

JOSÉ CARLOS MUNHOL JÚNIOR

**PEER SMOKING AMONG BRAZILIAN ADOLESCENTS: ARE THE EFFECTS OF
SOCIAL INTERACTION IN SCHOOL CLASSES RELEVANT?**

Dissertation submitted to the Applied Economics
Graduate Program of the Universidade Federal de
Viçosa in partial fulfillment of the requirements
for the degree of *Magister Scientiae*.

Adviser: Maria Micheliana da Costa Silva

**VIÇOSA - MINAS GERAIS
2022**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

M966p
2022
Munhol Júnior, José Carlos, 1992-
Peer smoking among brazilian adolescents: are the effects
of social interaction in school classes relevant? / José Carlos
Munhol Júnior. – Viçosa, MG, 2022.
1 dissertação eletrônica (52 f.): il. (algumas color.).

Texto em inglês.

Inclui apêndices.

Orientador: Maria Micheliana da Costa Silva.

Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Economia Rural, 2022.

Referências bibliográficas: f. 43-47.

DOI: <https://doi.org/10.47328/ufvbbt.2022.452>

Modo de acesso: World Wide Web.

1. Fumo - Vício - Aspectos sociais. 2. Econometria.
3. Jovens - Saúde e higiene. I. Silva, Maria Micheliana da Costa,
1986-. II. Universidade Federal de Viçosa. Departamento de
Economia Rural. Programa de Pós-Graduação em Economia
Aplicada. III. Título.

CDD 22. ed. 362.2961

Bibliotecário(a) responsável: Euzebio Luiz Pinto CRB 6/3317

JOSÉ CARLOS MUNHOL JÚNIOR

**PEER SMOKING AMONG BRAZILIAN ADOLESCENTS: ARE THE EFFECTS OF
SOCIAL INTERACTION IN SCHOOL CLASSES RELEVANT?**

Dissertation submitted to the Applied Economics
Graduate Program of the Universidade Federal de
Viçosa in partial fulfillment of the requirements
for the degree of *Magister Scientiae*.

APPROVED: April 28, 2022.

Assent:

JOSÉ CARLOS MUNHOL JÚNIOR

José Carlos Munhol Júnior

Author

Maria Micheliana da Costa Silva

Maria Micheliana da Costa Silva

Adviser

ACKNOWLEDGEMENTS

I would like to thank all those who contributed so that this research could be carried out, such as my family, master's advisor Maria Micheliana da Costa Silva, professors from the Department of Rural Economics.

To the Federal University of Viçosa, for the opportunity to complete the postgraduate course.

To the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), to granting the scholarship.

ABSTRACT

MUNHOL Jr, José Carlos, M.Sc., Universidade Federal de Viçosa, April, 2022. **Peer smoking among brazilian adolescents: are the effects of social interaction in school classes relevant?** Adviser: Maria Micheliana da Costa Silva.

This research aims to analyze how smoking among young Brazilians is influenced or not by the effects of peers at the level of school classes (colleagues), weighted by individual characteristics, contextual to individuals and institutions in the school households surveyed. The contribution of this study is to fill an existing gap in the Brazilian literature, where the existence of drug use pairs in general is still little explored by empirical academic research in the area of health and social economy. The analytical and econometric approach taken here aggregates issues related to non-linear models of consumption choices and lagged social interaction effects, where the issue of causality is resolved with such empirical modeling. Using individual-level data from the 2012 and 2015 of Pesquisa Nacional de Saúde do Escolar, estimates were performed using a multinomial logit model and an additional series of models were estimated by inputting control variables, adolescent family background, effects of year and region to mitigate the biases of the social interaction effects of school classes, in addition to the temporal effect for the year 2012. The results found showed that increasing the proportion of peers in a smoking room increases the relative risks of be an occasional and regular user, for the model with all variables (RRR=2,501; 2,064); while for the lagged pair effect, the relative risks of being an occasional smoker decrease in the 2015 period (RRR=0.614; 0.493). The estimation of the various econometric models, where the other regressors are added gradually, showed a reduction in the relative risks of smoking occasionally and regularly, albeit in a discrete way and still being statistically significant to fully explain smoking.

Keywords: Peer effects. Smoking. Health adolescents. Pseudo-panel. Social econometrics

RESUMO

MUNHOL Jr, José Carlos, M.Sc., Universidade Federal de Viçosa, abril de 2022. **Tabagismo entre pares sobre adolescentes brasileiros: os efeitos de interação social nas turmas escolares são relevantes?** Orientadora: Maria Micheliana da Costa Silva.

Esta pesquisa tem como objetivo analisar como o tabagismo entre jovens brasileiros é influenciado ou não pelos efeitos dos pares no nível das turmas escolares (colegas), ponderados por características individuais, contextuais aos indivíduos e instituições nos domicílios escolares pesquisados. A contribuição deste estudo é preencher uma lacuna existente na literatura brasileira, onde a existência de pares de uso de drogas em geral ainda é pouco explorada por pesquisas acadêmicas empíricas na área da saúde e economia social. A abordagem analítica e econométrica aqui feita agrega questões relacionadas a modelos não lineares de escolhas de consumo e efeitos de interação social defasados, onde a questão da causalidade é resolvida com tal modelagem empírica. Usando dados em nível individual da Pesquisa Nacional de Saúde do Adolescente de 2012 e 2015, as estimativas foram realizadas a partir de um modelo logit multinomial e uma série adicional de modelos foram estimados inserindo variáveis de controle, antecedentes familiares do adolescente, efeitos de ano e região, para mitigar os vieses dos efeitos da interação social das classes escolares, além de efeito temporal para o ano de 2012. Os resultados encontrados mostraram que o aumento da proporção de pares em uma sala de fumantes aumenta os riscos relativos de ser usuário ocasional e regular, para o modelo com todas as variáveis (RRR=2.501; 2.064); enquanto para o efeito de par defasado, os riscos relativos de ser fumante ocasional diminuem no período de 2015 (RRR=0.614; 0.493). A estimação dos vários modelos econométricos, onde os outros regressores são adicionados gradativamente, mostrou uma redução nos riscos relativos de fumar ocasionalmente e regularmente, embora de forma discreta e ainda sendo estatisticamente significativo para explicar plenamente o tabagismo.

Palavras-chave: Efeitos de pares. Tabagismo. Saúde dos adolescentes. Pseudo-painel. Econometria social

LIST OF TABLES

Table 1. Variables used in the research.....	22
Table 2. Tobacco consumption (industrialized cigarettes) in the entire sample	23
Table 3. Drugs versus participation in labour market.....	28
Table 4. Model I (benchmark) - Individual and peer effects that explain individual behavior.....	31
Table 5. Model (II) only with the effects of pairs with regressors.....	32
Table 6. Model III - Background of each individual's family and peer effects to explain smoking.....	33
Table 7. Model V - Presence of consumption of complementary drugs (Other drugs) along with other variables to explain smoking.....	34
Table 8. Model V (presence of consumption of other drugs)	35
Table 9. Model VI (controlling for the effect of year and Major Brazilian Regions)	36
Table 10 – Model VII - Pseudo-panel with lagged mean social interaction effect.....	37

LIST OF FIGUERES

Graph 1. Tobacco smoke due a sex.....	24
Graph 2. Occasional and regular smoker due to participation in the labor market.....	25
Graph 3. Occasional and regular smokers in large Brazilian regions	26
Graph 4. Occasional and regular smoker's due mother's education.....	27
Graph 5. Physical activity due occasional and regular smokers.....	29
Graph 6. Incidence of smoking due to ban policies in schools and lack of information on this.....	30

