

INTRODUCTION

Tobacco use is the paramount preventable tenet of deaths around the world. Each year an estimated figure of seven million deaths are ascribed to the use of tobacco (WHO, 2017). On average, users of tobacco lose around 15 years of their life. Up to 50 percent of all tobacco users die untimely because of tobacco-related issues (Goodchild, 2018). The majority of these deaths happened in the middle and low-income countries, which reported nearly 80 percent of tobacco-related deaths. The costs imposed by tobacco use on the economy are colossal, adding up to more than US\$ 1.4 trillion in medical care costs and bringing about lost productivity which is around 1.8 percent of the world's GDP and 40% of total spending on school education by the world (Goodchild, 2018). Despite these facts, tobacco stays the absolute most generally accessible and available drug whose purchase is legitimate all over.

Curbing the use of tobacco has been the major goal of Sustainable Development Goals (SDGs) as enshrined in the third goal of good health and well-being. It has been argued that reduced tobacco consumption generates positive externalities by reducing premature mortality and toppling down the number of deaths due to non-transmissible diseases (NTDs) such as cancer, COPDs, and other cardiovascular diseases. Despite of its ill-effects, every tenth adult (10.4 percent) currently smokes tobacco and every fifth adult (21.4 percent) currently uses smokeless tobacco in India and is estimated to cause more than 1.3 million deaths (GATS, 2017).

Beedi recorded the highest proportion of smoked tobacco consumption in India, particularly for people belonging to lower-income groups, with smoking intensity as high as eight times than smoking cigarettes. Despite its use at the large scale and toxicity attached with it, they usually go neglected. One of the archaic rationales for higher beedi consumption was in response to the Swadeshi Movement, where cigarette imports were banned and native beedi got more prominent openness, and its intended use increased from other tobacco products (Kannan, 1988). A study conducted by (Mishra S, 2016) however noticed that cigarettes were uprooting beedis among men because of increasing income and expanded affordability and this has been consistently observed among people belonging to the bottom of the society.

India's population is expected to grow by 2030 and will be around 400 million, the majority of them in the age scope of 15-59 years (Goodchild, 2018). It has been cited that the marketing strategies aim to target the younger population and most of the times children are

used by women to buy tobacco for them. This in turn exposes children towards tobacco at a very tender age bringing about early inception and addiction. GATS 2 reported that tobacco marketing strategies are usually targeted towards youngsters and most of the anti-tobacco messages are not understood effectively by the sub-population. A study in this context on schoolchildren in the age group of 13 to 15 years by (Warren W, 2006) have shown that one in every five individuals started smoking because of the peer pressure and the impact of advertisements and other promotional activities thereon.

Having understood the ill effects of tobacco consumption, it is quintessential to take a broader view and empirically examine the factors influencing the cigarette consumption by the youngsters. Cigarette consumption could be influenced by both demand as well as supply side factors but this paper has only taken demand side constraints into consideration. The present paper has analysed the average cigarette and rolled tobacco consumption for the age group of 15-24 years. The remaining of the term paper is organized as follows. Following the introduction, Section II reviews the available literature. Section III elaborates the data structure and methodology explaining the source of data, dependent as well as independent variables and the model used for conducting analysis. Data analysis and interpretation has been done in Section IV of this paper followed by conclusion and policy suggestions.

SECTION II: REVIEW OF THE AVAILABLE LITERATURE

The factors influencing the consumption of tobacco products can be categorized into four broad categories. Concerning personal level characteristics, the sex of the respondent is quite influential and in practically all nations, the consumption of tobacco products by females are less than their male counterpart. There exists a number of competing and contrasting hypothesis that helps explain the gender differences in smoking behavior. Lombroso and Ferrero (1895) gave biological explanations and believe that females consider smoking tobacco because of hormonal imbalances and possess a number of male characteristics. In response to the theory, Smart (1979) expels a well-balanced argument by solving the unemployment-drug puzzle and argues that women experience the same humiliating experience as experienced by men at times of unemployment and when permissible or legalized labor market doesn't permit women entry, illegitimate labor market becomes more attractive which spills over into anger, excessive consumption of drugs such as alcohol, and cigarettes and further exasperation. These arguments stand in sharp contrast to Clark (1980) who argues that to account for the gender gap in tobacco consumption, one has to take a sociological perspective, rather than just

accounting for a biological phenomenon. Thus it can be seen that there is a sense of discontentment to consider smoking tobacco across genders.

