between price and cigarette consumption. One would conclude that previous findings that considered only one of these approaches (the RE and FE panel regression or pooled OLS) to tease out the effect of price on cigarette use should be evaluated very critically.

However, in interpreting and making causal inferences of the described and discussed results, some limitations embedded in the analysis should be considered. The analysis only includes four panel waves and average annual prices by province. This limits the within variation in the model and may explain why the FE estimates are not significant, although descriptive statistics indicate that the amount of within-unit variance is adequate. The second limitation is that the panel is not balanced suggesting the possibility of panel attrition and explain why we employ the pooled OLS estimation approach. However, evidence from this data indicates that the attrition rate is consistently low for the possibility of systematic bias (Baigrie and Eyal 2014, Chinhema et al. 2016).

To conclude, the findings indicate that in the presence of a changing market structure, tobacco excise tax increases remain a desirable policy tool for reducing cigarette consumption. The traditional economic analysis of cigarette demand considers the demand for total cigarette consumed in a given period (smoking intensity). This approach assumes that the frequency (how often and individual smokes) and smoking intensity (how many cigarettes smoked at each given period) have no bearing on the utility an individual receives (frequency and intensity are treated as perfect substitutes). Understanding the separate effects of a policy change on frequency and intensity decision is important since both affect the overall quantity decision. Therefore, further research must focus on the effect of price on smoking frequency.

References

Abedian, I. and Dorrington, R.: 1994, An evaluation of a recent attempt to assess the social benefits of cigarette smoking, *Journal for Studies in Economics and Econometrics* **18**(1), 59–72.

Abedian, I. and Jacobs, R.: 2001, Tobacco taxes and government revenue in south africa, *Journal of Economic Studies* **28**(6), 397–407.

Ates, M.: 2017, Does grandchild care influence grandparents' self-rated health? evidence from a fixed effects approach, *Social Science & Medicine* .

Baigrie, N. and Eyal, K.: 2014, An evaluation of the determinants and implications of panel attrition in the national income dynamics survey (2008-2010), *South African Journal of Economics* **82**(1), 39–65.

Baltagi, B. H. and Griffin, J. M.: 2001, The econometrics of rational addiction: the case of cigarettes, *Journal of Business & Economic Statistics* **19**(4), 449–454.

Becker, G. S. and Murphy, K. M.: 1988, A theory of rational addiction, *Journal of political Economy* **96**(4), 675–700.

Belotti, F., Deb, P., Manning, W. G., Norton, E. C. et al.: 2015, twopm: Two-part models, *Stata J* **15**(1), 3–20.

Blecher, E.: 2015, Taxes on tobacco, alcohol and sugar sweetened beverages: Linkages and lessons learned, *Social Science & Medicine* **136**, 175–179.

Boshoff, W. H.: 2008, Cigarette demand in south africa over 1996-2006: The role of price, income and health awareness, *South African Journal of Economics* **76**(1), 118–131.

Cameron, A. C. and Trivedi, P. K.: 2005, *Microeconometrics: methods and applications*, Cambridge university press.

Chaloupka, F. J., Hu, T.-w., Warner, K. E., Jacobs, R. and Yurekli, A.: 2000, The taxation of tobacco products, *In: Jha, P.; Chaloupka, F.(Eds.). Tobacco Control in Developing Countries*, Citeseer.

Chinhema, M., Brophy, T., Brown, M., Leibbrandt, M., Mlatsheni, C. and Woolard, I.: 2016, *National Income Dynamics Study Panel User Manual*.

Coats, R. M.: 1995, A note on estimating cross-border effects of state cigarette taxes, *National Tax Journal* pp. 573–584.

Ezzati, M. and Lopez, A. D.: 2003, Estimates of global mortality attributable to smoking in 2000, *The Lancet* **362**(9387), 847–852.

Guindon, G. E., Paraje, G. R. and Chaloupka, F. J.: 2016, The impact of prices and taxes on the use of tobacco products in latin america and the caribbean, *Revista Panamericana de Salud Pública* **40**(4), 272–284.

Heckley, G., Jarl, J. and Gerdtham, U.-G.: 2017, Frequency and intensity of alcohol consumption: new evidence from sweden, *The European Journal of Health Economics* **18**(4), 495–517.