TABLE OF CONTENTS

1. INTRODUCTION.....	9
2. LITERATURE REVIEW.....	11
3. THE MODEL.....	13
3.1 Social interaction environment incorporated too discrete consumer choices.....	13
3.2 Theoretical approach of a model of choice.....	15
4. METHODOLOGY AND DATA.....	15
4.1 Econometric Model	16
4.2 Data.....	18
5. RESULTS.....	23
5.1 Drug use, Brazilian students and socioeconomic profile: stylized facts.....	23
5.2 Econometric exercises.....	30
5.3 Discussion.....	38
6. FINAL REMARKS.....	40
REFERENCES.....	42
APPENDIX.....	47

1 - INTRODUCTION

Studies in economics of how consumer's choices depend not only on individual factors, but on factors external to them - those who make decisions - have recently been established in the theoretical and empirical literature (HSIEH and LIN 2021; MATSON, 2021). Within a wider universe of possibilities for analyzing preferences and choices, health-related behaviors are also research objects and how social interactions between agents contribute or not to outputs in this area, especially in adolescence/school socialization environment.

The issue of conformity in which individuals fit, perform social roles, imitate each other, producing the general phenomenon of homophilia does not escape the economic analysis of health (BERRY et al., 2018) where young people establish their relationships with other friends/companions and form their preferences for consumption by the most varied types of goods, especially the use of tobacco cigarettes.

The behavior of smoking cigarettes, which is the object of analysis of this research is inserted in the area called economics of social interactions (SI), which emphasizes how preferences and subsequent decision-making are influenced and formed by the context of the insertion of individuals in the most diverse social groups. The question of how smoking among adolescents is an activity possibly resulting from interactions between peers of friends and family is also in the specialized literature that treats the topic of Tobacco Smoking as both an individual and a collective health problem (HSIEH; VAN KIPPERLUIS, 2018).

More precisely in the Brazilian reality, the importance of studies on the social economy of health in line with theoretical and empirical findings in the specialized literature of the medical sciences is necessary, especially in dialogue with scientific trends in other national contexts (PIÑA et al., 2018). Brazilian stylized facts denote that smoking is associated with the development of various types of cancer, with 428 people dying each day due to direct dependence on nicotine, in addition to diseases caused by secondhand smoke (INCA, 2018). In the case of smoking among adolescents, Brazil has legislation that prohibits the sale of cigarettes to minors. However, statistics show that the age of smoking initiation is between 14 and 15 years old, below the age of majority (BASTOS, 2017) and defying the current norms of ordinary legislation. Despite the factors mentioned and the absolute prevalence of smokers not exceeding 10%, the consequences throughout life impact on economic productivity, on general

health indices, of those who use tobacco, which is in accordance with the number of cigarettes consumed daily (BRITO et al., 2021).

The need to investigate socioeconomic and demographic characteristics that make young people start using a legal drug has two reasons; the first is the discussion applied, at the Brazilian level, of strategies not necessarily in terms of public policies - private leisure, for example - to prevent individuals in puberty from starting tobacco consumption or ending this decision. The second reason is more specific to the SI literature, in which problems arising from the formation of expectations, preferences, social utility, contextual and endogenous effects can be detected by estimating the parameters of the empirical model in order to discuss the common elements in this type area of studies – such as, for example, the presence of a social multiplier effect (FLETCHER, 2010). Both motivations are linked to each other.

Given the aforementioned justifications, this article analyzes how tobacco consumption is influenced or not by social interactions. It is empirically investigated whether the smoking behavior of Brazilian adolescents in the period between 2012 and 2015 is dependent or not on the effects of social interaction in the classroom environment. We work with the decisions of young people between 11 and 19 years old regarding the consumption of industrialized tobacco cigarettes, in addition to other illicit substances, socioeconomic context of their families and institutional characteristics of schools, contained in the “*Pesquisa Nacional de Saúde do Escolar*” (PeNSE), of the respective years mentioned above.

The identification of whether the effects of social interaction explain changes in drug use in general is still relatively unexplored in the Brazilian academic literature. At the level of research in applied economics, Becker (2017) is one of the few studies that measured peer and contextual effects using statistical data from 9th grade adolescents at the Brazilian level, via the binary choice model of drug use. Internationally, contributions from Gaviria and Raphael (2001), Powell et al. (2005) and McVicar (2011) are classic studies, which also explore issues related to tobacco/drug use in school grades.

All the works mentioned above make use of theoretical approaches that focus only on whether an individual consumes cigarettes or not. The nature of the problem requires that a broader categorization be made, beyond the binary yes or no to tobacco consumption. Thus, the first contribution of this article is to expand the analysis, where a possible chemical dependency can be studied, in which the use of cigarettes is measured in polytomous categories, together

with a broader theoretical framework, which are the multinomial choice models of Brock and Durlauf (2005); the latter is almost absent in research in the area of medical sciences and health econometrics at the Brazilian level.

A second contribution is to applied literature, which is concerned with how to solve identification problems that arise when peer, contextual, and correlated effects cannot be distinguished from each other and which of them impact individual behavior; or whether it is from individual social interactions that aggregate social behaviors emerge. The solution to these problems and the analytical and quantitative tools used in this article are the use of control variables for individual characteristics of the students, fixed effects of year and region, institutional context of the schools and the estimation of a pseudo-panel with the lag in the time of the social contagion effect.

Thus, the individual smoking behavior of Brazilian adolescents can be better explained using these controls to circumvent possible biases that arise from self-selection of student pairs about the same classroom and school. Therefore, there is a dialogue, aligning this approach, that of Health Economics and Social Economics, with the study of young people aged 11 to 19, in an applied manner to Brazilian sample data.

In addition, it is verified whether there are effects of the complementary consumption of other drugs, the use of tobacco cigarettes converging with Brazilian research on the subject (Ferreira et al., 2019; Sanchez et al., 2015), mainly in the empirical model of consumption. Thus, there are contributions in methodological terms, bringing a topic widely studied outside Brazil. Issues related to monitoring individual consumption behavior over time, school performance and other activities that are exposed to social interactions are addressed in the applied economics literature (ROBALINO and MARCY, 2018; REYNOLDS, 2020).

The remainder of this research is divided into the following sections. Section 2 presents a literature review on how social interaction effects (contextual, peer and correlated) influence choices, focusing attention on drug and youth issues. Section 3 explains the theoretical model that supports the peer effects approach. Section 4 contains the methodology and data used. Section 5 presents the results and discussions and section 6 presents some final comments.

2 - LITERATURE REVIEW

The identification of which agents interrelate through social groups of the most diverse sizes and scope, changing their behaviors of choice, including, but going beyond the sphere of consumption of economic goods (drugs, marriage, criminality and others), forms the core of the so-called Social Economics (BECKER, 1974). These considerations open paths for the study within the economics of themes, previously restricted to sociological research, such as urban gang behavior, use of illicit drugs by minors via social influence, paired school performance, among other agendas.

That said, there are three theoretical concepts that guide and at the same time justify that attention to the social environment is important to be analyzed: peer effects (also synonymous for endogenous effect, neighborhood effect, social contagion, imitation), contextual and correlated effects.