Tobacco consumption seems to be more profound among the uninformed individuals in the country. It has often been cited in literature that education plays a significant role in the consumption of tobacco. Uninformed and uneducated citizens in India were at a higher risk of consuming tobacco. This can be ascribed to information asymmetry among the uneducated individuals. Being poor was fundamentally connected with a higher risk of consumption of smokeless tobacco among males, and use of smokeless tobacco, and its dual-use among females in India. The connection between these socioeconomic markers and tobacco utilization is coherent to what is noticed in developed nations and other studies conducted in the context of India (Subramanian SV, 2004). Given their higher degree of fatalism, destitute individuals in India are bound to start tobacco utilization at a very early stage of their life. Tobacco prices are another important determinant. The intensity of smoking Beedis are eight times higher than cigarettes as the average cost of beedi is around (Rs 0.4) concerning cigarettes (Rs 3.1) (Mishra S, 2016). It has often been argued that increased taxes on tobacco consumption will increase the price of cigarettes and act as a major factor for cessation. However from the Indian setting, since the average costs are lower, users often switch towards their substitutes.

Concerning demographic level characteristics as far as castes and ethnic groups are concerned, the beliefs of society as a whole towards tobacco influence how people belonging to these groups view cigarette and alcohol consumption. For instance, consuming alcohol and smoking tobacco has been considered a sin in Islam whereas it often argued that the trend of smoking is more obvious among people belonging to SC and ST and other lower caste groups (Surani NS, 2012). Social interaction also has an important role to play for the increased tendency of working people to prefer smoking. Guidelines that restrict smoking in enclosed areas, including workspaces, reduces the intensity of smoking. One might say that these policies are significant not exclusively to forestall passive smoking but also to diminish the impact of social interaction. Using either smoke or smokeless tobacco at home can likewise be considered socializing since smoking at home additionally affect others. One more issue for social connection is schools. Almost certainly, children of school-going age may impact and incent each other on this issue to maintain the so-called affiliation needs.

Concerning knowledge and perception characteristics, youths throughout the world are subjected to a barrage of images, many designed specifically to encourage impulse purchasing

of tobacco products without attention to legally required health warnings (Anon., 2002). Medical personnel are embodiments of tobacco abeyance and perform a critical part in the triumph of abeyance services and a significant role in tobacco cessation, with a chance for aiding tobacco consumers to quit and forestall initiation. However, studies have shown that around half of the Indian male doctors smoke (Cattaruzza, 2013). On a similar note, studies on the consumption of tobacco by healthcare professionals (physicians, dentists, nurses, and pharmacists) highlighted a higher utilization rate, particularly among male medical students (Surani NS, 2012). This is inimical not just for the concerned individual but to society as a whole as it reduces their ability to convey meaningful anti-tobacco messages in their lifetime clinical practises.

As can be seen there are many studies on tobacco consumption and the socio-economic roots of it. India, being a country of diversity, has numerous socio-economic issues, which can affect the smoking levels of individuals. Presence of large number of poor and illiterate people, religious and caste-wise beliefs, area of residence and marital status can all be potential factors affecting tobacco consumption. Hence in this circumstance, it become imperative to conduct this study, taking three dimensions: the personal, demographic and, knowledge and perception factors that can impact the smoking habits of people in India. A graph illustrating the relationship of above factors with current percentage of tobacco users is illustrated below.