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

Huang, B.-N. and Yang, C.-W.: 2006, Demand for cigarettes revisited: an application of the threshold regression model, *Agricultural Economics* **34**(1), 81–86.

Huang, B.-N., Yang, C.-w., Hwang, M.-j. et al.: 2004, New evidence on demand for cigarettes: a panel data approach, *International Journal of Applied Economics* **1**(1), 81–97.

IARC: 2011, *Effectiveness of tax and price policies for tobacco control*, World Health Organization, International Agency for Research on Cancer.

Jha, P. and Peto, R.: 2014, Global effects of smoking, of quitting, and of taxing tobacco, *New England Journal of Medicine* **370**(1), 60–68.

Koch, S. F.: 2018, Quasi-experimental evidence on tobacco tax regressivity, *Social Science & Medicine* **196**, 19–28.

Ku, L.-J. E., Stearns, S. C., Van Houtven, C. H., Lee, S.-Y. D., Dilworth-Anderson, P. and Konrad, T. R.: 2013, Impact of caring for grandchildren on the health of grandparents in taiwan, *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* **68**(6), 1009–1021.

Linegar, D. J. and van Walbeek, C.: 2017, The effect of excise tax increases on cigarette prices in south africa, *Tobacco Control* pp. tobaccocontrol–2016.

Malan, M. and Leaver, R.: 2003, Political change in south africa: New tobacco control and public health policies, *In: Beyer de J, Brigden W, eds. Tobacco Control policy: Strategies, Successes & Setbacks. Washington and Ottawa: World Bank and Research for international Tobacco Control, Chapter 6* pp. 121–153.

Mukong, A. K.: 2017, Peer networks and tobacco consumption in south africa, *South African Journal of Economics* **85**(3), 341–367.

Mukong, A. K., Van Walbeek, C., Ross, H. et al.: 2017, Lifestyle and income-related inequality in health in south africa, *Int J Equity Health* **16**(103).

NCI: 2016, The economics of tobacco and tobacco control. national cancer institute tobacco control monograph 21., *Technical report*, NIH Publication No. 16-CA-8029A. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute; and Geneva, CH: World Health Organization;.

Plümper, T. and Troeger, V. E.: 2007, Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects, *Political Analysis* **15**(2), 124–139.

Reekie, W. D.: 1994, Consumers' surplus and the demand for cigarettes, *Managerial and Decision Economics* **15**(3), 223–234.

USDOHS et al.: 2014, The health consequences of smoking 50 years of progress: a report of the surgeon general, *Atlanta, GA: US Department of Health and Human Services (DOHS), Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health* **17**.

Vallejo-Torres, L. and Morris, S.: 2010, The contribution of smoking and obesity to income-related inequalities in health in england, *Social Science & Medicine* **71**(6), 1189–1198.

Van der Merwe, R. and Annett, N.: 1998, The effects of taxation on consumption in south africa, *The Economics of Tobacco Control in South Africa. Report submitted to the International Tobacco Initiative* .

Van Walbeek, C.: 2000, Impact of the recent tobacco excise tax increases on the future government revenue potential in south africa, *Economics of Tobacco Control Project* .

van Walbeek, C.: 2015, The economics of tobacco control (part 2): evidence from the itc project.

Van Walbeek, C. P.: 1996, Excise taxes on tobacco: How much scope does the government have?, *South African Journal of Economics* **64**, 20–42.

Vellios, N. and van Walbeek, C.: 2016, Determinants of regular smoking onset in south africa using duration analysis, *BMJ open* **6**(7), e011076.

Warner, K. E.: 1990, Tobacco taxation as health policy in the third world., *American journal of public health* **80**(5), 529–531.

Wooldridge, J. M.: 2010, *Econometric analysis of cross section and panel data*, MIT press.

Yurekli, A. A. and Zhang, P.: 2000, The impact of clean indoor-air laws and cigarette smuggling on demand for cigarettes: an empirical model, *Health economics* **9**(2), 159–170.

