The effect of cigarette price changes on smoking prevalence by gender: the case of South Africa

Zachary Gitonga, Nicole Vellios, Corne van Walbeek

ERSA working paper 859

April 2021

The effect of cigarette price changes on smoking prevalence by gender: the case of South Africa

Zachary Gitonga [*], Nicole Vellios[†], Corné van Walbeek[‡§]

April 19, 2021

Abstract

South Africa successfully reduced smoking prevalence by substantially increasing tobacco excise tax and therefore real cigarette prices between 1993 and 2010. Tobacco market structure changed in 2010 following the entry of local tobacco companies and the introduction of cheaper cigarette brands. Illicit cigarettes have also increased significantly. This paper estimates price elasticities of smoking prevalence by gender and examines the effect of an increase in illicit cigarettes and changes in tobacco market structure on smoking behavior in South Africa. Two nationally representative longitudinal data sets and cigarette price data from Statistics South Africa, are used. We use pooled fractional probit correlated random effects and panel LPM models for estimation. Smoking prevalence and price sensitivity are higher among males than among females. Price elasticity of smoking prevalence is about -0.33 overall, -0.43 for males and -0.20 for females. The increase in illicit cigarettes and availability of cheaper brands reduce the effect of price on smoking prevalence and undermine tobacco control policy. The relatively price-inelastic demand implies that there is room for an increase in excise tax on cigarette. We recommend a further increase in excise taxes on tobacco and implementing a track and trace system to control illicit trade.

[*]Corresponding Author. Email address: zachgitonga@gmail.com

[†]Email address: nicolevellios@gmail.com

[‡]Email address: cornelis.vanwalbeek@uct.ac.za

[§]Research Unit on the Economics of Excisable Products (REEP), School of Economics, University of Cape Town. Private, Rondebosch, Cape Town, 7701. Tel: (021) 650–5751.

Key words: Price elasticity, smoking prevalence, illicit trade, fractional probit, South Africa, structural break, tobacco control

Acknowledgment

The authors thank Micheal Boachie and an anonymous reviewer for their valuable critique of an earlier version of this paper. Sincere gratitude to Statistics South Africa for providing price data

Funding

This research was supported by the International Development Research Centre under the Global Alliance for Chronic Diseases research programme (grant ID number: 108442-002 and IRMA 25761)

Data availability

NIDS and AMPS data is publicly available through DataFirst (www.datafirst.uct.ac.za)

1 Introduction

Article 6 of the World Health Organization's (WHO) Framework Convention on Tobacco Control (FCTC) encourages Parties to increase tax and prices to reduce the demand for tobacco products and enhance public health (International Agency for Research on Cancer, 2011; World Health Organization, 2003). WHO provides guidelines to help FCTC parties to implement article 6 (World Health Organization, 2014). The global target set under the Noncommunicable Diseases Global Action Plan 2013-2020 is a 30% global reduction in smoking prevalence by 2025, relative to 2010 levels (World Health Organization, 2018). Empirical research has shown that the most effective way to reduce tobacco consumption and enhance public health is a substantial increase in excise tax and prices (Chaloupka et al., 2011; International Agency for Research on Cancer, 2011). A reduction in tobacco consumption implies a reduction in smoking intensity (average number of cigarettes smoked per day) and/or a reduction in smoking prevalence. Since all levels of smoking exposure are likely associated with lasting and progressive lung damage (Oelsner et al., 2020), smoking prevalence, rather than smoking intensity, is a more appropriate variable to follow when looking at the public health effects of, say, a change in tobacco prices or taxes. A reduction in smoking prevalence is achieved if current smokers quit and if young people do not become regular smokers (International Agency for Research on Cancer, 2011; World Health Organization, 2010). However, these gains in consumption reduction can easily be eroded by the availability of low-cost brands and illicit tobacco products in the market, which undermine both the fiscal and health goals (Jha & Peto, 2014; van der Zee et al., 2019).

Adopting a strong tobacco control stance at the dawn of democracy in

1994, South Africa was one of the first middle-income countries to effectively use taxation to reduce smoking prevalence. In 1994, the government set a target to raise the total tax burden[1] to 50% of the retail price, from its level of about 33% at the time (which was achieved in 1997), and increased the target to 52% in 2004 (Treasury & Manuel, 2004). Since the tobacco industry over-shifted[2] excise tax increases, retail prices increased sharply, resulting in a significant drop in smoking prevalence between the early 1990s and 2010 (Linegar & van Walbeek, 2018). The largest tax growth occurred between 1995 and 2011 when the real excise tax increased at an average rate of 9.7% per year, but it slowed markedly to 1% per year between 2011 and 2017 (Vellios et al., 2019).

Before 2010, the market was quite ordered. Historically, British American Tobacco South Africa (BATSA) has dominated the market and maintained the lead in manufacturing and distribution of tobacco products in South Africa. Other multinational companies like Philip Morris International (PMI) and Japan Tobacco International (JTI) accepted BATSA's price leadership and followed passively. The whole atmosphere changed when the small local producers gradually came into the market from about 2010 onward and they did not play by the established rules. Post-2010, the structure of the cigarette market changed from a near-monopoly to a more competitive market. The new firms introduced cheaper brands and are largely responsible for the increase in illicit trade in cigarettes (van der Zee et al., 2020; Vellios et al., 2019). Although the South African Revenue Service was starting to make headway in reducing illicit trade in 2014, a devastating management crisis that started at the end of 2014 eroded these gains. Five specialized units that were investigating the illegal cigarette market were disbanded (Du Toit, 2018). With no oversight, the illicit market expanded. Manufacturers were able to evade excise taxes, resulting in cigarettes sold at very low prices. In order to maintain their market shares, tobacco firms under-shifted excise taxes, i.e. increased the retail price by less than the increase in the excise tax, absorbing some of the tax burden (Linegar & van Walbeek, 2018).

Post 2010, National Treasury increased the nominal excise tax roughly in line with the inflation rate. Between 2010 and 2020 the real excise tax rate increased by less than 2% annually (Budget Reviews, various years). Since many manufacturers evaded excise taxes, excise taxes increases become

[1] Tax burden refers to the percentage total tax component (excise + sales taxes) of the retail price of the most-sold popular cigarette brand.

[2] Overshifting describes an activity whereby the industry increases the price of cigarette by an amount that exceeds the excise tax. By doing this, the industry increases their profit margins and is compensated for any reduction in sales resulting from higher prices (Ross et al., 2016).

less effective. The illicit market worsened following the COVID-19-related tobacco sales ban, which was in place in South Africa from 27 March 2020 to 17 August 2020. The rationale for the ban was to reduce the pressure on the health sector in light of the COVID-19 pandemic. According to Euromonitor, BATSA dominated the cigarette market with a 71% market share followed by Japan Tobacco International (JTI) with 12.2% share and Philip Morris International (8.9%) in 2019, before this disruption (sales ban), (Euromonitor International, 2020). Local companies like Gold Leaf Tobacco Corporation, Carnilinx, and Amalgamated Tobacco Manufacturers shared the remainder. Some of the smaller tobacco companies have a significant foothold in the illicit market, which is not reflected in these market shares, (Van Walbeek et al 2020).

So far, none of the reviewed studies have investigated how the increase in illicit trade in cigarettes and introduction of cheaper brands in South Africa since 2010 has affected smoking behaviour. To our knowledge, no study has estimated the price elasticity of smoking prevalence disaggregated by gender and race in Africa, and specifically in South Africa. The current study aims to fill this gap. We analyze the sensitivity of smoking prevalence to changes in cigarette prices by gender.

Selected price elasticity of tobacco demand studies

Most studies that estimate price elasticity of tobacco demand, using cross-sectional data, do so using a two-part model in which individuals are assumed to be faced with two sequential decisions. The first decision is about whether to smoke (i.e. smoking participation or smoking prevalence) and the second involves deciding the number of cigarettes to smoke (smoking intensity). The first part (price elasticity of smoking prevalence[3]) is usually estimated by either logit or probit. The second part (price elasticity of smoking intensity[4]) is estimated using Ordinary Least Squares (OLS) or generalized linear models (GLM) (International Agency for Research on Cancer, 2011).

Although there has been an increase in the number of studies about the responsiveness of tobacco demand to price changes in low- and middle-income countries (LMICs), those focusing on the price elasticity of smoking prevalence are scarce (Gallego et al., 2020). Most studies in LMICs estimate price elasticity of smoking intensity rather than prevalence. For instance, price elasticity of smoking intensity is estimated to be between -0.78 and -0.44 in Colombia (James et al., 2019; Maldonado et al., 2016), -0.77 in El Salvador (Paraje et al., 2020), -0.64 in Serbia (Vladisavljevic et al., 2020), -1.37 in

[3]Price elasticity of smoking prevalence measures the extent to which smoking prevalence (participation) changes for a given percentage change in cigarette price.

[4]Price elasticity of smoking intensity measure the extent to which demand for cigarette (consumption) changes for a given percentage change in cigarette prices

Bosnia and Herzegovina (Gligorić et al., 2020), between -0.26 and -0.33 in Uganda (Chelwa & van Walbeek, 2019), between -0.30 and -0.72 in South Africa (Boachie & Ross, 2020; Mukong & Tingum, 2020; Tingum et al., 2020), and -1.06 in Pakistan (Nayab et al., 2020). The average price elasticity of tobacco demand in LMICs is about -0.5 (Vladisavljevic et al., 2020).

Most studies on price elasticity of smoking prevalence have been conducted in developed countries (the USA, the UK, Canada, Korea, and Australia), but relatively few studies exist for LMICs. Studies conducted in the USA since the 1980s show that the estimates of price elasticity of smoking prevalence vary considerably, ranging from -0.54 to -0.05. The responsiveness to price changes differs by gender and other social economic variables (Farrelly et al., 2001; Lewit & Coate, 1982; Tauras, 2006; Wasserman et al., 1991; Yao et al., 2018). Using a two-way fixed effect, DeCicca and McLeod (2008) estimated price elasticity of smoking prevalence to be about -0.21 in USA. Prevalence estimates in Australia vary between -0.32 and -0.86 but differ by income; those in low-income categories are typically more sensitive to price changes (Cameron & Williams, 2001; Siahpush et al., 2009; Zhao & Harris, 2004). The price elasticity of smoking prevalence ranges between -0.45 and -0.30 in Canada (Gospodinov & Irvine, 2009; Gruber et al., 2003), between -0.36 and -0.16 in Spain (Jiménez-Martín et al., 1998; Labeaga, 1999), -0.19 in the UK (Jones, 1989) and -0.02 in Korea (Chung et al., 2008).