SECTION III: METHODOLOGY

Data Source and Structure

The study has been conducted by using Global Adult Tobacco Survey (GATS) 2016-17 data. The Ministry of Health and Family Welfare, GoI, delegated the Tata Institute of Social Sciences (TISS), as the coordinating body to administer GATS 2016-17 survey in India. The chief ambition of the report was to measure the ramifications of tobacco restrictive measures through the administration of distinct provisions of COTPA 2003 and its canons. GATS is a large-scale, geographically clustered multi-stage sampling design. GATS 2016-17 covers representative data for 74,037 households, with 33,772 men and, 40,265 men and provides data on socio-economic characteristics of the households, usage of tobacco products (both smoke and smokeless products), exposure to second-hand smoke, media, knowledge, and attitude towards consumption of alcohol.

The 2016-17 report noticed that 28.6 percent of the population use tobacco in any form. Among them, 24.9 percent use tobacco on a daily basis and 3.7 percent of them use it occasionally (GATS, 2017). A comparison with the GATS-1 report highlighted a decline of 6 percent in tobacco utilization and also the NFHS-4 has shown a decline in the prevalence rate with respect to the NFHS-3 report. However, a point worth noticing is that there has been a one-year increase in the commencement of tobacco use, and between NFHS-3 and NFHS-2, consumption of all forms of tobacco have increased with the greatest numbers being observed between 15-24 years age group. Therefore it is quintessential to do an assessment of daily tobacco consumption and look for the factors encouraging or discouraging tobacco use among the youngsters.

GATS 2016-17 report provides data for consumption of various smoke as well as smokeless tobacco products. The study has taken only smoking tobacco products into consideration. Mediating smoking tobacco products, there are different categories namely manufactured cigarettes, rolled tobacco, bidis, cigars, cheroots, or cigarillos, and hookah. However, the prices of smoking products vary with an average price of bidis as low as Rs. 0.20 per stick and an average of Rs. 700 per hookah session (IGTC, 2017). Thus it can be argued that smoking tobacco products vary by the status a person holds in the social stratum of the society. It has been noticed that subjects belonging to higher income groups are not seen smoking bidis and they drift for conspicuous consumption which includes cigars, cheroots, and cigarettes; whereas people belonging to the lower-income category are often seen smoking

bidis due to its lower prices. To account for the price differential, only cigarette and rolled tobacco consumption have been taken into consideration as their prices on average are similar with Rs. 12 and Rs.13 respectively (IGTC, 2017).

Dependent variable

The study has been restricted by undertaking observations only for the age-group of 15-24 years to examine the characteristics that may influence the daily cigarette and rolled tobacco consumption of young people. The dependent variable therefore is the count of daily cigarette and rolled tobacco consumption.

Determinants of cigarette and rolled tobacco consumption

The explanatory variables have been selected on the basis of literature review that has been conducted and on the basis of their relative importance. The independent variables have been classified into three broad categories namely personal, demography, and knowledge and perception characteristics.

- **Personal Characteristics** which include the gender (male, female), education status (no schooling, primary education, secondary education, graduate, and postgraduate) and marital status of the respondents (single, married, divorced, separated and widowed).
- **Demographic Characteristics** which include the area of residence (urban, rural), religion (Muslim, Hindu, Christian, and others), and caste/tribe (general, other backward castes, scheduled tribe, and scheduled caste).
- **Knowledge and Perception Characteristics** which includes knowledge that exposure to smoking causes lung cancer, COPDs, and other cardiovascular diseases (yes, no), knowledge about various media promotion activities towards smoking (yes, no), and media awareness against tobacco consumption (yes, no).

Media exposure has been captured by taking into account newspapers and magazines, or watching TV or listening to radios. A person has received media awareness against tobacco use if they have noticed information about the dangers of smoking tobacco in at least one of these - newspaper or magazine, TV, radio, internet, cinemas, billboards, public walls. Dummies for each variable was created with yes being equal to 1 and no being equal to 0. All such dummies were summed up to create a total awareness variable. If the total was greater than equal to 1 then the person was said to have received awareness against smoking. Promotion

activities and Knowledge about health impacts of smoking tobacco have been captured in a similar way. Mean value of the explanatory variables (in percentages) are shown in Table 1.