Table 6: Pooled OLS estimates for price elasticity of demand for cigarette

Panel A: Economy price						
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Log of average price of cigarette	-0.305***	-0.256***	-0.232***	-0.197***	-0.206***	-0.182**
	(0.062)	(0.064)	(0.061)	(0.063)	(0.063)	(0.080)
Individual smoking duration (in years)		0.007***		0.006***	0.006***	0.020***
		(0.000)		(0.000)	(0.000)	(0.001)
Log of household per capita income			0.145***	0.139***	0.138***	0.117***
			(0.006)	(0.006)	(0.006)	(0.008)
Individual drinks alcohol					0.075***	0.052***
					(0.016)	(0.018)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.648***	2.358***	1.476***	1.283***	1.258***	1.261***
	(0.177)	(0.181)	(0.180)	(0.184)	(0.184)	(0.238)
Observations	12,700	11,534	12,694	11,529	11,513	8,535
R-squared	0.005	0.026	0.047	0.065	0.067	0.095
Panel B: Mid price						
	(1)	(2)	(3)	(4)	(5)	(6)
Log of average price of cigarette	-0.354***	-0.248***	-0.569***	-0.472***	-0.480***	-0.291***
	(0.091)	(0.093)	(0.089)	(0.092)	(0.092)	(0.100)
Individual smoking duration (in years)		0.008***		0.007***	0.007***	0.021***
		(0.000)		(0.000)	(0.000)	(0.001)
Log of household per capita income			0.136***	0.131***	0.129***	0.111***
			(0.006)	(0.006)	(0.006)	(0.008)
Individual drinks alcohol					0.066***	0.049***
					(0.016)	(0.018)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Constant	2.878***	2.399***	2.629***	2.227***	2.211***	1.672***
	(0.288)	(0.298)	(0.283)	(0.293)	(0.293)	(0.317)
Observations	12,700	11,534	12,694	11,529	11,513	8,535
R-squared	0.001	0.023	0.040	0.059	0.060	0.091

Results presented in this table are estimates from the Pooled OLS estimation procedure. In column (6) other possible individual observable characteristics are included (see column (1) for Pane A and column (2) for Panel B in Table **??**). Testing the null of no effect $***p \prec 0.01; **p \prec 0.05; *p \prec 0.10$.

Table 7: Pooled OLS and Random effect estimates including individual characteristics

	Pooled OLS and probit			Random effect (RE) and RE probit		
	(1)	(2)	(3)	(4)	(5)	(6)
Log of average price of cigarette	-0.202***	-0.291***	-0.208***	-0.195***	-0.233**	-0.175***
	(0.065)	(0.100)	(0.013)	(0.063)	(0.092)	(0.013)
Individual smoking duration (in years	0.020***	0.021***		0.018***	0.019***	
	(0.001)	(0.001)		(0.001)	(0.001)	
Log of household per capita income	0.109***	0.111***	0.008***	0.092***	0.094***	0.007***
	(0.008)	(0.008)	(0.002)	(0.008)	(0.008)	(0.002)
Individual drink alcohol	0.050***	0.049***	0.225***	0.052***	0.051***	0.182***
	(0.018)	(0.018)	(0.003)	(0.018)	(0.018)	(0.003)
Individual is a male	0.004	0.002	0.141***	0.019	0.016	0.169***
	(0.017)	(0.017)	(0.003)	(0.020)	(0.020)	(0.004)
General education	0.086***	0.087***	0.043***	0.089***	0.090***	0.034***
	(0.030)	(0.030)	(0.006)	(0.033)	(0.033)	(0.007)
Secondary education	0.176***	0.177***	-0.013**	0.181***	0.183***	-0.019***
	(0.030)	(0.030)	(0.006)	(0.033)	(0.033)	(0.007)
Tertiary education	0.237***	0.237***	-0.087***	0.231***	0.231***	-0.071***
	(0.042)	(0.042)	(0.008)	(0.044)	(0.045)	(0.009)
Individual is age 25 - 34	-0.053**	-0.054**	0.072***	-0.031	-0.031	0.068***
	(0.025)	(0.025)	(0.004)	(0.025)	(0.025)	(0.005)
Individual is age 35 - 44	-0.175***	-0.179***	0.092***	-0.131***	-0.133***	0.089***
	(0.032)	(0.032)	(0.005)	(0.033)	(0.033)	(0.006)
Individual is age 45 - 54	-0.328***	-0.332***	0.103***	-0.265***	-0.267***	0.096***
	(0.041)	(0.041)	(0.006)	(0.041)	(0.041)	(0.006)
Individual is age 55 - 64	-0.510***	-0.516***	0.085***	-0.413***	-0.415***	0.080***
	(0.052)	(0.052)	(0.006)	(0.053)	(0.053)	(0.007)
Individual living with Partner	-0.054**	-0.055**	0.064***	-0.035	-0.036	0.058***
	(0.023)	(0.023)	(0.005)	(0.025)	(0.025)	(0.006)
Individual is a widow/widower	-0.002	-0.006	0.016*	-0.000	-0.004	0.015*
	(0.045)	(0.045)	(0.008)	(0.046)	(0.046)	(0.009)
Individual is divorced	0.081*	0.079*	0.057***	0.066	0.066	0.053***
	(0.046)	(0.046)	(0.010)	(0.048)	(0.048)	(0.011)
Individual is never married	-0.068***	-0.068***	0.028***	-0.058**	-0.057**	0.022***
	(0.021)	(0.021)	(0.004)	(0.023)	(0.023)	(0.005)
Religion is important to individual	-0.081***	-0.080***	-0.032***	-0.079***	-0.078***	-0.024***
	(0.021)	(0.021)	(0.005)	(0.020)	(0.020)	(0.004)
Constant	1.351***	1.672***		1.410***	1.566***	
	(0.198)	(0.317)		(0.193)	(0.290)	
Observations	8,535	8,535	51,362	8,535	8,535	51,362
R-squared	0.091	0.091				