Using a two-part model, Kostova et al. (2011) estimated the price elasticity of smoking participation to be -0.74 in 17 LMICs. Price elasticity of cigarette demand in LMICs in Asia and Pacific regions range from -0.04 to -1.30 (Ho et al., 2018). In China, price elasticity of prevalence range between -0.21 to -0.05 but low-income individuals are more sensitive to price changes than those in middle and high income groups (Bishop et al., 2007; Lance et al., 2004; Mao et al.). Price elasticity of prevalence was estimated to be -0.46 in Nepal (Karki et al., 2003), -1.28 in Myanmar (Kyaing, 2003), and -1.41 in Vietnam (Van Kinh et al., 2006).

Sayginsoy et al. (2002) estimated the price elasticity of smoking participation in Bulgaria to be -0.8 overall, -1.3 among low income, -1.02 among middle income and -0.52 among high income earners. Estimates are much lower in Ukraine at -0.46 (Krasovsky et al., 2002), in Turkey at -0.03 (Onder, 2002) and in Russia at -0.16 (Lance et al., 2004). In Latin America, price elasticity of smoking prevalence range between -0.17 and -0.06 in Mexico (Jiménez-Martín et al., 1998; Miera-Juárez & Iglesias, 2010) and -0.66 in Colombia (Gallego et al., 2020).

2 Methods

2.1 Data

Smoking behaviour data is sourced from two nationally representative surveys: The All Media and Products Survey (AMPS) and the National Income Dynamics Study (NIDS). AMPS and NIDS data is used to look at smoking prevalence over time by race and gender. NIDS data is used for the econometric analysis. Price data are sourced from government statistics.

AMPS is a repeated cross-sectional survey that contains individual data about the use and purchase of electronic and print media as well as consumer goods and services, including cigarettes. Although data before 2002 exists, we were only able to obtain data for the years 2002 to 2015. AMPS data was collected by the South African Audience Research Foundation (SAARF[5]), which ran annually from 1993 to 2015 (South African Audience Research Foundation, 2012). The smoking participation question is the variable of interest for our analysis.

NIDS is a longitudinal dataset collected over five waves (2008, 2010, 2012, 2015, and 2017). The study was funded by the South African government and was managed by the University of Cape Town (Brophy et al., 2018). The base survey used a stratified two-stage sample design. The first stage consisted of the primary sampling units (PSU), and the second stage sampled dwellings within in each PSU (Leibbrandt et al., 2010; Wittenberg, 2009). The smoking-related question asked to those aged 15 and older is: 'Do you smoke cigarettes?' We also use the questions that cover the demographic characteristics of respondents. We weighted the results using the panel weights to account for attrition between waves.

The average price per pack of 20 cigarettes was obtained from Statistics South Africa (Stats SA) (see description in Vellios and van Walbeek (2016)) and aggregated at province level. The provincial average price data was merged with the NIDS data by province and year of survey. Cigarette prices were adjusted for inflation using consumer price index (CPI) for tobacco with 2016 as the base year. The overall CPI, used to deflate cigarette prices, was constructed based on year-on-year annualized inflation rate. Table 1 shows price dispersion for each province in each wave. Data on excise and sales tax were obtained from published budget documents from the National Treasury. Tax variables are used in the control function regression as instruments for price.

[5]It was known as South African Advertising Research Foundation until 2012.

2.2 Estimation strategy

Fractional probit model (Correlated random effect) with exogeneity assumption

Studies that estimate the effect of changing tobacco prices on smoking prevalence do so using a two-part model. The first part uses either a logit or probit model to estimate the price elasticity of prevalence. The second part uses linear regression to estimate the price elasticity of intensity (also known as the conditional price elasticity of demand).

We use the panel data collected in five waves of the NIDS survey to estimate price elasticity of smoking participation. We use the correlated random effects (CRE) fractional probit model estimation proposed by Papke and Wooldridge (2008) for fractional response variables, in our estimation, given that our panel is characterized by a large N (number of cross-sectional observations >15,556 in each wave) and few time periods (small T=5). The approach explicitly allows time invariant fixed effects to be correlated with some explanatory variables by modelling the unobserved effects, conditional on strictly exogenous covariates, using Mundlak (1978) and Chamberlain (1980) device.

The alternative to CRE models is the fixed effects model which attempts to estimate unobserved heterogeneity for each observation in the sample but is often biased in nonlinear models due to the incidental parameters problem (Hahn & Newey, 2004; Jeffrey M. Wooldridge, 2019). The bias correction models proposed by Hahn and Newey (2004) and Fernández-Val (2009) cannot be applied here due to relatively small T. The dependent variable, y_{it}, is a binary dummy variable, indicating the decision to smoke. The first part of our estimation assumes that the explanatory variables are strictly exogenous such that,

$$E(\gamma_{it} \mid \mathbf{X}_{it}, M_t, c_i) = \Phi(\mathbf{X}_{it}\boldsymbol{\beta}+\mathrm{M}_t\delta + c_i) \quad t = 1, \ldots, T \tag{1}$$

$\mathbf{X}_{it}$ is a 1xK vector of time-varying explanatory variables, M_t captures changes in the cigarette market structure since 2010 (increased illicit trade, cheaper brands, competition) following the entry of local tobacco companies in South Africa (Linegar & van Walbeek, 2018; van der Zee et al., 2019; Corné van Walbeek, 2014; Vellios et al., 2019). $\Phi(\cdot)$ is the standard normal cumulative distribution (probit function) and c_i is the unobserved heterogeneity. Under strict exogeneity, the unobserved heterogeneity is specified as shown in eq. 2.

$$\begin{aligned} c_i &= \psi + \bar{\mathbf{X}}_i\gamma + \alpha_i, \alpha_i \mid \mathbf{X}_{it} \sim N(0, \sigma_\alpha^2) \\ c_i &\mid \mathbf{X}_{i1}, \ldots, \mathbf{X}_{iT} \sim N(\psi + \bar{\mathbf{X}}_i\gamma, \sigma_\alpha^2) \end{aligned} \tag{2}$$

The conditional distribution assumption of $c_i|\mathbf{X}_{it}$ allows for consistent estimation of the average partial effects (APEs). Substituting Eq. 2 into Eq.1, we estimate Eq.3 by maximum likelihood estimation (MLE) using a generalized linear model (GLM) with probit as the link function.

$$E(\gamma_{it} \mid \mathbf{X}_{it}, M_t, c_i) = \phi(\mathbf{X}_{it}\boldsymbol{\beta} + \mathrm{M}_t\delta + \psi + \bar{\mathbf{X}}_i y), \tag{3}$$

Where

$$\bar{\mathbf{X}}_i = \frac{1}{T}\sum_{t=1}^{T}\mathbf{X}_{it}$$

The inclusion of $\bar{\mathbf{X}}_i$ controls for possible correlation between unobserved heterogeneity and the explanatory variables. Pooled Bernoulli Quasi-MLE identifies the APEs without the conditional serial independence assumption (Papke & Wooldridge, 2008), by estimating scaled parameters using a scale factor

$$\alpha = 1/(1 + \sigma_\alpha^2)^{1/2}$$

$$E(\gamma_{it} \mid \mathbf{X}_{it}, M_t, c_i) = \Phi(\mathbf{X}_{it}\boldsymbol{\beta}_\alpha + \mathrm{M}_t\delta + \psi_{i\alpha} + \bar{\mathbf{X}}_i\gamma_\alpha), \tag{4}$$

Taking the derivative of Eq.4 with respect to price for example, $X_{it(price)}$, gives the average partial effect (APE) of price change to smoking prevalence in eq. 5.

$$\begin{aligned} APE = \left[\frac{\partial\Phi(\mathbf{X}_{it}\boldsymbol{\beta}_\alpha + \psi_{i\alpha} + \bar{\mathbf{X}}_i\gamma_\alpha)}{\partial x_{it(price)}}\right] = \boldsymbol{\beta}_{i\alpha(price)} E[\Phi[\mathbf{X}_{it}\boldsymbol{\beta}_\alpha \\ + \mathrm{M}_t\delta + \psi_{i\alpha} + \bar{\mathbf{X}}_i\gamma_\alpha \end{aligned} \tag{5}$$

Given that no individual smoker purchases enough cigarettes to influence the market price, the assumption of exogeneity of price variable is plausible (International Agency for Research on Cancer, 2011).

We also estimate the simple panel linear probability model (eq 6) and random effects probit model (eq.7) for robustness check.

$$\gamma_{it} = \mathbf{X}_{it}\boldsymbol{\beta} + \mathrm{M}_t\delta + c_i + \boldsymbol{\mu}_{iy} \tag{6}$$

$$\left\{\begin{array}{c}\gamma_{it}^{*}=\mathbf{X}_{it}\boldsymbol{\beta}+\mathrm{M}_{t}\delta+c_{i}+\boldsymbol{\mu}_{it}\\ \gamma_{it}=1[\gamma_{it}^{*}>0]\end{array}\right\} \tag{7}$$

y_{it}^{*} is the latent variable only observed when the $y_{it}=1$ and $\boldsymbol{\mu}_{it}$ is the identically independent error with mean zero. Other variables are defined like before. The unobserved individual time invariant effects (c_i) is assumed to be uncorrelated to $\mathbf{X}_{it}$ and $\mathbf{M}_t$.