The first effect briefly relates to how average or common behavior affects an individual's chances of behaving similarly within the social group; the second concept is related to how exogenous characteristics of the group affect the individual's behavior and finally, the third effect is a cross-relationship between the pair effect and the contextual one. Measuring both effects and verifying their magnitude in the Brazilian context is to converge with what has been done in the international scientific literature, as in Hsieh and Lin (2021). The respective effects can be grouped around a single definition, which are social externalities – social multiplier effects – which are studied in microeconomic theory (BOUNDREAUX and MAINERS, 2019). Social multiplier effects can be measured and have applicability in public drug policies, where there are spillover effects beyond the individual behavior that is intended to change, when a policy is applied, because there are peer effects (GLAESER, PRIEST and SCHEINKMAN, 2003).

Not least, detecting whether the individual is influenced by the social group or whether the aggregate behavior is explained by individual actions is also of paramount importance, which leads to the so-called Problem of Reflection (MANSKI, 1993). This problem has theoretical and applied ramifications, being extended to the study of tobacco consumption. The concurrency problem means that at the applied level, the three aforementioned effects may not be correctly distinguished and identified. That said, any debate applied in terms of public policies to tobacco use by young Brazilians must begin by overcoming the Problem of Reflection.

In the academic literature on smoking and peer social effects, attention has been paid to the type of social group that is the most relevant and significant to explain individual behavior. Various spheres such as school grade, school, friendship networks and influences from third parties, in addition to geographic homophiles, are modeled and identified if these social groups exert any weight on variations in individual consumption. It is precisely in the primary environment of the classroom, where social interactions take place, more directly and with constant impact during schooling, that smoking behavior has been extensively researched (AHO et al., 2019).

Adding to the questions of how individual's socialization processes at school and in the family impact their decisions per se, we have that the recurrent consumption of tobacco, nicotine and other complementary drugs is a problem in terms of health (in the most varied dimensions), in the capacity of young people to perform their potential in the labor market and in risk behaviors (COHEN; PRINSTEIN, 2006).

The dialogue that this literature promotes is also in connection with other researches that raise the issue of which sphere or groups of social references are the most adequate to explain the behavior of subjects; see (AXTELL et al., 2006). This is explained by the reason that certain interaction structures are more influential, such as the family or group of friends in the same classroom, than the participation of a young person in a group in a distant neighborhood. A second point derived from this is how agents make decisions, the latter being dependent on how individuals form their expectations, given that they are inserted in a social group. In this specific case in question, both the school class and the family are the reference groups for the formation of action expectations. Last but not least, detecting whether the individual is influenced by the social group or whether the aggregate behavior is explained by individual actions is also of paramount importance, which leads to so-called simultaneity problems (MANSKI, 1993).

International studies that focus on cross-sectional dimensions of social interactions, such as network and genetic pair effects, using cohort data and/or nonlinear response models with instrumental variables are markers of the state-of-the-art of current research on peer effects (BRAMOULLÉ; DJEBARRI; FORTIN 2020; SOTOUDEH, MURRAN HARRIS; CONLEY 2019). These researches show how a different social reference group can change individual

behavior, going beyond local contexts such as classrooms, schools and even genetic-environmental conditions (TANKSLEY, 2019).

Also recurrent in the state-of-the-art literature on social interactions, research that measures how genetic factors at the individual level interrelate and correlate with environmental factors can show the peculiarities of behavior in relation to health in general. This is because, as shown in Brunello (2019), Cawley (2017), Domingue et al. (2018), Griffin et al. (2015), Li and Guo (2021) and Sotoudeh (2019), individual genetic endowments for a given phenotype are moderated by environmental factors, when surveyed over time and by the type of characteristics of the pairs.

3 - THE MODEL

In this section, the theoretical model underlying consumer choice decisions is explained, taking into account aspects of how the relationship between decision makers is intertwined.

The theoretical model chosen is not exhaustive regarding Social Economy and related areas, and the reason for the choice is due to the empirical implementation already established in other applied contexts and which derives from the theory and the object of study being considered a social interaction of non-market.

Intuitively, the choices of an agent who is faced with a set of individuals, also choosing certain courses of action, is influenced by the latter.

Formally, the discrete choice theoretical model with social interactions is presented in what follows in the next two subsections. In the first, the concept of social interaction is approached, in which the decision-making of individuals is directly dependent on it. Subsequently, both the environment of social interaction and the description of the properties that give cohesion to the choices that are the basis for the econometric model are formalized.

3.1 - Social interaction environment incorporated too discrete consumer choices

The study of social interactions, or similarly, non-pecuniary social externalities produced in the contact of individuals with each other is formalized from the work of Blume, Brock and Durlauf (2000, 2003, 2005).

In this sense, the externalities that matter in this analysis are the social ones that are generated from the relationships between agents who make decisions, and the latter produce

different courses of action when we look at an isolated individual, within a group, where we have interactions that are of any type. The notion of peer effects is precisely the close relationships that agents have with each other, with various types of scenarios for this, such as family, neighbors, neighborhood, co-workers and other social groups (Manski, 2000).

From this, initially consider an agent i that is within a social group g , where, for reasons of coherence, the same individual has some kind of social interaction with the others. In short, what we want to explain is the probability of discrete choice of each subject - where their likely choices, we will name w_i - within a possible set of behaviors, Ω_i . To give empirical content to what was exposed, it is assumed that the set of behaviors is the use or non-use of cigarettes (smoking in the last thirty days).

Since it is assumed that the choices in probabilistic terms of individuals i are affected by various influences, these are as follows:

- 1) a h_i vector of individual-level deterministic characteristics specified for each subject i .
- 2) a vector π_i of random characteristics at the individual level that is associated with each i .
- 3) the expected expectation μ_i^e and (w) of beliefs that individual i has in relation to other members of the social group.

The aforementioned influences are distinct elements that make up the argument of a payoff function (a function that depends on results) of individual choices.

Basically, each choice made by i individuals of a social reference group provides a respective utility, the latter to be maximized. Formally, individual choices represent the maximization of a payoff function (V), so that:

$$w_i = \arg \max_{\xi \in \Omega} V(\xi, h_i, \mu_i^e(w), \pi) \quad (1)$$

The decision-making agent is faced with the problem involving (1), which can be understood as a preference function (where the function arguments are summarized by V), where a set of choice constraints is added, which is Ω termed in the function and the expected beliefs, $\mu^e(w)$. The formation of beliefs, which is also subjective, is based on the interactions of an individual within the group, in terms of the expected expectation that an element has in relation to others. In this sense, the expected belief seen from the perspective of this type of

choice model must obey the so-called self-consistency, which says that the expectation must be rational, in the sense that each individual i in relation to group g expects in relation to the others.

The self-consistency property of the choice model is applied both to subjective beliefs, as well as to objective conditional probability $\mu(\omega | F_i)$. In general, a decision maker has objective information on how other agents choose their options, as well as the structure of choices made by others, in a rational and satisfactory way - a denser rationality in this sense (Brock and Durlauf, 2003). Therefore:

$$\mu_i(w) = \mu(w | h_i, \mu_i^e(w)) \quad (2)$$

In the equation, the right side represents objective conditional probability. The expected belief seen from the perspective of this type of choice model must obey the so-called self-consistency, which says that the expectation must be rational, in the sense that each individual i in relation to group g expects in relation to the others. The concept of rationality here is similar to the axioms of preference of the theory of consumer, adding the fact that we are working with preferences conditional to social contexts. Expectations are also rational, in order to constitute a Nash equilibrium, since the situation described under social interactions is a type of strategic interaction, which can be modeled within Game Theory (Arrow, 1994).