Estimates in column (1) and (4) are respectively pooled OLS and random effect estimates using economy price, column (2) and (5) are respectively pooled OLS and random effect estimates using the mid price, column (3) and (6) are estimates for smoking participation. The reference category male, individual does not drink alcohol, individual is female, no formal education, individual is aged 15 - 24, married and religion is unimportant to individual. Testing the null of no effect $***p \prec 0.01; **p \prec 0.05; *p \prec 0.10$.

Table 8: Regression results: Pooled OLS and Random effect (RE) estimates including possible shocks

Panel A: Pooled OLS estimates using economy price						
	(1)	(2)	(3)	(4)	(5)	(6)
Log of average price of cigarette	-0.303***	-0.245***	-0.259***	-0.210***	-0.221***	-0.151*
	(0.062)	(0.063)	(0.060)	(0.062)	(0.062)	(0.080)
Price war/regulations	0.004	0.009	-0.043**	-0.036**	-0.036**	-0.022
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.020)
Individual smoking duration (in years)		0.007***		0.007***	0.007***	0.020***
		(0.000)		(0.000)	(0.000)	(0.001)
Log of household per capita income			0.135***	0.130***	0.129***	0.110***
			(0.006)	(0.006)	(0.006)	(0.008)
Individual drinks alcohol					0.068***	0.049***
					(0.016)	(0.018)
Constant	2.642***	2.327***	1.620***	1.379***	1.365***	1.210***
	(0.174)	(0.179)	(0.177)	(0.182)	(0.182)	(0.237)
Observations	12,700	11,534	12,694	11,529	11,513	8,535
R-squared	0.003	0.024	0.040	0.060	0.061	0.091
Panel B: Random effect estimates using economy price						
	(1)	(2)	(3)	(4)	(5)	(6)
Log of average price of cigarette	-0.233***	-0.181***	-0.228***	-0.178***	-0.187***	-0.177**
	(0.062)	(0.064)	(0.061)	(0.063)	(0.063)	(0.081)
Price war/regulations	0.024	0.015	-0.018	-0.024	-0.024	-0.007
	(0.016)	(0.017)	(0.016)	(0.017)	(0.017)	(0.019)
Individual smoking duration (in years)		0.008***		0.007***	0.007***	0.018***
		(0.001)		(0.001)	(0.001)	(0.001)
Log of household per capita income			0.113***	0.108***	0.106***	0.093***
			(0.006)	(0.007)	(0.007)	(0.008)
Individual drinks alcohol					0.071***	0.052***
					(0.016)	(0.018)
Constant	2.405***	2.115***	1.649***	1.409***	1.391***	1.357***
	(0.176)	(0.181)	(0.179)	(0.184)	(0.184)	(0.242)
Observations	12,700	11,534	12,694	11,529	11,513	8,535

Results presented in this table are estimates from the Pooled OLS and random effect estimation procedures. In column (6) other possible individual observable characteristics such as gender, education, age, marital status and religion are included. Testing the null of no effect $***p \prec 0.01; **p \prec 0.05; *p \prec 0.10$.