Test for endogeneity of price variable using the Control Function Approach

It is possible that there are omitted time-varying variables that could be related to the price of cigarettes and the decision to smoke. If that is the case, our estimates would be inconsistent and inefficient. We test for endogeneity of cigarette prices by running a control function regression using tax to instrument prices. The effect of taxes on tobacco prices is well documented in the literature and consumers' decision to smoke is not influenced by taxes but the price they pay for cigarettes. Taxes therefore affect the decision to smoke indirectly through cigarette prices. The tax instrument is made up of excise tax and sales tax. Liu (2010) used excise taxes as an instrument for cigarette prices to look at the effect of prices on smoking initiation, cessation, and relapse. Ngeh Tingum et al. (2019) used taxes to instrument cigarette prices to estimate the price elasticity of cigarette demand.

$$\begin{array}{c}E(\gamma_{it}\mid \Pr ice_{it},\mathbf{Z}_{it},c_{i1,}u_{it1})=\Phi(\theta_{1}\Pr ice_{it}\\ +\mathbf{Z}_{it}\beta_{1}+M\delta_{1}+c_{i1,}+\upsilon_{it1}),\end{array} \tag{8}$$

where υ_{it1} is the time-varying omitted variable; $\mathbf{Z}_{it}=(\mathbf{Z}_{it1}+\mathbf{Z}_{it2})$ is a vector of exogenous variables and M is the dummy capturing changes in the cigarette market structure as before.

$$\begin{array}{c}c_{i1}=\psi_{1}+\bar{\mathbf{Z}}_{i}y_{1}+\alpha_{i1},\quad \alpha_{i}\mid \mathbf{Z}_{i}\sim N(0,\sigma_{\alpha}^{2})\\ \text{and } D(\alpha_{i}\mid \mathbf{Z}_{i})=D(\alpha_{i})\end{array}$$

Substituting Eq.9 into Eq.8

$$\begin{array}{c}E(\gamma_{it}\mid \Pr ice_{it},\mathbf{Z}_{it},c_{i},\upsilon_{it1})=\Phi(\theta_{it}\Pr ice_{it}\\ +\mathbf{Z}_{it}\beta+M\delta+\psi_{1}+\bar{\mathbf{Z}}_{i}y_{1}+\alpha_{i1}+\upsilon_{it1}\end{array} \tag{9}$$

Let $\xi_{it}=\alpha_{i1}+\upsilon_{it1}$

We estimate the model in two steps as follows: First, we run a reduced form regression of endogenous variable (price) against other exogenous variables and the instruments (exercise tax and sales tax) and obtain the residuals υ_{it2}.

$$\Pr ice_{it} = \psi + \mathbf{Z}_{it}\delta + \bar{\mathbf{Z}}_i y + \upsilon_{it2} \tag{10}$$

The correlation between the errors in the estimation equation (10) and the reduced form eq.11 determines the nature of endogeneity of the price variable in our case.

$$\xi_{it} = \eta\upsilon_{it2} + e_{it} \tag{11}$$

Eq.13 presents error distribution assumptions.

$$\begin{aligned} &\xi_i \mid (\mathbf{Z}_i, \upsilon_{it2}) \sim N(\eta\upsilon_{it2}, w_i^2) \\ &\text{and} \quad e_{it} \mid (\mathbf{Z}_i, \upsilon_{it2}) \sim N(0, \sigma_e^2) \end{aligned} \tag{12}$$

The F-statistic [F (2,37724) = 15734.25] obtained using the testparm routine in Stata after regressing log of real price against other independent variables and instruments indicates that the instruments are relevant and strong.In the second step, we use fractional probit quasi-maximum likelihood estimator to estimate the scaled parameters of Eq.14. The scaling factor is $\omega = 1/(1 + \sigma_{ie}^2)^{1/2}$

$$\begin{aligned} E(y_{it}|Price_{it}, Z_i, \upsilon_{it2}) = \Phi(&\hat{\theta}_\omega Price_{it} + Z_{it}\beta_\omega \\ &+M\delta_\omega + \psi_\omega + \bar{Z}_i\gamma_\omega + \eta_\omega\upsilon_{it2}), \end{aligned} \tag{13}$$

The estimated scaled parameters give the direction of the effect, but the model also indexes the APEs

$$\begin{aligned} APE_{price} = \hat{\theta}_\omega * \lceil \frac{1}{N}\sum_{i=1}^{N} \Phi(&\hat{\theta}_\omega Price_{it} + \mathbf{Z}_{it}\hat{\beta}_\omega \\ &+M\hat{\delta}_\omega + \hat{\psi}_\omega + \bar{\mathbf{Z}}_i\hat{\gamma}_\omega + \hat{\eta}_\omega\upsilon_{it2})\rceil \end{aligned} \tag{14}$$

Since the fitted residuals from the first stage is included in the second stage regression, we corrected the standard errors by bootstrapping, as recommended by Jeffrey. M. Wooldridge (2010). We used the Chow test to check for a structural break in price elasticity parameters following the 2010 changes in cigarette market structure in South Africa.

3 Descriptive analysis

3.1 Overview of smoking prevalence trends by gender

Descriptive statistics presented in Table 2 are calculated using the NIDS data. Smoking prevalence decreased between 2008 - 2010, followed by an increase in the period 2010 - 2015 (Table 2). Overall prevalence decreased by about one percentage point from 20.4% in 2015 to 19.3% in 2017, but it is still higher than the 17.8% prevalence reported in 2010. The prevalence for males was 33.6% in 2017 compared to 30.3% in 2010. Female smoking prevalence was 9.0% in 2010 and remained below 8.0% for the period 2012 - 2017. The male/female prevalence ratio has increased from 4.0 in 2008 to 4.9 in 2017.

Smoking prevalence is lower for each additional level of education attained. University graduates have the lowest prevalence in each wave. Smoking prevalence is highest among those with primary qualification or less. Smoking prevalence among those with university degree and post matric qualification have been declining across waves. Prevalence for university graduates declined from 18.5% in 2008 to 9.9% in 2017.

Smoking prevalence initially increases with age but starts to decline for those in the age category 50-59 years and beyond. Prevalence among the youth in the age category 15-19 years increased from 5.5% in 2010 to 8.1% in 2017. There was a substantial jump in the average inflation-adjusted prices for a pack of 20 cigarettes between 2008 and 2010 but the prices remained nearly flat thereafter. The size of the illicit cigarette market has increase substantially post-2010, from 14.1% in 2010 to 34.5% in 2017 (Vellios et al., 2019).

Evolution of cigarette prices, illicit trade, and smoking prevalence (1980-2017)

Figure 1 shows a slight decline in real cigarette prices before 1992 followed by a steady increase from the early 1990s until 2010 after which prices remained largely unchanged. As inflation-adjusted prices started to increase, smoking prevalence, estimated at 33% in 1993, started to decline until 2010. The entry of local tobacco companies and introduction of cheaper cigarette brands in the market depressed the real prices (orange bars, Figure 1). There is a slight increase in prevalence in the post-2010 period. Since our data only start in 2002, smoking prevalence estimates prior to 2002 were obtained from previous published work using earlier AMPS cross section surveys (Corne van Walbeek, 2002).

The estimates of the share of illicit trade in cigarette market were obtained from Vellios et al. (2019). The authors used gap analysis in their estimation,

by comparing the self-reported consumption from survey data and the official tax-paid cigarette sales. The data shows a sharp increase in the share of illicit trade in cigarette market from 2010 (black line, Figure 1)

4 Multiple regression results using NIDS data.

Overview

We estimate price elasticity of smoking prevalence by simple panel linear probability model (panel LPM), the Mundlak-Chamberlain approach using correlated random effects pooled fractional probit (CRE-PFP) and the random effects probit model (RE Probit). Table 3 presents the results of the three models while controlling for the effect of illicit trade (estimated share of the illicit cigarettes) on the responsiveness of smoking prevalence to price changes. Table 4 presents the same estimations while controlling for the changes in cigarette market structure (dummy variable, 0 for 2009 and earlier, 1 for 2010 and later) in South Africa since 2010. In all the three model estimations, we control for province-specific effects to account for regional differences in cigarette pricing and smoking prevalence.

The models are estimated for males and females combined, and for males and females separately. We also estimated eq.14 (a control function) to check for endogeneity of price variable. Since the coefficients on the residuals v_{it2} in equation [14] are not significant (exogeneity of the price variable is maintained), we only discuss the results presented in Table 3 and Table 4 and present the control function estimation in the appendix (Table 5).

Price elasticity of smoking participation and the effect of illicit trade and structural change in cigarette market

All models show a significant reduction in smoking prevalence when cigarette prices are increased. For example, in Panel A of Table 3, the coefficient on log(price) is estimated to be between -0.27 (PFP) and -0.29 (panel LPM). This implies that, for every 10% increase in the price of cigarettes, smoking prevalence decreases by between 2.7% and 2.9% overall. Smoking prevalence among males is more sensitive to price changes (elasticity of between -0.42 and -0.47, Panel A Table 3) than females (elasticity of -0.15 and -0.19 Panel C Table 3). An interaction variable (price multiplied by the share of the illicit market) was included to determine whether the illicit market modified the impact of price changes on smoking prevalence. The coefficient on the interaction term is in the range 0.013- 0.016 overall (between 0.02 and 0.025 in males and 0.010 in females), which means that for every 10 percentage points increase in the illicit market share, the smoking prevalence elasticity

becomes less negative (i.e., less elastic) by 0.16 (0.25 for males and 0.10 for females).

When we control for structural change in the cigarette market since 2010, (Table 4), the estimated price elasticity of smoking prevalence is between -0.27 (PFP) and -0.34 (panel LPM) overall, -0.43 and -0.60 for males and -0.14 &-0.20 for females.

The effect of the structural change in the cigarette market on the impact of price changes on smoking prevalence is captured by the interaction variable (price multiplied by the post2010 dummy). The coefficient on the interaction term is 0.014 overall, 0.02 for males and 0.01 for females. The estimates are very close to those obtained when we interact price with share of illicit market. The similarity in magnitude of the coefficients on illicit trade and structural break indicators suggested that the main effect of the structural change in the cigarette market is through the size of illicit trade. The net effect is a reduction in the responsiveness of smoking prevalence to changes in real cigarette prices.