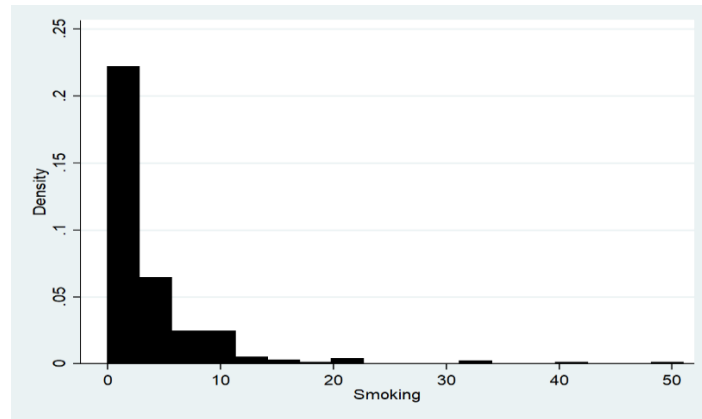
Table 1: Mean value of the explanatory variables (in percentages)

		India	Age Group (15-24 years)
Personal characteristics	Gender of the respondent		
	Male	20.22	23.21
	Female	2.05	2.44
	Education status		
	No Schooling	12.99	13.48
	Primary Education	14.37	15.48
	Secondary	8.4	10.74
	Graduate	4.73	5.76
	Post Graduate	2.81	3.09
Demographic characteristics	Marital Status		
	Single	4.66	10.32
	Married	11.74	12.40
	Divorced	12.05	12.54
	Separated	13.27	14.43
	Widowed	7.64	7.67
	Area of residence		
	Urban	7.29	8.28
	Rural	12.07	14.04
	Religion		
	Hindu	9.26	10.83
	Muslim	11.76	14.11
	Christian	17.66	19.47
	Others	9.09	9.94
	Caste/Tribe		
	SC	10.89	12.88
	ST	16.97	19.22
	OBC	8.35	9.84
	General	8.83	10.17
Knowledge and Perception characteristics	Knowledge About Illness		
	Yes	10.32	11.94
	No	11.4	12.52
	Knowledge about media promotion activities		
	Yes	12.17	14.26
	No	9.97	11.50
	Media awareness against tobacco consumption		
	Yes	10.14	11.85
	No	11.06	12.30
Total	Number of respondents	73520	60,286

Source: Computed by the Author using GATS 2016-17 data

Figure 1 represents the count of the daily cigarette and rolled tobacco utilization for the age group of 15-24. It is noticeable that the distribution is positively skewed. The summary statistics of the dependent variable show mean and variance of 3.22 and 31.72 respectively. The variance for the dependent variable is around nine times higher than the mean value. Thus above information shows that there are a large proportion of zeros and because variance is greater than mean, there is overdispersion in the data. Excess zeros can be accorded to both structural as well as sampling zeros. Sampling zeros could result because the following two reasons: some subjects who usually smoke but did not smoke in the period of survey; and some participants might not have told the interviewer due to personal reasons. For instance, some youngsters would not have talked about it openly when their parents were around.

Figure 1: Count of daily consumption of cigarettes and rolled tobacco for 15-24 age group



Source: Generated by the Author using GATS 2016-17 data

The analysis has been conducted with the help of zero-inflated negative binomial model (ZINB). To understand ZINB, it is important to understand Negative Binomial (NB) model. When there is over-dispersion i.e. when variance is greater than mean, NB distribution is used. It contains a dispersion parameter (k) to model individual level heterogeneity and thus manoeuvre over-dispersion. The probability density function for a Negative Binomial Distribution has the form:

$$P(Y_i/X_i) = \frac{\Gamma(Y_i + \frac{1}{k})}{\Gamma(Y_i + 1) \Gamma(\frac{1}{k})} \left(\frac{1}{1+k\mu}\right)^{1/k} \left(\frac{k\mu}{1+k\mu}\right)^{Y_i}$$

The highlighting feature of Negative Binomial (NB) distribution is that its theoretical distribution is skewed allowing for a variance to be greater than the mean. However, when the value of (k converges to zero, the NB model resembles the Poisson model. Thus, NB model can be considered as an overdispersed Poisson model.