3.2 - Theoretical approach of a model of choice

Given that agents are faced with a problem of choosing certain types of actions that give them maximum utility, we formally have that:

1) the agents within a reference group are faced with a set of possibilities $\Omega_i = (0, 1, 2, \dots, L-1)$, where, for the purpose of empirical content, L is assumed to be a “consumption space” possible in terms of the number of days that tobacco was used in the last month. The objective here is to give the intuitive idea of restricting the courses of choices that are feasible to be made and available, for each individual i .

2) Given that each agent i obtains utility from each choice l , within the possible set Ω , one has to:

$$V_{i,l} = h_{i,l} + Jp_{i,l}^e + e_{i,l} \quad (3)$$

where the term “ h ” is the deterministic private utility of agent i , given a choice l ; the term “ Jp^e ” is the social utility, the term J being a weighting of how strong the social utility is, multiplied

by the subjective expectation that individual i has in relation to beliefs that are close to him, within the social group. The term “ e ” translates to private random utility, which is not observed in econometric implementations; details are in Anderson, De Palma and Thiesse (1992).

From the theory exposed above, it is derived in terms of empirical application, the contextual and endogenous effects, which come from the terms “ h ” and “ Jp^e ”. The term “ h ” in econometric works is composed of the individual factors X of each agent i – previously determined by the researcher – plus the so-called contextual effect $Y_{g(i)}$ which describes how indirect characteristics of members of a social group affect individual decisions. In turn, the term “ Jp^e ” is operationalized as the endogenous effect, in which it can be summarized, in the way in which the behaviors of other individuals impact upon the decision of an agent i . Also in regressions, specifying “ e ” is known as random error and associated with unobserved variables that affect the decision to choose.

From the above, maintaining the hypothesis that the errors are independent of X_i and $Y_{g(i)}$ for every individual i and choice i , the likelihood function for the set of omega choices is:

$$Z_l^{-1} \Pi [\Sigma \exp(\beta k_1 + \beta c_1 X + \beta d Y_{g(i)} + \beta J_1 p_{g(i),l}^e) 1(w = 1)] \quad (4)$$

where Z is the normalization of:

$$Z_1 \Pi [\Sigma \exp(\beta k_1 + \beta c_1 X + \beta d_{g(i)} + \beta J_1 p_{g(i),l})] \quad (5)$$

Thus, the econometric model is consistent with the approach to choices under the context of social interaction.

4- METHODOLOGY AND DATA

This section is divided into two parts. The first presents the empirical strategy to understand the relationship between peer effects and tobacco consumption. Next, the database used is displayed, as well as descriptive statistics.

4.1 - Econometric model

In order to empirically study the summed relationships of individual and pair effects on the dependent variable, a multinomial logistic regression is used, this model having desirable properties from a theoretical and statistical point of view.

The properties that are desirable and that the multinomial choice model meets, according to Brock and Durlauf (2003) are linked to 1) how to overcome the Reflection Problem (individual behavior is affected by the social group or vice versa) and 2) discern the contextual effect of the social interaction effect. The problem (1) mentioned leads to the endogeneity between the consumption of industrialized cigarettes and the pair effect.

Therefore, using a multinomial logit model it is possible to deal with the issue of endogeneity, and the non-linear relationship between the effect of social interaction and individual behavior, as shown in Brock and Durlauf (2000), as well as in (Manski, 2000)

In addition to the reason listed above, in the empirical literature on the measurement of peer effects, the use of the multinomial logit model is made, as in Aho et al., (2019); Alves et al., (2020); Nguyen et al., (2020).

Thus, the equation that describes the empirical model (benchmark) for the decision of individual i , in social group g , who uses tobacco in period t is:

$$Y_{igt} = \beta_{0t} + \beta_1 P_{igt} + e_{ig} \quad (6)$$

where Y is the dependent categorical variable that takes discrete values in the interval $[0,4]$. More specifically denoting choices an individual makes in choosing to smoke. While the regressor P is just our main independent variable of interest in relation to j students.

Although the use of logistic regression is supported by a sufficiency condition and empirical issues in the literature, there are also other sources that compromise the analysis of the results. That is, in addition to the endogeneity between Y and the social interaction effect, there are two other sources of bias. The first arises when common characteristics, rules and shared institutional environment among young people in a school are missing; this can lead to inflated or overestimated pair effects. Finally, there is the existence of group self-selection (endogenous peer selection), where, in the case of smoking, those responsible (parents) for adolescents choose schools based on similar or similar preferences and attitudes; In this sense, families more likely to allow their children to smoke, for example, may choose their schools, with flexible smoking policies, leading their children to have more likely social relationships with smokers, increasing their propensity to smoke (NIKAJ, 2017).

More specifically for the problems raised above, the empirical strategy used is the estimation of a series of models, where the individual behavior is regressed in relation to the effects of the pairs only, and in the other stages, variables of an individual nature, family context, rules institutional shared at school, consumption of complementary goods, year effect and Brazilian regions are added. Following this approach, the institutional effect of the school is controlled, which leads to distinguishing the correlation effect from peer effects, which is in the theoretical literature of social interaction. Year and region variables are also added, and this approach is similar in studies such as Lundborg (2006)

The specifications below denote the models that are estimated via the pseudo-likelihood log:

$$Y_{igt} = \beta_{0t} + \Sigma\beta_1 I_{igt} + \beta_2 P_{igt} + e_{ig} \quad (7)$$

$$Y_{igt} = \beta_{0t} + \Sigma\beta_1 I_{igt} + \beta_2 P_{igt} + \Sigma\beta_3 F_{igt} + e_{ig} \quad (8)$$

$$Y_{igt} = \beta_{0t} + \Sigma\beta_1 I_{igt} + \beta_2 P_{igt} + \Sigma\beta_3 F_{igt} + \Sigma\beta_4 S_{igt} + e_{ig} \quad (9)$$

$$Y_{igt} = \beta_{0t} + \Sigma\beta_1 I_{igt} + \beta_2 P_{igt} + \Sigma\beta_3 F_{igt} + \Sigma\beta_4 S_{igt} + \beta_5 C_{igt} + e_{ig}$$

$$Y_{igt} = \beta_{0t} + \Sigma\beta_1 I_{igt} + \beta_2 P_{igt} + \Sigma\beta_3 F_{igt} + \Sigma\beta_4 S_{igt} + \beta_5 C_{igt} + \beta_6 T_{igt} + \beta_7 R_{igt} + e_{ig}$$

From equations 7 to 10, the variables I, F, S and C indicate, respectively, the individual characteristics of adolescents and their demographic aspects, socioeconomic background of the families of the *i* students, a control variable at school level and the presence or absence of consumption of other complementary drugs. From the aforementioned specifications, our main estimate of interest is equation 11, in which we would like to verify how the additions of other variables, such as time and place, can affect smoking behavior. These are denoted by T which is a dummy for year and R categorical values of Major Brazilian Regions.

Finally, and trying to circumvent the first source of endogeneity, which is between social interaction and smoking behavior - whose use of a logistic model is not a necessary and sufficient condition to dissipate the Reflection Problem - a pseudo-panel is estimated. This methodology is justified, since the sample database of this research is not longitudinal and assumes whether a temporal structure of lagged social interaction is present in the behavior of young Brazilians. This approach is similar to other studies, where the econometrics of genuine

panel data is used as time-lagged effects, in the most diverse types of pair effects (CLARK and LOHÉAC, 2007; MANSKI, 2000; BRAMOULLÉ; DJEBARRI; FORTIN 2020).