Other determinants of smoking prevalence

Smoking prevalence is higher among males than females by between 24 and 25 percentage points. There is an inverse relationship between smoking prevalence and education. The panel LPM consistently underestimates the effect of education on prevalence when compared to the CRE fractional Probit estimates. Smoking prevalence among individuals with some high school education is 7.5 percentage points lower than the reference category (adults with completed primary school education or less). Those with matric qualification and university graduates have 13.9 and 19.3 percentage points lower, respectively, than the reference category.

Smoking prevalence varies significantly by age and race. It is significantly higher among adults than teenagers, aged 15-19. The prevalence appears to increase with age up to 40-49 years and then start to decline both for males and females. Overall, smoking prevalence among Coloureds, Indians and Whites is substantially higher than among Africans. Compared to the refence group (Africans), smoking prevalence in males is higher by 24.1% among Coloureds, 9% among Indians and 22.4% among Whites. Smoking prevalence in females is higher by 23% among Coloureds, 5.3% among Indians and 27.9 among Whites than among Africans.

Smoking prevalence differ significantly by province. Except for the Northern and Free state provinces, smoking prevalence in Western Cape is higher than in the other six provinces. However, smoking prevalence of males in North West Province is not significantly different from that of males in Western Cape (Table 3 and Table 4).

5 Discussion

This paper provides new insights into the price elasticity of smoking prevalence by gender in South Africa. The paper tests how increased share of illicit cigarettes and changes in the cigarette market structure in 2010 has affected smoking prevalence in South Africa. The changes in market structure has resulted in many more and cheaper brands coming to the market affecting the price paid by cigarette consumers. Trading in illicit cigarettes has also increased substantially during this period (van der Zee et al., 2020; Vellios et al., 2019). There has not been any non-price tobacco control legislation since the 2008 amendment of the Tobacco Products Control Act of 1993 that could have changed the smoking prevalence. Given that the taxes are embedded in cigarette prices and our regression analysis covers the period 2008 - 2017, the inclusion of the estimated size of illicit trade in cigarette market and post-2010 structural change indicator in the tobacco industry captures the heterogeneity in the tobacco market. Increasing share of Illicit trade makes tobacco products more accessible and affordable while also leading to a substantial loss of tax revenue. We capture the effect of these market dynamics on the price elasticity of smoking prevalence by interacting the price variable with both the share of illicit trade and the dummy variable for structural change in tobacco market.

The paper provides further evidence of the effectiveness of the price-based interventions mediated through the excise tax as an effective policy tool in tobacco control. The overall price elasticity of smoking prevalence is estimated to be between -0.27 and -0.34 after controlling for the changes in the cigarette market structure and illicit cigarettes but males are more responsive to changes in real cigarette prices than females. The results also demonstrate that the change in the cigarette market since 2010 and the increasing share of illicit trade in the post-2010 period is undermining the current price-based tobacco control policy. According to the 2012 WHO CFTC Protocol to Eliminate Illicit Trade in Tobacco Products, illicit trade poses a serious threat to public health. The coefficients on the interaction between the illicit trade proportion and price and that of price interacted with the structural change dummy indicator, are positive, which weakens the effect of price increases on smoking prevalence. The results are robust to different model specifications.

6 Conclusion

In conclusion, increasing excise taxes and therefore cigarette prices remain an effective tool to reduce tobacco demand. However, the post-2010 price

wars, availability of cheap cigarettes and increase in illicit tobacco in South Africa market is undermining the gains made in tobacco control since 1990s. In order to reverse the trend of increasing smoking prevalence, the government should increase cigarette excise taxes, which in turn should increase prices. However, an important proviso is that the government should prioritize the fight against illicit trade. A good start would be to ratify the 2012 WHO CFTC Protocol to Eliminate Illicit Trade in Tobacco Products and implement the measures that the country would commit itself to.

7 Study limitations

The entry of new tobacco companies since 2010 and increasing share of illicit trade in cigarettes means that the prices paid by tobacco consumers are more dynamic. The actual cigarette prices paid by smokers were not available for the analysis. The study uses official cigarette prices from STAT South Africa, which do not capture the prices paid for the illicit cigarettes. We believe that including the estimated share of illicit cigarettes and a dummy indicator for structural change in cigarette market in post-2010 period, addresses those shortcomings in the price data by capturing the dynamics in the tobacco market. An area for further research would include estimating the model using the actual prices paid by consumers.

Ethics

The study used secondary data and therefore seeking ethical approval from the University was not necessary.

References

[1] Bishop, J. A., Haiyong, L., & Qi, M. (2007). Are Chinese smokers sensitive to price? *China Economic Review, 18*(2), 113-121.

[2] Boachie, M. K., & Ross, H. (2020). Determinants of smoking intensity in South Africa: evidence from township communities. *Preventive Medicine Reports*, 101099.

[3] Brophy, T., Branson, N., Daniels, R. C., Leibbrandt, M., Mlatsheni, C., & Woolard, I. (2018). National Income Dynamics Study Panel User Manual. In (Vol. Version 1. Cape Town): Release 2018.

[4] Cameron, L., & Williams, J. (2001). Cannabis, alcohol and cigarettes: substitutes or complements? *Economic Record, 77*(236), 19-34.

[5] Chaloupka, F. J., Straif, K., & Leon, M. E. (2011). Effectiveness of tax and price policies in tobacco control. *Tobacco control, 20* (3), 235-238.

[6] Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. *Review of Economic Studies, 47*(1), 225-238.

[7] Chelwa, G., & van Walbeek, C. (2019). Does cigarette demand respond to price increases in Uganda? Price elasticity estimates using the Uganda National Panel Survey and Deaton's method. *BMJ open, 9*(3), e026150.

[8] Chung, W.-J., Lee, S.-M., Shin, K.-Y., Lim, S.-J., & Cho, K.-S. (2008). Analysis of willingness-to-quit cigarette price among Korean male adults. *Journal of Preventive Medicine and Public Health, 41* (3), 136-146.

[9] DeCicca, P., & McLeod, L. (2008). Cigarette taxes and older adult smoking: Evidence from recent large tax increases. *Journal of health economics, 27*(4), 918-929.

[10] Du Toit, P. (2018). Sars is a mess, thanks to Tom Moyane. Here's how he did it. Retrieved from https://www.huffingtonpost.co.uk/2018/03/20/the-damage-tom-moyane-wrought-on-sars_a_23390102/

[11] Euromonitor International. (2020). Cigarettes in South Africa. Retrieved 25Nov 2020 https://www.portal.euromonitor.com/portal/analysis/tab

[12] Farrelly, M. C., Bray, J. W., Pechacek, T., & Woollery, T. (2001). Response by adults to increases in cigarette prices by sociodemographic characteristics. *Southern Economic Journal*, 156-165.

[13] Fernández-Val, I. (2009). Fixed effects estimation of structural parameters and marginal effects in panel probit models. *Journal of Econometrics, 150*(1), 71-85. doi:https://doi.org/10.1016/j.jeconom.2009.02.007

[14] Gallego, J. M., Otálvaro-Ramírez, S., & Rodriguez-Lesmes, P. A. (2020). Price smoking participation elasticity in Colombia: estimates by age and socioeconomic level. *Tobacco control*.

[15] Gligorić, D., Pepić, A., Petković, S., Ateljević, J., & Vukojević, B. (2020). Price elasticity of demand for cigarettes in Bosnia and Herzegovina: microdata analysis. *Tobacco control*, tobaccocontrol-2019-055258. doi:10.1136/tobaccocontrol-2019-055258

[16] Gospodinov, N., & Irvine, I. (2009). Tobacco taxes and regressivity. *Journal of health economics, 28*(2), 375-384.

[17] Gruber, J., Sen, A., & Stabile, M. (2003). Estimating price elasticities when there is smuggling: the sensitivity of smoking to price in Canada. *Journal of health economics, 22*(5), 821-842.

[18] Hahn, J., & Newey, W. (2004). Jackknife and analytical bias reduction for nonlinear panel models. *Econometrica, 72*(4), 1295-1319.

[19] Ho, L.-M., Schafferer, C., Lee, J.-M., Yeh, C.-Y., & Hsieh, C.-J. (2018). Raising cigarette excise tax to reduce consumption in low-and middle-income countries of the Asia-Pacific region: a simulation of the anticipated health and taxation revenues impacts. *BMC public health, 18*(1), 1187.

[20] International Agency for Research on Cancer. (2011). *Effectiveness of tax and price policies for tobacco control* (Vol. 14): World Health Organization.

[21] James, E. K., Saxena, A., Franco Restrepo, C., Llorente, B., Vecino-Ortiz, A. I., Villar Uribe, M., . . . Verguet, S. (2019). Distributional health and financial benefits of increased tobacco taxes in Colombia: results from a modelling study. *Tobacco control, 28*(4), 374-380. doi:10.1136/tobaccocontrol-2018-054378

[22] Jha, P., & Peto, R. (2014). Global effects of smoking, of quitting, and of taxing tobacco. *New England Journal of Medicine, 370*(1), 60-68.

[23] Jiménez-Martín, S., Labeaga, J. M., & López, A. (1998). Participation, heterogeneity and dynamics in tobacco consumption: evidence from cohort data. *Health Economics, 7*(5), 401-414.

[24] Jones, A. (1989). The UK demand for cigarettes 1954-1986, a double-hurdle approach. *Journal of health economics, 8*(1), 133-141.

[25] Karki, Y. B., Pant, K. D., & Pande, B. R. (2003). A Study on the Economics of Tobacco in Nepal.

[26] Kostova, D., Ross, H., Blecher, E., & Markowitz, S. (2011). Is youth smoking responsive to cigarette prices? Evidence from low- and middle-income countries. *Tobacco control, 20*(6), 419-424. doi:10.1136/tc.2010.038786

[27] Krasovsky, K., Andreeva, T., Krisanov, D., Mashliakivsky, M., & Rud, G. (2002). Economics of tobacco control in Ukraine from the public health perspective. *Kiev: Polygraph Center TAT.*

[28] Kyaing, N. N. (2003). Tobacco economics in Myanmar. *Health, Nutrition and Population (HNP) Discussion Paper, Economics of Tobacco Control No.14. The World Bank and WHO.*

[29] Labeaga, J. M. (1999). A double-hurdle rational addiction model with heterogeneity: estimating the demand for tobacco. *Journal of Econometrics, 93*(1), 49-72.