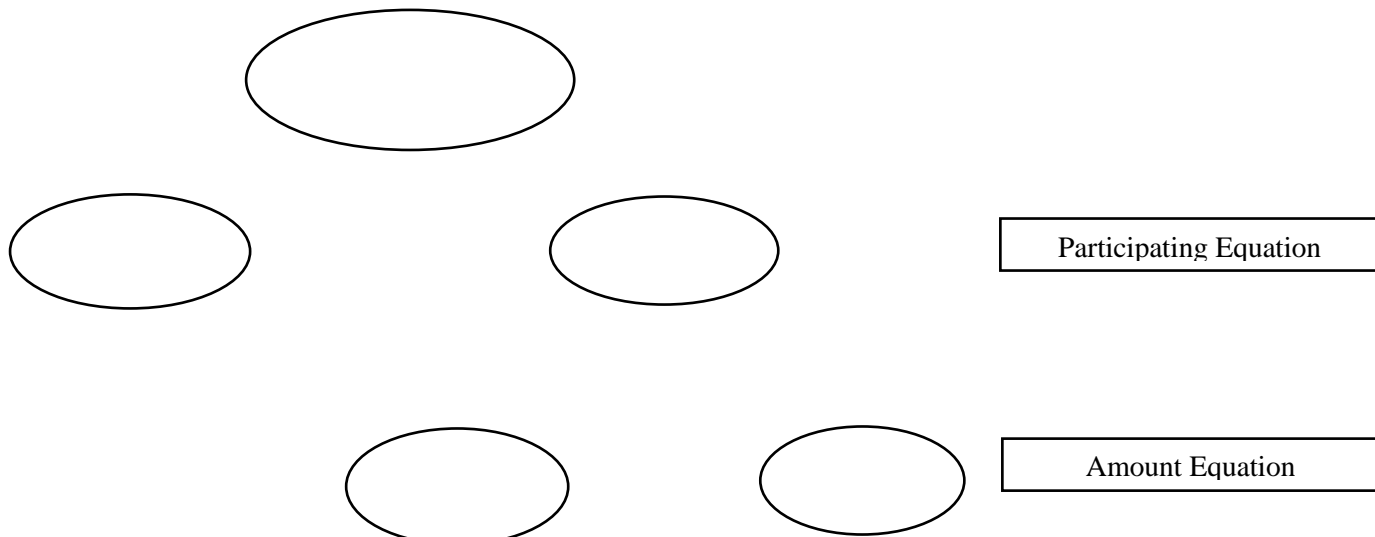
In the last few years, however, to address excess zeros, zero-inflated and hurdle models have become quite popular. In the words of (Cameron, 2007), these models are most often regarded as finite combination models with a degenerate dispersion whose mass is consolidating at zero. Immoderate zeros emerge when the occasion of interest is not faced by several individuals selected for the study under consideration. For scanty events or events with small mean value, Poisson or NB model can foresee a large number of zeros, with NB having an option to anticipate even more zeros.

Hence, zero inflated model articulates two separate models, one for the participating equation and one for the amount equation. In the participating equation, we take into account the structured zeros (generating only zero count, e.g. non-smokers), whereas in the amount equation, we account for random zeros (e.g. among those who smoke, but did consider it on a particular day). Thus, it can be argued that zero inflated models combines two different modelling processes; the logistic regression model and NB distribution model. ZINB model can be explained with the help of a give framework:

From the given framework, we can understand that ZINB is a two stage model: the logistic regression model predicting structured zeros (categorising as smokers and non-smokers), whereas the NB model takes into account both positive counts and as well as random zeros. The ZINB model can be written as:

$$P(Y) = \begin{cases} \omega + (1 - \omega) \left(\frac{1}{1+k\mu} \right)^{\frac{1}{k}} & \text{if } Y = 0 \\ (1 - \omega) \frac{\Gamma(Y + \frac{1}{k})}{\Gamma(Y+1) \Gamma(\frac{1}{k})} \left(\frac{1}{1+k\mu} \right)^{\frac{1}{k}} \left(\frac{k\mu}{1+k\mu} \right)^Y & \text{if } Y \geq 1 \end{cases}$$