Taking this into account, we applied the time lagged effects approach for independent cross-section, where cohorts (c) of year of birth grouped with the sex of the individuals are performed, in addition to the average of the pair effects lagged ($P_{c,t-1}$) for cohort c of the year 2012.

In this way, the following model is estimated:

$$Y_{ict} = \beta P_{ic,t-1} + \Sigma \beta X_{ict} + \alpha_c + v_i + u_{ict}, \quad c=1, \dots, C; \quad (12)$$

where in the equation above, α_c represents dummies for the cohorts, v_i is a random error per hypothesis. The analysis is done using the year 2015 as the current base. Others variables (X_{ict}) summarized by the summation are the same as in equations 7 to 11.

4.2 - Data

For this article, the *Pesquisa Nacional de Saúde do Escolar* (PeNSE) for the years 2012 and 2015 was used as a database. This survey, which is produced by the *Instituto Brasileiro de Geografia e Estatística*, contains information on health protection and risk factors of Brazilian adolescents, complemented with data on individual schooling of students and parents, inclusion of young people in the labor market, socioeconomic characteristics of the school environment and surroundings, behaviors related to the use of various types of drugs, mental and physical health, hygiene and psychological well-being.

The final sample that is used contains information from 204145 (sum of the years 2012 and 2015) students, and 211176 individuals were originally surveyed. This is due to the fact that some adjustments were made, removing students who did not answer questions related to the use and non-use of drugs from the original sample design. Data were collected from 5882 schools and, in order to pile up samples from different years, it was necessary to use information at the level of large Brazilian regions (“Centro-Oeste”, “Norte”, “Nordeste”, “Sul” and “Sudeste”). Below are detailed descriptions of all the variables that are used in the econometric model, which comes from PeNSE 2012 and 2015:

- Dependent variable (Y). To measure tobacco consumption among adolescents, the question "In the last 30 days, on how many days did you smoke cigarettes?" The answers

to this question are grouped into 7 categories, with the items “Jump in the questionnaire” and “Not informed” are added to the answers of the variation in the number of days a young person smoked. For the case of the response variable, a recoding strategy is used, in which the original numerical values are grouped into four new categories. In this direction, we have: I never smoked-0 (1) and I jump in the questionnaire-0 (1); I haven't smoked in the last 30 days-1 (1 and 2); Occasional smoker-2 (3,4,5 and 2,3,4); Regular smoker-3 (6,7,8; 5,6,7). The numbers in parentheses indicate the original value in the questionnaire, while those that accompany the variable name are the discrete values in the multinomial logit model.

The variable recoding procedure is used in research on smoking in general; examples are Banzer et al., (2017); Scalici and Schulz (2016).

- Main independent variable (P). For this case, which are the effects of peers at the class level, the measurement is made taking into account a formula, which calculates the percentage (share) of students who smoke in a room, except for individual i : $P_{ics} = 1/n_{cs} - 1 (\sum_{j \neq i} S_{jcs})$. S is a binary variable, where 0 is assigned to non-smokers and 1 to smokers.
- This adopted strategy is used in the most varied types of research (ALEXANDER et al., 2001; GREEN et al., 2007; VAN DE BRAND et al., 2019), where the social reference group that influences the individual ranges from the class from school to co-workers.
- Other independent variables (I, S, C, T and R). The individual effects take into account information collected from questions about age, sex, participation in the labor market, practice of sports activities outside school, of the respective adolescents. The institutional characteristic of the school represents whether the unit implements anti-smoking actions (if it has policies and norms that prohibit the consumption of tobacco in its physical facilities). The consumption of complementary drugs to cigarettes is defined if the young person uses any other tobacco product, alcoholic beverage or illicit drugs (marijuana, crack, etc.). In addition, fixed weather effects (2012 and 2015) and for large geographic regions are added: “Centro-Oeste”, “Norte”, “Nordeste”, “Sul” e “Sudeste”.

The definition of the answers and the measurement of the aforementioned variables follows the original questionnaire, and in the case of the use of complementary drugs, a binary category is

constructed, and if an adolescent consumes any other product, except tobacco cigarettes, the value is assigned of 1, and otherwise 0. Below are summary statistics for all variables in the article.

Table 1 – Variables and responses used in the research

Variables	Proportion of the answer of students	Variables	Proportion of the answer of students
Y		Mother's education	
Non smoking	0.816	Did not study	0.209
Have not smoked in the last 30 days	0.130	Elementary school incomplete	0.062
Ocassional smoker	0.040	Complete primary education	0.209
Regular smoker	0.012	Incomplete high school	0.065
Peer effects		Complete high school	0.067
Age		Incomplete college	0.188
11 to 13 years	0.188	Complete college	0.053
14 years	0.482	I do not know	0.142
15 to 19 years	0.329	Parent attention	
Sex		None	0.189
Men	0.479	Sometimes	0.377
Women	0.520	Always and often	0.433
Participation in labour market		Parents smoke	
No	0.872	No	0.746
Yes	0.127	Yes	0.253
Others drugs		Anti-smoking policy	
No	0.930	Not anti-smoking policy	0.059
Yes	0.069	Yes anti-smoking policy	0.453
Physical activity		Not information	0.486
None	0.330	Year	
Few	0.411	2012	
1 to 3 days a week	0.254	2015	
4 to 7 days a week	0.003	Region	
		Centro-Oeste	0.228
		Norte	0.331
		Nordeste	0.183
		Sul	0.121
		Sudeste	0.135

Source: Elaborated by the author.

5- RESULTS

5.1- Drug use, Brazilian students and socioeconomic profile: stylized facts.

Questions related to smoking behavior, as well as the description of information in the sample of students from the National School Health Survey, are important for us to have a first overview of Brazilian students as an initial step towards statistical inference.

Such behavior in the practical life of young people develops in their constant interactions within classrooms, schools and surroundings. The act of experimenting, using cigarettes sporadically or constantly, can produce different emotional and social reactions and a sense of belonging to some group. It can be associated with this practical description that adolescents both react positively to the use of tobacco, as well as a sign that when they smoke, they become or enter the "adult" phase.

The first statistic of interest refers to the percentages of each response to smoking behavior, given in Table 1 below.