[30] Lance, P. M., Akin, J. S., Dow, W. H., & Loh, C.-P. (2004). Is cigarette smoking in poorer nations highly sensitive to price?: Evidence from Russia and China. *Journal of health economics, 23*(1), 173-189.

[31] Leibbrandt, M., Woolard, I., & De Villiers, L. (2010). The South African national income dynamics study: design and methodological issues. *Studies in Economics and Econometrics, 34*(3), 7-24.

[32] Lewit, E. M., & Coate, D. (1982). The potential for using excise taxes to reduce smoking. *Journal of health economics, 1*(2), 121-145.

[33] Linegar, D. J., & van Walbeek, C. (2018). The effect of excise tax increases on cigarette prices in South Africa. *Tobacco control, 27*(1), 65-71. doi:10.1136/tobaccocontrol-2016-053340

[34] Liu, F. (2010). Cutting through the smoke: separating the effect of price on smoking initiation, relapse and cessation. *Applied Economics, 42*(23), 2921-2939.

[35] Maldonado, N., Llorente, B., & Deaza, J. (2016). Cigarette taxes and demand in Colombia/Impuestos y demanda de cigarrillos en Colombia. *Revista Panamericana de Salud Publica, 40*(4), 229-237.

[36] Mao, Z., Sung, H.-Y., Hu, T.-w., & Yang, G. The Demand for Cigarettes in China. In *Tobacco Control Policy Analysis in China* (pp. 129-157).

[37] Miera-Juárez, B. S. d., & Iglesias, R. (2010). Impuestos para el control del tabaquismo: las experiencias de Brasil y México. *salud pública de méxico, 52*, S172-S185.

[38] Mukong, A. K., & Tingum, E. N. (2020). The demand for cigarettes: New evidence from South Africa. *Development Southern Africa, 37* (1), 40-54.

[39] Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica: journal of the econometric society*, 69-85.

[40] Nayab, D., Nasir, M., Memon, J. A., Khalid, M., & Hussain, A. (2020). Estimating the price elasticity for cigarette and chewed tobacco in Pakistan: evidence from microlevel data. *Tobacco control*, tobaccocontrol-2019-055232. doi:10.1136/tobaccocontrol-2019-055232

[41] Ngeh Tingum, E., Mukong, A. K., & Mdege, N. D. (2019). The effects of price and non-price policies on cigarette consumption in South Africa.

[42] Oelsner, E. C., Balte, P. P., Bhatt, S. P., Cassano, P. A., Couper, D., Folsom, A. R., . . . Mathew, A. R. (2020). Lung function decline in former smokers and low-intensity current smokers: a secondary data analysis of the NHLBI Pooled Cohorts Study. *The Lancet Respiratory Medicine, 8*(1), 34-44.

[43] Onder, Z. (2002). The economics of tobacco in Turkey: new evidence and demand estimates.

[44] Papke, L. E., & Wooldridge, J. M. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics, 145*(1-2), 121-133.

[45] Paraje, G., Araya, D., De Paz, A., & Nargis, N. (2020). Price and expenditure elasticity of cigarette demand in El Salvador: a household-level analysis and simulation of a tax increase. *Tobacco control*, tobaccocontrol-2019-055568. doi:10.1136/tobaccocontrol-2019-055568

[46] Ross, H., Tesche, J., & Vellios, N. (2016). *Undermining Government Tax Policies: Common strategies employed by the tobacco industry in response to tobacco tax increases.*

[47] Sayginsoy, O., Yurekli, A. A., & De Beyer, J. (2002). Cigarette demand, taxation, and the poor: a case study of Bulgaria.

[48] Siahpush, M., Wakefield, M. A., Spittal, M. J., Durkin, S. J., & Scollo, M. M. (2009). Taxation reduces social disparities in adult smoking prevalence. *American Journal of Preventive Medicine, 36* (4), 285-291.

[49] South African Audience Research Foundation. (2012). The Evolution Of AMPS In: Retrieved from SAARF: http://www.saarf.co.za/amps/amps-evolution.asp.

[50] Tauras, J. A. (2006). Smoke-free air laws, cigarette prices, and adult cigarette demand. *Economic Inquiry, 44*(2), 333-342.

[51] Tingum, E. N., Mukong, A. K., & Mdege, N. (2020). The effects of price and non-price policies on cigarette consumption in South Africa. *Tobacco Induced Diseases, 18*.

[52] Treasury, N., & Manuel, T. A. (2004). *Budget speech, 18 February 2004* (Vol. 19/2004). Cape Town: National Treasury.

[53] van der Zee, K., van Walbeek, C., & Magadla, S. (2019). Illicit/cheap cigarettes in South Africa. *Trends in Organized Crime*, 1-21.

[54] van der Zee, K., Vellios, N., van Walbeek, C., & Ross, H. (2020). The illicit cigarette market in six South African townships. *Tobacco control*.

[55] Van Kinh, H., Ross, H., Levy, D., Minh, N. T., & Ngoc, V. T. B. (2006). The effect of imposing a higher, uniform tobacco tax in Vietnam. *Health Research Policy and Systems, 4*(1), 6.

[56] van Walbeek, C. (2002). Recent trends in smoking prevalence in South Africa-some evidence from AMPS data. *South African Medical Journal, 92*(6), 468-472.

[57] van Walbeek, C. (2014). Measuring changes in the illicit cigarette market using government revenue data: the example of South Africa. *Tobacco control, 23*(e1), e69-e74.

[58] Van Walbeek, C., Filby, S., & Van der Zee, K. (2020). *Lighting up the illicit market: Smokers Responses to the cigarette sales ban in South Africa*. Retrieved from Research Unit on Excisable Products (REEP):

[59] Vellios, N., & van Walbeek, C. (2016). Determinants of regular smoking onset in South Africa using duration analysis. *BMJ open, 6* (7), e011076.

[60] Vellios, N., van Walbeek, C., & Ross, H. (2019). Illicit cigarette trade in South Africa: 2002-2017. *Tobacco control*, tobaccocontrol-2018-054798. doi:10.1136/tobaccocontrol-2018-054798

[61] Vladisavljevic, M., Zubović, J., Đukić, M., & Jovanović, O. (2020). Tobacco price elasticity in Serbia: evidence from a middle-income country with high prevalence and low tobacco prices. *Tobacco control*, tobaccocontrol-2019-055262. doi:10.1136/tobaccocontrol-2019-055262

[62] Wasserman, J., Manning, W. G., Newhouse, J. P., & Winkler, J. D. (1991). The effects of excise taxes and regulations on cigarette smoking. *Journal of health economics, 10*(1), 43-64.

[63] Wittenberg, M. (2009). Weights: report on NIDS Wave 1. *NIDS Technical Paper, 2*.

[64] Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*: MIT press.

[65] Wooldridge, J. M. (2019). Correlated random effects models with unbalanced panels. *Journal of Econometrics, 211*(1), 137-150. doi:https://doi.org/10.1016/j.jeconom.2018.12.010

[66] World Health Organization. (2003). *WHO framework convention on tobacco control*: World Health Organization.

[67] World Health Organization. (2010). *WHO technical manual on tobacco tax administration*: World Health Organization.

[68] World Health Organization. (2014). *WHO Framework Convention on Tobacco Control. Guidelines for implementation of Article 6: Price and tax measures to reduce the demand for tobacco*. Retrieved from https://apps.who.int/iris/handle/10665/145110

[69] World Health Organization. (2018). *WHO global report on trends in prevalence of tobacco smoking 2000-2025*: World Health Organization.

[70] Yao, T., Ong, M. K., Max, W., Keeler, C., Wang, Y., Yerger, V. B., & Sung, H.-Y. (2018). Responsiveness to cigarette prices by different racial/ethnic groups of US adults. *Tobacco control, 27*(3), 301-309. doi:10.1136/tobaccocontrol-2016-053434

[71] Zhao, X., & Harris, M. N. (2004). Demand for Marijuana, Alcohol and Tobacco: Participation, Levels of Consumption and Cross-equation Correlations. *Economic Record, 80*(251), 394-410.

Table 1: Average real price of a pack of 20 cigarettes by year and province

			2008					2010		
Province	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Western Cape	572	30.37	5.45	19.94	50.94	4 359	35.12	3.61	17.58	44.29
Eastern Cape	609	31.73	5.59	24.66	47.54	2 792	34.18	3.83	17.98	42.89
Northern Cape	454	30.90	5.08	21.01	47.31	677	33.31	3.80	17.23	39.57
Free State	557	30.32	5.11	18.63	49.47	2 075	34.06	4.21	16.05	47.73
KwaZulu-Nata	570	31.98	6.07	17.01	51.40	2 187	35.56	4.16	17.99	45.66
North West	477	31.03	4.54	20.08	48.68	941	35.10	3.65	18.68	41.50
Gauteng	742	32.35	5.55	25.96	49.86	6 028	35.90	3.54	17.71	46.98
Mpumalanga	331	30.51	5.01	16.85	48.68	725	35.39	3.66	18.68	44.96
Limpopo	416	31.04	5.20	21.01	46.67	272	35.38	3.55	20.07	42.89
			2012					2015		
Western Cape	1 894	35.21	3.43	18.35	44.92	329	35.80	5.59	12.77	46.81
Eastern Cape	511	34.14	3.34	12.48	40.68	147	34.70	3.53	18.62	40.43
Northern Cape	195	32.81	3.42	19.97	36.82	126	35.08	4.53	21.28	44.74
Free State	927	33.89	4.60	19.08	43.62	218	35.29	4.60	21.28	42.56
KwaZulu-Nata	1 036	35.27	4.58	19.22	43.06	250	36.65	4.55	19.15	42.55
North West	606	35.29	2.47	24.95	41.81	155	36.60	2.14	31.92	40.43
Gauteng	3 261	35.60	3.26	19.96	43.68	564	37.01	3.74	23.41	53.20
Mpumalanga	359	34.00	4.69	18.71	40.55	145	34.13	5.45	19.14	41.92
Limpopo	193	35.29	2.61	19.35	39.93	90	36.77	3.21	33.51	45.74
			2017							
Western Cape	692	34.67	5.34	22.32	53.47					
Eastern Cape	263	34.64	3.87	18.99	52.23					
Northern Cape	302	33.79	5.65	19.94	52.23					
Free State	417	34.72	4.51	22.32	52.23					
KwaZulu-Nata	485	34.99	4.34	19.47	53.18					
North West	275	36.48	2.18	31.34	52.23					
Gauteng	980	35.94	3.67	21.46	52.23					
Mpumalanga	299	33.80	5.38	19.46	53.18					
Limpopo	153	32.96	5.20	12.35	37.98					