Where μ indicates the mean value and ω denotes the probability for a given subject having zero count. In this model, $k \geq 0$ is called as dispersion parameter which is presumed to be independent of the explanatory variables. A simple ZINB framework can be represented as:



SECTION IV: EMPIRICAL ANALYSIS, RESULTS AND INTERPRETATION

One of the challenges in capturing the determinants of childhood immunization across all four states is that there exists a possibility of collinearity among the explanatory variables and it is quintessential to look at the correlation matrix to account for multicollinearity. However, in real life and in social science research studies, explanatory variables are directly or indirectly correlated with each other. In the context of present study, for example, education status is closely associated with knowledge about illness. Similarly, education and media awareness about the ill effects of tobacco are correlated with each other. However if the explanatory variables are correlated but none can be excluded from theoretical point of view, removing an explanatory variable to solve the problem of multicollinearity is not often recommended unless the correlation coefficient is greater than 0.8. The constructed correlation matrix among the explanatory variables shows that the correlated values are less than threshold value and can be retained (Appendix). Also, there is considerable degrees of freedom due to large dataset (close to 12,000 observations).

A Negative Binomial regression model has been estimated and results are presented in Table 2. As noticed there is a sense of discontentment to consider the factors affecting tobacco consumption for males and females, separate models have been estimated for them since only 2 percent of the female respondents smoke in the given dataset. It is important to note here that the value of $\ln\alpha$, which is the overdispersion parameter is significant in both the models which clearly implies that ZINB model is a correctly fitted model.

On the note of personal characteristics, it can be seen that the result of education attainment increases the rate of expected count of smoking tobacco consumed, holds for both Males and Females.

In comparison to urban areas, both males and females living in rural areas have a lower rate of expected count of smoking tobacco consumed.

As per Religion, Muslim men have a higher expected count while Muslim women have a lower expected count in comparison to Hindu men and women. Both Christian male and females have a higher expected count. Religious rules regarding smoking and substance use can be clearly seen in the results while being consistent with gender norms of said religions.

Marital Status of a female truly defines the expected count as married, divorced and separated women are seen to have a decrease in their expected count consumed in comparison to single women. For males this result is reverse as married men have an increase in rate of incidence ratio in comparison to their single counterparts.

Caste doesn't provide any clear patterns but as per results people belonging to Scheduled Tribes smoke higher counts.

Other factors and characteristics which define the average smoking count could be the influence of media in people's lives which might influence their decisions.

Males who have received awareness about the ill effects of smoking from their surroundings and media do smoke less to a certain degree. Media can be seen as a manipulator in changing people's behaviour.

Similarly a significant number of Males who are aware about the major diseases caused by smoking (Heart Attack, TB, Stroke etc.) do not smoke as much as people who are ignorant or unaware about such illness. Such information can be attributed to media campaigns about making people aware about the ill effects of smoking.

Our participation variable (hurdle variable) which categorises the sampled population into smokers and non-smokers. Such zero creating variable is necessary to include to deflate the excess zeros present in the survey.