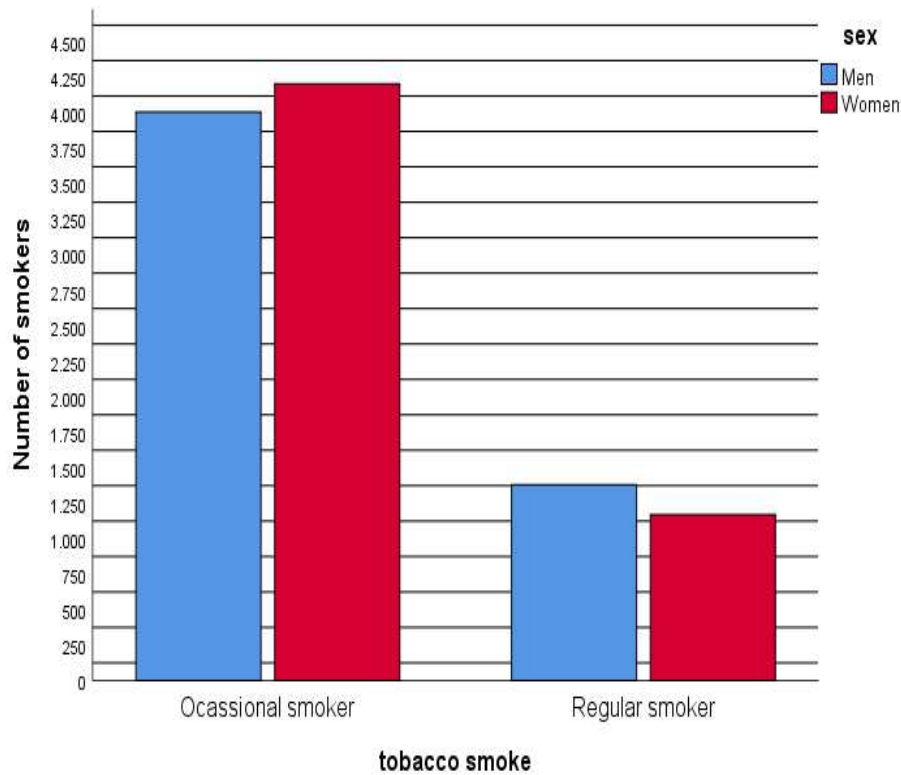
Table 2- Tobacco consumption (industrialized cigarettes) in the entire sample.

Tobacco smoke		
	Smokers	Percent
Non smoking	166702	81,7
Have not smoked in the last 30 days	26667	13,1
Occasional smoker	8223	4,0
Regular smoker	2553	1,3
Total	204145	100,0

Source: Elaborated by the author.

Looking specifically at those who smoke considering the sex of students and the answers given by them, there is an overview of the sample as shown in Graph 1.

Graph 1- Tobacco smoke due a sex

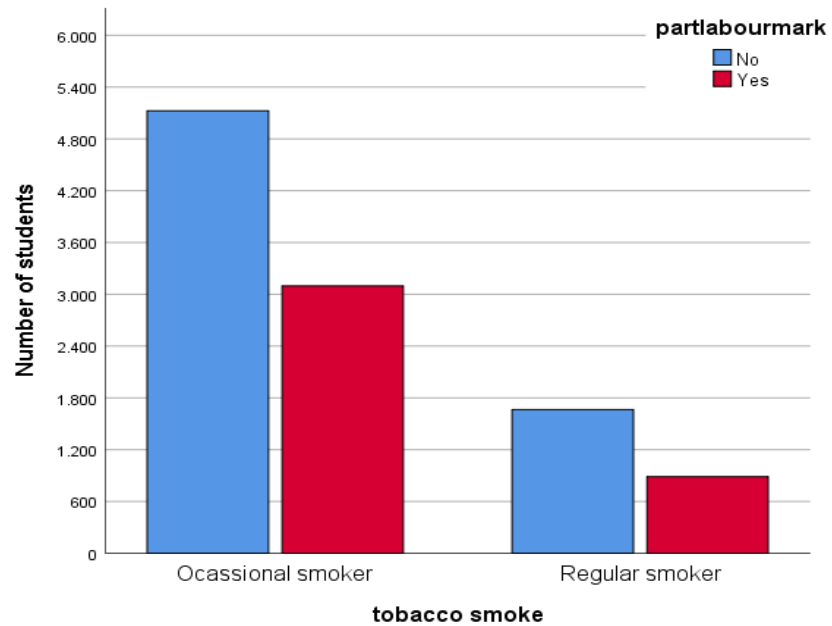


Source: Elaborated by the author

From Graphs 1, cigarette consumption is only a small portion of the sample universe, with occasional and regular smokers. One must take into account the share of students who smoke alone in the classes, which mischaracterizes the purpose of the effects of pairs - at least two students smoking. In addition to the fact that the number of individuals who did not report whether they used cigarettes may be related not only to the mere negative statement of the question, but to some hypothetical type of embarrassment, given the age and context of answering if they are a cigarette user.

Regarding the issue of student participation and how much this corresponds to being an occasional and regular smoker, Graph 2 shows this relationship, being of economic importance, because in econometric exercises, the proxy for income is precisely the insertion of the labor market of teenagers.

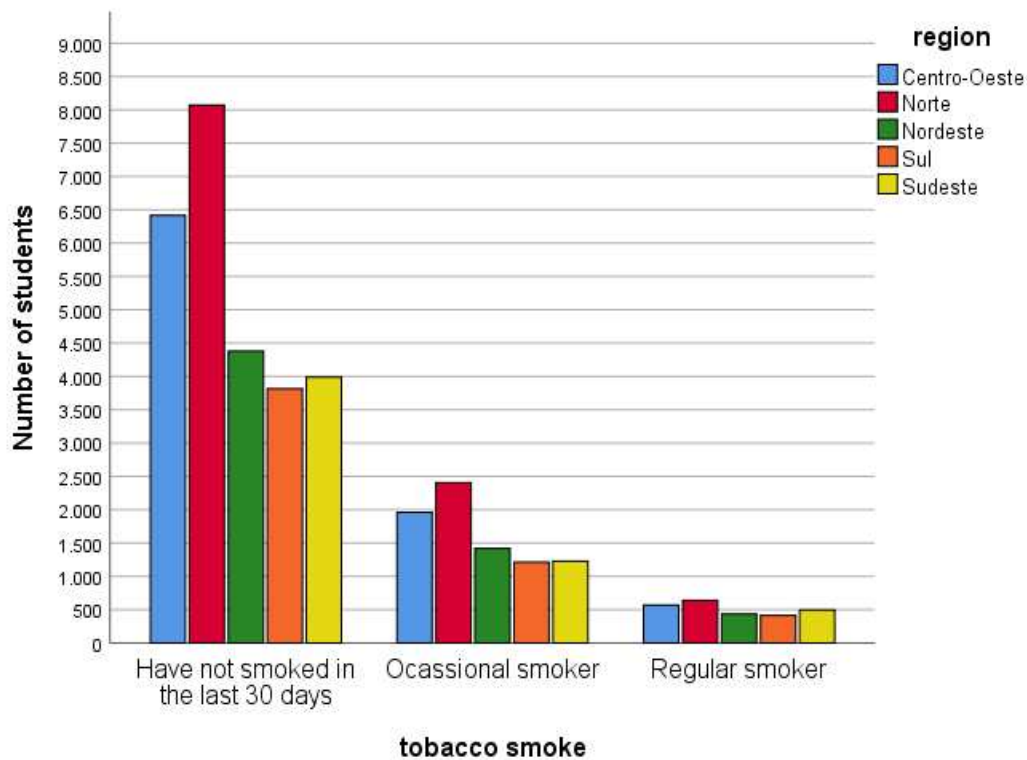
Graph 2 - Occasional and regular smoker due to participation in the labor market.



Source: Elaborated by the author based on 2012 and 2015 PeNSE

Finally, with information on the two main response categories, occasional and regular smokers, the prevalence is higher in the “Norte” region, for the first category, while the distribution for the second category is more uniform for all major Brazilian regions. Just with descriptive questions, the incidence of regular smokers may not be related to a student being from a certain region of Brazil. But any kind of inference can only be made through econometric modeling.