Source: Authors calculations using data from STAT SA

Table 2: Weighted sample proportions and smoking prevalence

	Sample proportion					Smoking prevalence (%) and 95% confidence interval in parenthesis				
	wave 1	wave 2	wave 3	wave 4	wave 5	wave 1 [2008]	wave 2 [2010]	wave 3 [2012]	wave 4 [2015]	wave 5 [2017]
Overall (%)						20.8 (20.2-21.5)	17.8 (17.2-18.4)	19.5 (19-20.1)	20.4 (19.9-20.9)	19.3 (18.8-19.8)
Female (%)	56.1	54.0	54.1	53.2	53.3	9.0 (8.4-9.6)	7.2 (6.7-7.7)	7.7 (7.2-8.2)	7.9 (7.4-8.4)	6.9 (6.5-7.3)
Male (%)	43.9	46.0	45.9	46.8	46.7	35.9 (34.7-37.1)	30.3 (29.2-31.3)	33.5 (32.4-34.5)	34.6 (33.7-35.6)	33.6 (32.6-34.5)
Prevalence ratio (M/F)						4.0	4.2	4.3	4.4	4.9
Education										
<=primary (%)	29	26.1	23.9	20.6	17.3	24.4 (23.3-25.5)	19.7 (18.6-20.7)	24.3 (23.2-25.4)	25.8 (24.7-26.8)	21.0 (19.9-22.1)
Some secondary (%)	42	44.4	45.2	46.8	47.0	19.9 (18.9-20.9)	17.7 (16.8-18.6)	19.2 (18.4-20.1)	21.3 (20.5-22.1)	21.9 (21.2-22.7)
Matric (%)	17	16.6	16.7	16.6	18.6	18.5 (16.8-20.1)	16.1 (14.6-17.6)	18.5 (17.0-20.0)	17.6 (16.3-18.9)	16.1 (14.9-17.3)
University degree (%)	4	3.6	3.7	3.6	4.4	18.5 (14.2-22.7)	16.7 (12.1-21.3)	10.9 (7.5-14.3)	9.1 (6.4-11.9)	9.9 (7.8-12.1)
Other tertiary training (%)	8.6	9.4	10.6	12.5	12.6	18.8 (16.3-21.3)	16.8 (14.5-19.1)	14.4 (12.6-16.3)	15.0 (13.5-16.6)	15.3 (13.9-16.8)
Age group: (%)										
15-19	16	15	14	13	12	7.8 (6.7-8.8)	5.5 (4.7-6.3)	5.2 (4.4-6.0)	6. (5.7-7.2)	8.1 (7.2-9.0)
20-24	13	14	13	13	13	19.5 (17.8-21.2)	15.5 (14.1-16.9)	17.1 (15.7-18.5)	19.8 (18.5-21.2)	16.3 (15.0-17.5)
25-29	12	12	13	14	14	24.9 (22.8-27.1)	20.7 (18.8-22.5)	21.1 (19.3-22.7)	24.1 (22.4-25.5)	22.3 (20.8-23.8)
30-39	21	21	21	21	24	24.0 (22.4-25.6)	23.7 (22.1-25.3)	24.8 (23.3-26.3)	23.1 (21.8-24.4)	23.9 (22.6-25.2)
40-49	16	16	16	16	15	28.6 (26.8-30.4)	25.9 (24.1-27.6)	26.6 (24.9-28.3)	26.8 (25.2-28.3)	22.3 (20.9-23.7)
50-59	11	11	11	11	11	24.5 (22.5-26.4)	17.4 (15.7-19.1)	21.0 (19.3-22.7)	26.3 (24.6-27.9)	23.6 (22.0-25.2)
60-69	7	7	7	7	7	19.9 (17.7-22.0)	15.2 (13.2-17.2)	21.1 (19.0-23.3)	17.2 (15.4-19.0)	17.3 (15.7-19.0)
70+	4	4	4	4	4	11.4 (9.3-13.4)	8.6 (6.8-10.3)	10.4 (8.6--12.2)	9.2 (7.6-10.8)	6.0 (4.8-7.3)
Real price of a pack of 20 cigarettes (Rands)	25.0	29.7	30.9	31.1	31.2					
Adjusted income per capita	5564.7	5888.9	6052.1	6208.2	6181.7					

(USD)										
Illicit cigarette market (%)	0.8	14.1	17.2	20.5	34.5					
Observations	15,556	16,831	18,671	22,709	23,816	15,556	16,831	18,671	22,709	23,816

Table 3: Price elasticity and determinants of smoking prevalence (share of illicit trade)

	Panel A: Overall			Panel B: Male			Panel C: Female		
	CRE-FP	LPM	RE Probit	CRE-FP	LPM	RE Probit	CRE-FP	LPM	RE Probit
Log real price	-0.27***	-0.29***	-0.276***	-0.428***	-0.423***	-0.472***	-0.150***	-0.188***	-0.115***
	(0.057)	(0.027)	(0.027)	(0.113)	(0.057)	(0.063)	(0.048)	(0.025)	(0.019)
Illicit*price	0.016***	0.013***	0.014***	0.025***	0.022***	0.023***	0.008***	0.011***	0.006***
	(0.003)	(0.001)	(0.001)	(0.006)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)
Illicit market (%)	-0.056***	-0.049***	-0.049***	-0.089***	-0.077***	-0.081***	-0.029***	-0.038***	-0.021***
	(0.011)	(0.005)	(0.005)	(0.021)	(0.010)	(0.011)	(0.009)	(0.005)	(0.003)
Male smokers	0.251***	0.242***	0.226***						
	(0.006)	(0.003)	(0.003)						
Education (base: <=primary)									
Some high school	-0.075***	-0.04***	-0.051***	-0.073***	-0.052***	-0.060***	-0.068***	-0.042***	-0.042***
	(0.010)	(0.004)	(0.004)	(0.016)	(0.007)	(0.008)	(0.009)	(0.004)	(0.003)
Matric	-0.139***	-0.091***	-0.094***	-0.189***	-0.143***	-0.148***	-0.102***	-0.065***	-0.058***
	(0.011)	(0.005)	(0.004)	(0.019)	(0.009)	(0.010)	(0.010)	(0.004)	(0.004)
University degree	-0.193***	-0.179***	-0.140***	-0.316***	-0.310***	-0.279***	-0.120***	-0.093***	-0.065***
	(0.012)	(0.009)	(0.005)	(0.033)	(0.019)	(0.013)	(0.013)	(0.009)	(0.004)
Other post matric	-0.158***	-0.109***	-0.102***	-0.242***	-0.181***	-0.176***	-0.101***	-0.066***	-0.058***
	(0.011)	(0.005)	(0.005)	(0.020)	(0.011)	(0.011)	(0.011)	(0.005)	(0.004)
Age group: base 15-19									
20-24	0.102***	0.131***	0.084***	0.208***	0.233***	0.204***	0.060***	0.047***	0.029***
	(0.008)	(0.004)	(0.004)	(0.013)	(0.007)	(0.007)	(0.009)	(0.003)	(0.003)
25-29	0.157***	0.182***	0.135***	0.315***	0.344***	0.319***	0.070***	0.059***	0.038***
	(0.009)	(0.005)	(0.004)	(0.014)	(0.009)	(0.008)	(0.008)	(0.004)	(0.003)
30-39	0.175***	0.193***	0.153***	0.362***	0.380***	0.359***	0.060***	0.063***	0.040***
	(0.008)	(0.006)	(0.004)	(0.015)	(0.013)	(0.008)	(0.008)	(0.006)	(0.003)
40-49	0.188***	0.191***	0.158***	0.375***	0.385***	0.365***	0.072***	0.069***	0.043***
	(0.013)	(0.009)	(0.006)	(0.026)	(0.018)	(0.014)	(0.012)	(0.008)	(0.004)
50-59	0.158***	0.175***	0.154***	0.352***	0.371***	0.353***	0.047***	0.063***	0.043***
	(0.018)	(0.012)	(0.009)	(0.037)	(0.024)	(0.021)	(0.014)	(0.010)	(0.006)
60-69	0.122***	0.141***	0.131***	0.297***	0.321***	0.304***	0.030*	0.047***	0.036***
	(0.023)	(0.015)	(0.011)	(0.049)	(0.030)	(0.028)	(0.017)	(0.013)	(0.007)
70+	0.057**	0.086***	0.082***	0.178***	0.223***	0.194***	0.005	0.023	0.023***
	(0.022)	(0.018)	(0.012)	(0.065)	(0.037)	(0.032)	(0.017)	(0.015)	(0.008)
Race: Base African									
Coloured	0.241***	0.247***	0.244***	0.148***	0.186***	0.200***	0.233***	0.296***	0.213***
	(0.019)	(0.006)	(0.009)	(0.025)	(0.012)	(0.014)	(0.020)	(0.006)	(0.010)
Asian/Indian	0.087***	0.092***	0.105***	0.117**	0.109***	0.128***	0.053***	0.080***	0.092***
	(0.027)	(0.013)	(0.016)	(0.049)	(0.025)	(0.030)	(0.018)	(0.013)	(0.018)
White	0.224***	0.162***	0.187***	0.093***	0.102***	0.114***	0.279***	0.217***	0.222***
	(0.026)	(0.008)	(0.011)	(0.030)	(0.015)	(0.018)	(0.028)	(0.008)	(0.015)
Province dummies: Base Western Cape									