The hurdle variable

Level of characteristics	Gender Variables	Males (IRR)	Females (IRR)
Personal level characteristics	Education Status		
	Primary Education	1.201 (0.096)	1.738 (0.384)
	Secondary Education	1.53 (0.146)	1.315 (0.4)
	Graduate	1.917 (0.319)	3.731 (4.263)
	Post Graduate	2.255 (0.561)	4.738 (9.399)
Demographic characteristic	Marital Status		
	Married	1.015 (0.121)	0.719 (0.632)
	Separated	0.891 (0.349)	0.273 (0.363)
	Divorced	1.2 (0.521)	0.28 (0.304)
	Widowed	0.63 (0.146)	1.085 (0.962)
	Area of Residence		
	Rural	0.782 (0.062)	0.324 (0.255)
	Religion		
	Muslim	2.215 (0.218)	.622 (0.168)
	Christian	2.573 (0.357)	2.228 (0.596)
	Others	2.041 (0.305)	1.106 (0.351)
	Caste/Tribe		
	ST	1.116 (0.126)	2.816 (0.808)
	OBC	1.014 (0.09)	1.613 (0.418)
	General	1.07 (0.104)	1.105 (0.32)
Knowledge and Perception Characteristics	Knowledge about illness		
	Yes	0.612 (0.1)	1.94 (0.3)
	Media Promotion		
	Yes	0.917 (0.064)	1.389 (0.385)
	Media Awareness		
	Yes	0.855 (0.068)	1.923 (0.34)
	Constant	4.395 (1.127)	0.779 (0.719)
	Inflate Hurdle		
	Yes	-0.509	-0.041
	Inalpha	0.931	1.336

	Mean Dependent variable	2.796	2.675
	Prob. > chi2	0.000	0.000
	SD Dependent variable	7.297	9.312
	Chi-square	278.680	128.803
	Akaike Crit. (AIC)	22340.423	2739.317

Results have shown that for Indian youngsters

This could be because of the restriction on the age that we have taken into consideration.

Policymakers, guardians, and school directors have a significant role to play by encouraging youngsters to deflect and surrender smoking. In spite of the fact that there are forbiddances on the sale of cigarettes to youngsters, it is quintessential to control this better. The other significant variable was sex. In practically all nations, the consumption of tobacco products by females are less than their male counterpart. However, from a policy viewpoint, it is of great interest whether the sex contrasts in smoking predominance could fundamentally be clarified by contrasts in core attributes or by different smoking behaviour. Having sufficient evidence on the sources of the differences in cigarette utilization among males and females might help, for instance, make antismoking strategies more compelling by empowering the policymakers to address the specific target group.

CONCLUSION AND POLICY IMPLICATIONS

Substantial, evidence-based, pictorial wellbeing admonitions on all tobacco packs are effective mediations, as they are noticed by both users as well as non-users of tobacco. The enormous pictorial health warning should proceed; pictures and messages should be changed at regular intervals of time to clinch supported impact. Endeavours should be made to make the warning labels more effective by giving information about diseases brought about by tobacco, including cancer, stroke, coronary illness, lung sickness, etc.

The pervasiveness of tobacco usage among children of 15-24 age group and specifically students of school-going age stays a space of concern. There is a need to reinforce coordination with higher educational institutions, youth, and local area-based associations to arrive at this segment of the populace all the more viably. Schools and colleges should take a holistic approach sensitizing students towards the environment and tobacco consumption. Imaginative initiatives should be taken in this regard to address this group, complementing custom

messages through web-based media and other emerging platforms. Tobacco-free institutional guidelines need to be embraced and executed across the country.

There is a need to put forth distinctive efforts to reach out to the people placed at the lowest ladder of social stratum, the most un-instructed, and the old since they are the ones least composed to bear the results of addiction. In this manner, tobacco tax policies should be surveyed consistently to decrease the affordability of tobacco products to protect vulnerable population groups. Civil society and private area commitment are urgent to contact different stakeholders, particularly in enervating, transfer of knowledge, instructing, social media campaigns, espousal, and building partnerships.

APPENDIX

Pairwise Correlation Matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Area of Residence	1.000							
(2) Caste/Tribe	-0.150	1.000						
(3) Knowledge about illness	-0.063	0.043	1.000					
(4) Media Promotion Activities	-0.050	0.045	0.047	1.000				
(5) Media Awareness Activities	-0.233	0.091	0.193	0.176	1.000			
(6) Religion	0.018	-0.071	0.010	-0.034	-0.015	1.000		
(7) Education Status	-0.303	0.183	0.130	0.103	0.370	0.029	1.000	
(8) Marital Status	0.020	-0.007	-0.059	-0.052	-0.129	-0.031	-0.222	1.000