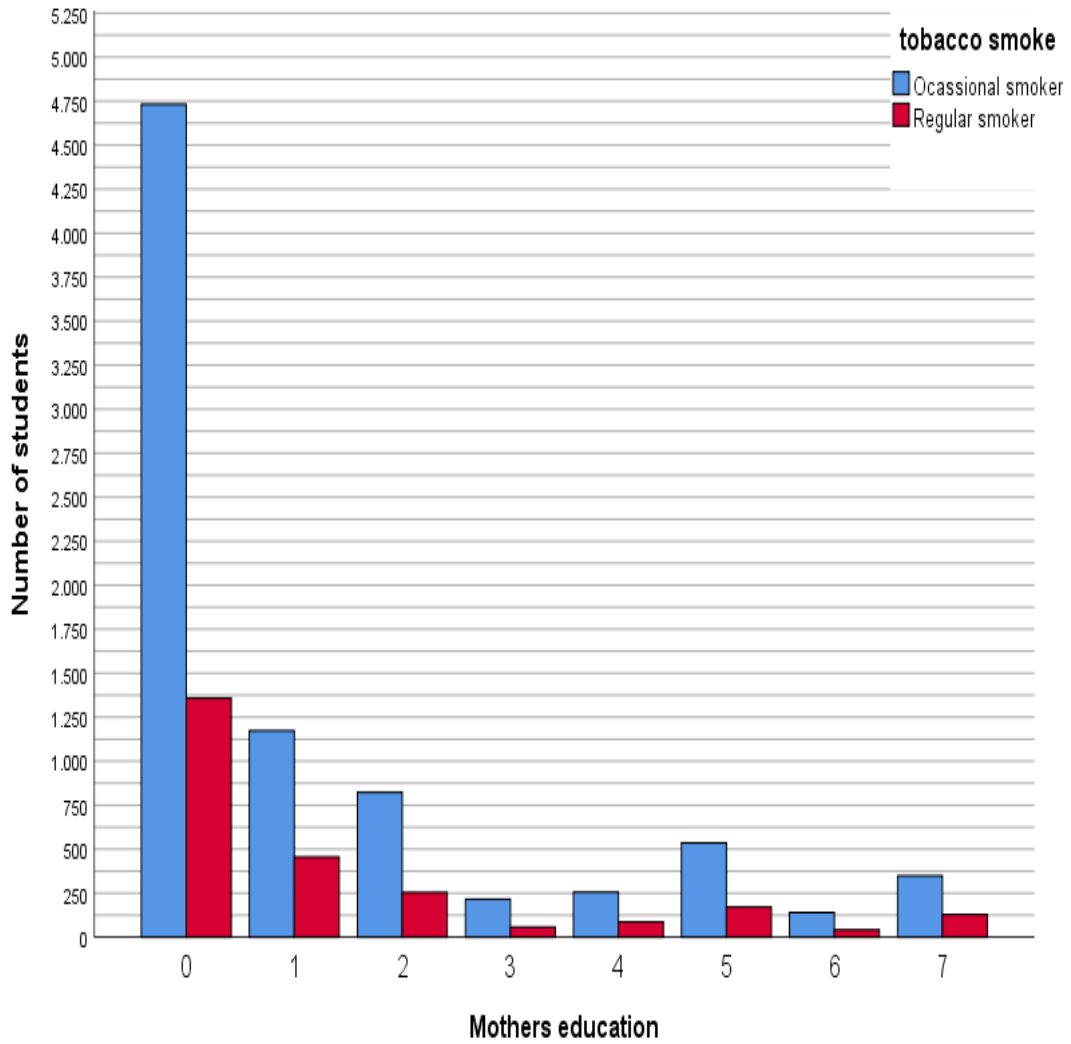
Graph 3 - Occasional and regular smokers in large Brazilian regions.



Source: Elaborated by the author based on 2012 and 2015 PeNSE

From Graph 4, there is an important stylized fact which is how much the occasional and regular smokers are distributed due to the education of the respective students' mothers. It is noteworthy that in the first levels of education there is a high incidence for the two response variables, but there is no clear pattern decreasing along the spectrum of the level of education of the teenagers' mothers.

Graph 4- Occasional and regular smokers due mother’s education



0 - Did not study; 1-Elementary school incomplete; 2-Complete primary education; 3-Incomplete high school; 4-Complete high school; 5-Incomplete college; 6-Complete college; 7- I do not know

Source: Elaborated by the author based on 2012 and 2015 PeNSE.

The relationship between the labor market and the consumption of other drugs is an indication that access to the labor market, even in adolescence, does not denote, at a descriptive level, that being working increases the consumption of complementary drugs. An individual who indicates that he uses another drug and has never worked is more numerically present in our sample than a student who works - 22,699 drug users who do not work versus 3,422 who use other drugs and at least work.

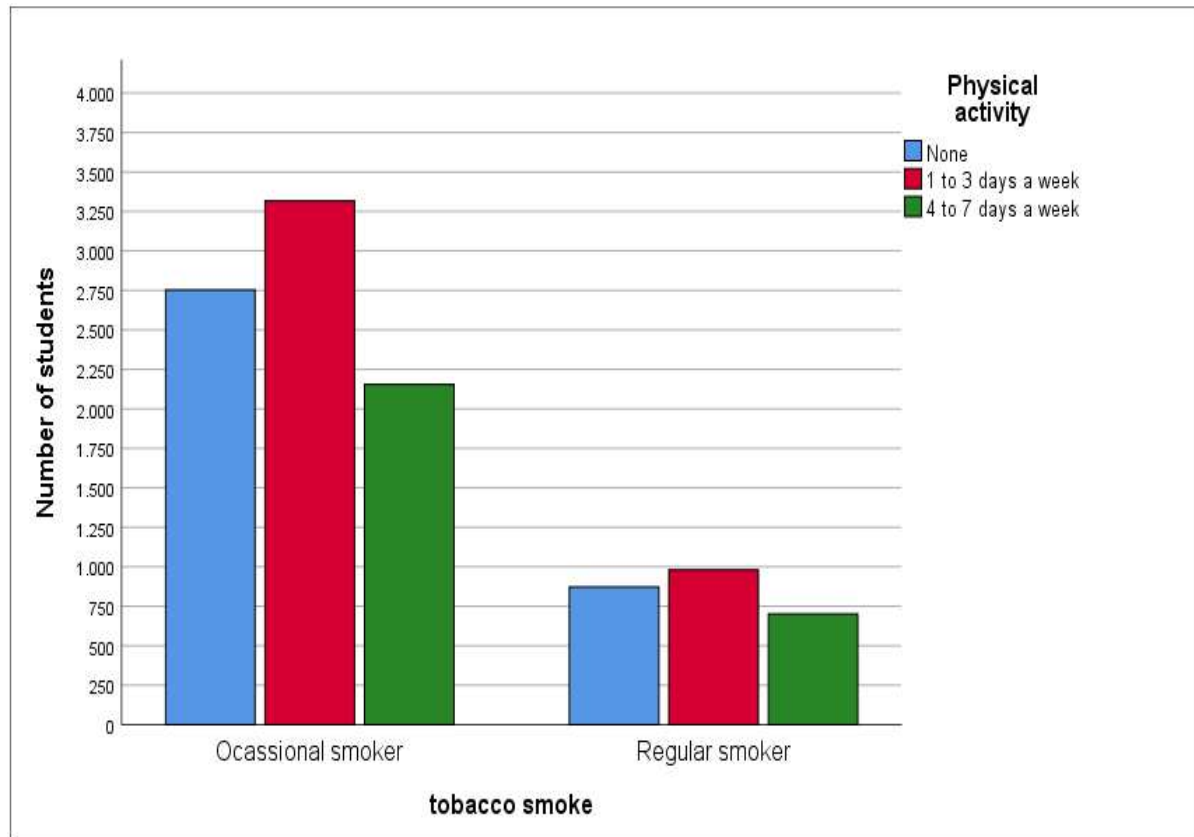
Table 3- Drugs versus participation in labour market

	partlabourmarket		Other drugs	
	No	Yes	No	Yes
No	167,181	10,843	10,843	178,024
Yes	22,699	3,422	3,422	26,121
Total	189,880	14,265	14,265	204,145

Source: Elaborated by the author based on 2012 and 2015 PeNSE.