Eastern Cape	-0.038***	-0.029***	-0.022***	-0.069***	-0.027**	-0.027*	-0.010	-0.034***	-0.011***
	(0.014)	(0.007)	(0.006)	(0.025)	(0.013)	(0.014)	(0.011)	(0.007)	(0.004)
Northern Cape	-0.010	-0.007	-0.006	-0.028	-0.016	-0.015	0.004	-0.004	-0.001
	(0.013)	(0.007)	(0.006)	(0.032)	(0.013)	(0.014)	(0.008)	(0.007)	(0.003)
Free State	-0.007	0.001	0.005	-0.008	0.027*	0.027	-0.004	-0.022***	-0.003
	(0.015)	(0.009)	(0.007)	(0.027)	(0.016)	(0.017)	(0.013)	(0.008)	(0.005)
KwaZulu-Natal	-0.052***	-0.054***	-0.053***	-0.070***	-0.053***	-0.062***	-0.034***	-0.053***	-0.040***
	(0.014)	(0.007)	(0.006)	(0.027)	(0.013)	(0.014)	(0.012)	(0.006)	(0.004)
North West	-0.029*	-0.027***	-0.017**	-0.039	-0.011	-0.004	-0.025**	-0.044***	-0.016***
	(0.015)	(0.008)	(0.007)	(0.029)	(0.016)	(0.017)	(0.012)	(0.008)	(0.005)
Gauteng	-0.041***	-0.029***	-0.024***	-0.068***	-0.022*	-0.020	-0.018*	-0.042***	-0.018***
	(0.013)	(0.007)	(0.006)	(0.025)	(0.013)	(0.014)	(0.010)	(0.007)	(0.004)
Mpumalanga	-0.044***	-0.049***	-0.045***	-0.068**	-0.050***	-0.058***	-0.024**	-0.046***	-0.030***
	(0.014)	(0.008)	(0.007)	(0.028)	(0.015)	(0.016)	(0.011)	(0.008)	(0.005)
Limpopo	-0.060***	-0.052***	-0.053***	-0.095***	-0.062***	-0.080***	-0.024**	-0.046***	-0.027***
	(0.015)	(0.008)	(0.007)	(0.029)	(0.015)	(0.017)	(0.012)	(0.007)	(0.005)
urban	0.040***	0.018***	0.026***	0.045***	0.030***	0.036***	0.042***	0.009***	0.017***
	(0.007)	(0.003)	(0.003)	(0.012)	(0.006)	(0.007)	(0.006)	(0.003)	(0.002)
_cons		0.997***			0.000			0.746***	
		(0.096)			(.)			(0.087)	
N	158163304	97141	97141	72560091	39754	39754	85603213	57387	57387

Robust standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.010

Notes for Table 3 and Table 4: CRE-FP is correlated random effects (CRE) pooled fractional probit model estimation (chamberlain-Mundlak approach). The results are compared to the simple panel linear probability model (LPM) and Random Effects Probit Model (RE Probit). CRE-FP model includes time averages of education and age to allow them to be correlated with individual unobserved heterogeneity. After controlling for the size of the illicit cigarette market and province specific effects, the models show a significant effect of the price on smoking prevalence although the elasticity is quite low. The model in Table 3 is estimated controlling for the estimated size of the illicit cigarette market.

Table 4: Price elasticity and determinants of smoking prevalence (structural change in cigarette market)

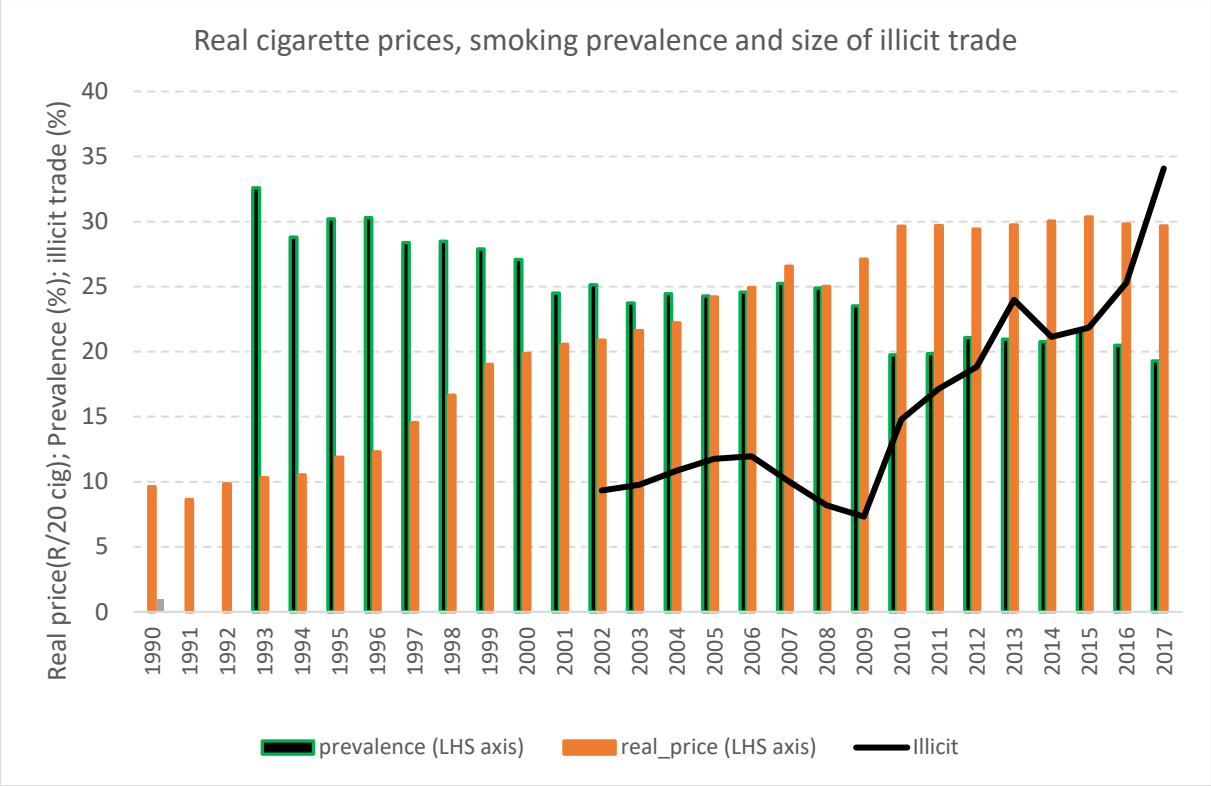
	Panel A: Overall			Panel B: Male			Panel C: Female		
	CRE-FP	LPM	RE Probit	CRE-FP	LPM	RE Probit	CRE-FP	LPM	RE Probit
Log real price	-0.275***	-0.332***	-0.338***	-0.457***	-0.532***	-0.604***	-0.141***	-0.195***	-0.138***
	(0.050)	(0.023)	(0.023)	(0.099)	(0.049)	(0.055)	(0.040)	(0.021)	(0.017)
Post2010*Price	0.012***	0.014***	0.014***	0.019***	0.021***	0.023***	0.006***	0.009***	0.006***
	(0.002)	(0.001)	(0.001)	(0.005)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Post2010	-0.408***	-0.476***	-0.459***	-0.664***	-0.697***	-0.770***	-0.216***	-0.319***	-0.207***
	(0.083)	(0.036)	(0.036)	(0.162)	(0.076)	(0.084)	(0.067)	(0.033)	(0.025)
Male smokers	0.251***	0.242***	0.226***						
	(0.006)	(0.003)	(0.003)						
Education (base: <=primary)									
Some high school	-0.075***	-0.039***	-0.052***	-0.074***	-0.054***	-0.063***	-0.068***	-0.042***	-0.042***
	(0.010)	(0.004)	(0.004)	(0.016)	(0.007)	(0.008)	(0.009)	(0.004)	(0.003)
Matric	-0.139***	-0.090***	-0.094***	-0.189***	-0.145***	-0.150***	-0.102***	-0.066***	-0.058***
	(0.011)	(0.005)	(0.004)	(0.019)	(0.009)	(0.010)	(0.010)	(0.004)	(0.004)
University degree	-0.193***	-0.179***	-0.141***	-0.316***	-0.313***	-0.281***	-0.120***	-0.094***	-0.065***
	(0.012)	(0.009)	(0.005)	(0.033)	(0.019)	(0.013)	(0.013)	(0.009)	(0.004)
Other post matric	-0.159***	-0.108***	-0.103***	-0.242***	-0.185***	-0.179***	-0.101***	-0.067***	-0.058***
	(0.011)	(0.005)	(0.005)	(0.020)	(0.011)	(0.011)	(0.011)	(0.005)	(0.004)
Age group: base 15-19									
20-24	0.106***	0.131***	0.091***	0.213***	0.225***	0.212***	0.061***	0.045***	0.031***
	(0.008)	(0.004)	(0.004)	(0.012)	(0.007)	(0.007)	(0.009)	(0.003)	(0.003)
25-29	0.160***	0.182***	0.139***	0.317***	0.329***	0.319***	0.070***	0.056***	0.039***
	(0.009)	(0.005)	(0.004)	(0.014)	(0.009)	(0.008)	(0.009)	(0.004)	(0.003)
30-39	0.175***	0.195***	0.151***	0.358***	0.355***	0.348***	0.060***	0.059***	0.040***
	(0.008)	(0.006)	(0.004)	(0.015)	(0.012)	(0.009)	(0.008)	(0.005)	(0.003)
40-49	0.183***	0.192***	0.148***	0.362***	0.347***	0.338***	0.070***	0.062***	0.040***
	(0.013)	(0.008)	(0.006)	(0.025)	(0.017)	(0.014)	(0.011)	(0.007)	(0.004)
50-59	0.148***	0.178***	0.135***	0.331***	0.319***	0.310***	0.044***	0.054***	0.037***
	(0.016)	(0.011)	(0.008)	(0.035)	(0.022)	(0.020)	(0.013)	(0.009)	(0.005)
60-69	0.109***	0.144***	0.106***	0.268***	0.254***	0.246***	0.026	0.036***	0.028***
	(0.020)	(0.013)	(0.010)	(0.046)	(0.028)	(0.026)	(0.016)	(0.012)	(0.007)
70+	0.045**	0.090***	0.057***	0.145**	0.141***	0.129***	0.001	0.008	0.016**
	(0.019)	(0.016)	(0.010)	(0.058)	(0.035)	(0.028)	(0.016)	(0.014)	(0.007)
Race: Base African									
Coloured	0.241***	0.247***	0.244***	0.148***	0.186***	0.200***	0.233***	0.296***	0.213***
	(0.019)	(0.006)	(0.009)	(0.025)	(0.012)	(0.014)	(0.020)	(0.006)	(0.010)
Asian/Indian	0.087***	0.092***	0.105***	0.116**	0.110***	0.128***	0.053***	0.080***	0.092***
	(0.027)	(0.013)	(0.016)	(0.049)	(0.025)	(0.030)	(0.018)	(0.013)	(0.018)
White	0.224***	0.162***	0.189***	0.094***	0.105***	0.117***	0.279***	0.218***	0.223***
	(0.026)	(0.008)	(0.011)	(0.030)	(0.015)	(0.018)	(0.028)	(0.008)	(0.015)
Province dummies: Base Western Cape									
Eastern Cape	-0.034**	-0.027***	-0.019***	-0.062**	-0.021	-0.021	-0.008	-0.031***	-0.010***
	(0.014)	(0.007)	(0.006)	(0.025)	(0.013)	(0.014)	(0.011)	(0.006)	(0.004)
Northern	-0.005	-0.005	-0.003	-0.021	-0.010	-0.010	0.006	-0.002	0.001

Cape									
	(0.013)	(0.007)	(0.006)	(0.031)	(0.013)	(0.014)	(0.007)	(0.007)	(0.003)
Free State	-0.003	0.003	0.007	-0.001	0.032**	0.031*	-0.002	-0.020**	-0.002
	(0.015)	(0.009)	(0.007)	(0.027)	(0.016)	(0.017)	(0.013)	(0.008)	(0.005)
KwaZulu-Natal	-0.051***	-0.054***	-0.052***	-0.067**	-0.051***	-0.060***	-0.033***	-0.053***	-0.039***
	(0.014)	(0.007)	(0.006)	(0.027)	(0.013)	(0.014)	(0.012)	(0.006)	(0.004)
North West	-0.025*	-0.028***	-0.015**	-0.033	-0.007	0.000	-0.024**	-0.043***	-0.015***
	(0.015)	(0.008)	(0.007)	(0.029)	(0.016)	(0.017)	(0.011)	(0.008)	(0.005)
Gauteng	-0.037***	-0.029***	-0.022***	-0.061**	-0.017	-0.015	-0.016*	-0.041***	-0.017***
	(0.013)	(0.007)	(0.006)	(0.025)	(0.013)	(0.014)	(0.010)	(0.007)	(0.004)
Mpumalanga	-0.040***	-0.045***	-0.041***	-0.061**	-0.043***	-0.050***	-0.022*	-0.043***	-0.028***
	(0.014)	(0.008)	(0.007)	(0.028)	(0.015)	(0.016)	(0.011)	(0.008)	(0.005)
Limpopo	-0.057***	-0.052***	-0.052***	-0.089***	-0.059***	-0.077***	-0.023**	-0.045***	-0.027***
	(0.015)	(0.008)	(0.007)	(0.029)	(0.015)	(0.017)	(0.012)	(0.007)	(0.005)
Urban	0.040***	0.018***	0.025***	0.044***	0.028***	0.034***	0.042***	0.009***	0.016***
	(0.007)	(0.003)	(0.003)	(0.012)	(0.006)	(0.007)	(0.006)	(0.003)	(0.002)
_cons		1.162***			2.019***			0.762***	
		(0.084)			(0.175)			(0.075)	
N	158163304	97141	97141	72560091	39754	39754	85603213	57387	57387

Robust standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.010

The model is estimated while controlling for the changes in the tobacco market structure since 2010.

Figure 1: Price per pack of 20 cigarettes 1990-2017 (Source: Authors' own computations using price data from Statistics South Africa and adjusted income data from World Development Indicators (WDI)).



Appendix: Control function estimation

v2hat are residual error obtained after estimating the reduced form for price (pooled across t) and included in the second stage pooled fraction probit QMLE regression model. Insignificant v2hat rejects the hypothesis that the price variable is endogenous.

Table 5: Correlated Random effects Fractional probit model using control function approach

Dependent variable: prevalence	Panel A; Includes estimates of illicit cigarette market			Panel B: Includes market structural break dummy		
	All	Male	Female	All	Male	Female
Log real price	-0.261***	-0.381***	-0.160***	-0.335***	-0.537***	-0.202***
	(0.036)	(0.075)	(0.032)	(0.031)	(0.065)	(0.029)
Illicit*price	0.010***	0.021***	0.006**			
	(0.003)	(0.007)	(0.003)			
Illicit market (%)	-0.034***	-0.075***	-0.020**			
	(0.012)	(0.026)	(0.010)			
Post2010*Price				0.011***	0.017***	0.006***
				(0.002)	(0.005)	(0.002)
Post2010 dummy				-0.384***	-0.566***	-0.198***
				(0.076)	(0.160)	(0.062)
Gender (1=Male; 0=Female)	0.231***			0.231***		
	(0.002)			(0.002)		
Education (base: <=primary)						
Some high school	-0.058***	-0.072***	-0.050***	-0.059***	-0.073***	-0.050***
	(0.003)	(0.006)	(0.003)	(0.003)	(0.006)	(0.003)
Matric	-0.114***	-0.175***	-0.076***	-0.115***	-0.176***	-0.077***
	(0.004)	(0.007)	(0.003)	(0.004)	(0.007)	(0.003)
University degree	-0.181***	-0.311***	-0.096***	-0.183***	-0.311***	-0.096***
	(0.005)	(0.010)	(0.004)	(0.005)	(0.010)	(0.004)
Other post matric	-0.128***	-0.213***	-0.078***	-0.129***	-0.214***	-0.079***
	(0.004)	(0.008)	(0.004)	(0.004)	(0.008)	(0.004)
Age group: base 15-19						
20-24	0.114***	0.214***	0.047***	0.117***	0.220***	0.049***
	(0.004)	(0.008)	(0.004)	(0.004)	(0.008)	(0.004)
25-29	0.170***	0.339***	0.062***	0.173***	0.341***	0.062***
	(0.004)	(0.008)	(0.004)	(0.004)	(0.008)	(0.004)
30-39	0.190***	0.386***	0.062***	0.188***	0.381***	0.061***
	(0.003)	(0.007)	(0.003)	(0.004)	(0.008)	(0.003)
40-49	0.196***	0.391***	0.070***	0.188***	0.374***	0.066***
	(0.007)	(0.014)	(0.006)	(0.007)	(0.015)	(0.006)
50-59	0.201***	0.387***	0.074***	0.187***	0.360***	0.068***
	(0.011)	(0.022)	(0.009)	(0.010)	(0.022)	(0.009)
60-69	0.176***	0.342***	0.061***	0.156***	0.304***	0.053***
	(0.014)	(0.031)	(0.012)	(0.013)	(0.029)	(0.011)
70+	0.115***	0.204***	0.044***	0.093***	0.161***	0.035***
	(0.016)	(0.036)	(0.014)	(0.015)	(0.033)	(0.013)
Race (base: Africans)						
Coloureds	0.234***	0.186***	0.227***	0.235***	0.184***	0.227***
	(0.005)	(0.010)	(0.007)	(0.005)	(0.009)	(0.007)
Indians	0.124***	0.155***	0.111***	0.124***	0.154***	0.110***
	(0.012)	(0.021)	(0.015)	(0.012)	(0.021)	(0.015)
Whites	0.209***	0.128***	0.273***	0.210***	0.128***	0.273***
	(0.008)	(0.013)	(0.012)	(0.008)	(0.013)	(0.012)
Province dummies: Base Western Cape						
Eastern Cape	-0.028***	-0.037***	-0.017***	-0.025***	-0.031***	-0.016***

	(0.005)	(0.010)	(0.004)	(0.005)	(0.010)	(0.004)
Northern Cape	-0.007	-0.027**	0.003	-0.005	-0.020**	0.004
	(0.005)	(0.011)	(0.003)	(0.004)	(0.010)	(0.003)
Free State	0.015**	0.041***	-0.001	0.016***	0.048***	0.000
	(0.006)	(0.013)	(0.005)	(0.006)	(0.012)	(0.005)
KwaZulu-Natal	-0.065***	-0.068***	-0.065***	-0.064***	-0.064***	-0.065***
	(0.004)	(0.010)	(0.004)	(0.004)	(0.010)	(0.004)
North West	-0.023***	-0.009	-0.030***	-0.021***	-0.003	-0.029***
	(0.006)	(0.013)	(0.005)	(0.006)	(0.012)	(0.005)
Gauteng	-0.019***	-0.017	-0.018***	-0.016***	-0.012	-0.017***
	(0.005)	(0.011)	(0.004)	(0.005)	(0.011)	(0.004)
Mpumalanga	-0.045***	-0.057***	-0.042***	-0.042***	-0.044***	-0.040***
	(0.006)	(0.013)	(0.006)	(0.006)	(0.012)	(0.006)
Limpopo	-0.070***	-0.093***	-0.056***	-0.069***	-0.086***	-0.055***
	(0.006)	(0.013)	(0.006)	(0.006)	(0.012)	(0.006)
Reduced form errors (v2hat)	0.131	-0.028	0.078	0.096	0.192	0.101
	(0.099)	(0.210)	(0.077)	(0.077)	(0.164)	(0.064)
N	97296	39754	57387	97141	39815	57387

* p<0.10, ** p<0.05, *** p<0.010

Notes: (i) bootstrapped standard errors from fractional probit estimation in parenthesis (ii) The model includes time averages of education and age to allow them to be correlated with individual unobserved heterogeneity (iii) excise and sales taxes are the excluded instruments in the control function estimation. The results obtained from the control function specification are remarkably close to the average partial effects (APEs) estimates obtained using the pooled fractional Probit and simple panel LPM.