The description of how the practice of physical activity outside of school related to tobacco use provides information about the profile of the students in the sample, especially the little variation of those who are regular smokers. Another fact is that for not practicing physical activity, the incidence of occasional smokers is lower than for other categories.

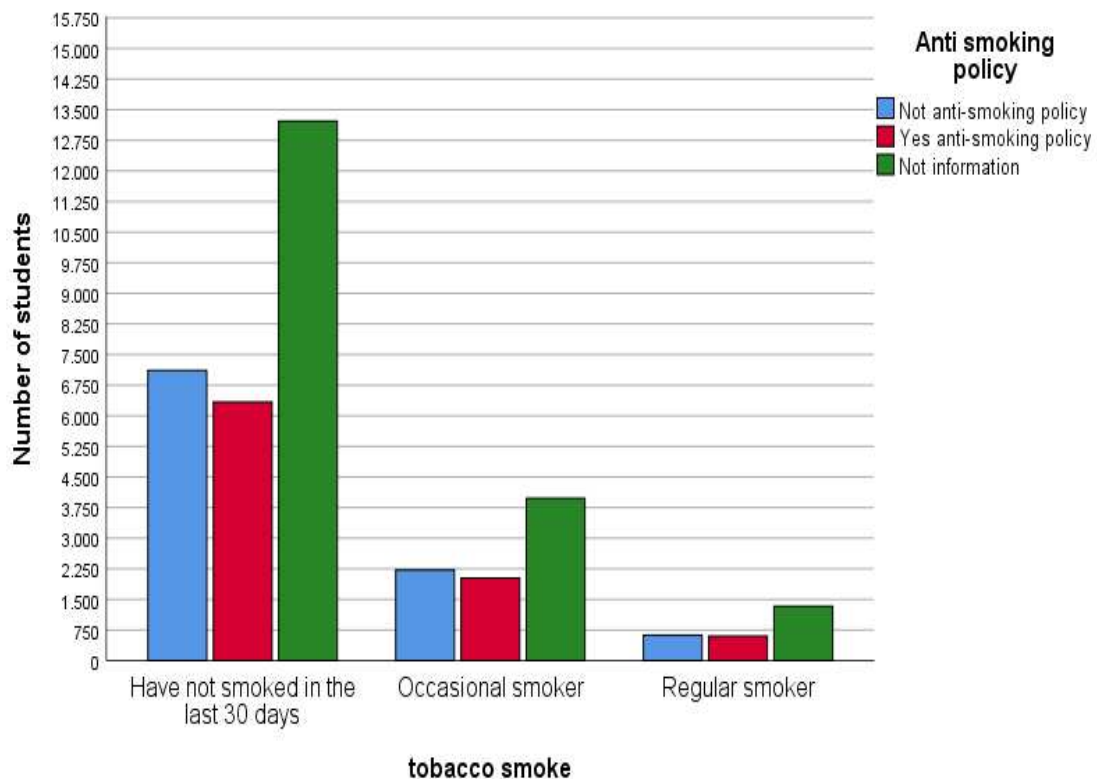
Graph 5- Physical activity due occasional and regular smokers



Source: Elaborated by the author based on 2012 and 2015 PeNSE.

Last but not least, the numerical incidence of those who have not smoked in the last 30 days or are occasional and regular smokers due to the smoke-free policy is shown in Graph 6.

Graph 6 – Incidence of smoking due to ban policies in schools and lack of information on this



Source: Elaborated by the author.

The prevalence of smoking in both categories of the horizontal axis is only partially linked to the existence of anti-smoking policies in schools, as, in the absence of information on this measure, it is not clear how important it is or not, in a descriptive way. as shown in the figure above.

5.2- Econometric results

Initially, the first point to be highlighted is how much the standard model of our analysis predicts or not different levels of choice in tobacco consumption, given the independent variables listed in the data section. By standard model (Model I), we mean the empirical form, considering that the pair effect is directly explanatory of smoking choices, that is, only with the individual effects of students.

Before performing the actual estimation of Model I, the verification of multicollinearity in the independent variables is made, and in this case, the Variance Infringement Factor (VIF)

for no variable is greater than the limit of 10 and significance less than 0.001. Therefore, the model does not present the aforementioned problem.

In general, Model 1, which has the characteristic of being a benchmark (multinomial logistic model) for the study of peer effects, is presented in Table 4. Our main explanatory variable is measured in terms of the proportion of student pairs who smoke plus the individual effects of students.

As explained in the introduction and in the methodology of this research, the estimation of several models, where possible sources of bias are corrected, is carried out. Issues related to self-selection of student pairs as well as school factors are inserted into the benchmark model used. As a first model, only pair effects are regressed against the categorical response variable.

The estimate when only the main regressor is in equation 6 is that all relative risk ratios decrease slightly, but still increase by a factor of 1,847, 3,444, and 3,213 for the 3 response categories, when compared to non-smokers.

Table 4 – Model I (benchmark) - only with the effects of pairs with regressors (Equation 6)

Variables (Omitted values are the bases for comparison; Non-smokers represent the base response category)	Have not smoked in the last 30 days		Occasional smoker		Regular Smoker	
	Relative ratio risk	Std. error	Relative ratio risk	Std. error	Relative ratio risk	Std. error
Peer effects	1.847** *	0.2773	3.444***	0.8966	3.213***	1.496

P-value: *p < 0,10; **p < 0,05; ***p < 0,01. 1- Robust standard errors. N=204145

Source: Elaborated by the author

Based on the main results, the 1-unit increase in the proportion of students (peer effects) causes the relative risk of being in the group of students who smoke occasionally to increase by a factor of 3.576, while the relative risk for regular smokers is 3,367. The comparison category (base) is in relation to those who have never smoked.

Regarding the other variables, it is noteworthy that an adolescent between 15 and 19 years of age has a relative risk of being an occasional and regular smoker, increased by a factor of 1,749 and 2,179, respectively, when compared to those who have never smoked.

Table 5 – Model (II) only with the effects of pairs with regressors Individual and peer effects that explain individual behavior (Equation 7)

Variables (Omitted values are the bases for comparison; Non-smokers represent the base response category)	Have not smoked in the last 30 days		Occasional smoker		Regular Smoker	
	Relative ratio risk	Std. error	Relative ratio risk	Std. error	Relative ratio risk	Std. error
Peer effects	1.888***	0.2841	3.576***	0.9336	3.367***	1.572
Age						
14 years old	1.194***	0.0229	1.167***	0.0398	1.169	0.0756
15 to 19 years old	1.523***	0.0305	1.749***	0.0607	2.179***	0.1383
Sex						
Women	1.009	0.0138	1.044	0.0245	0.878***	0.0370
Participation in labour market	1.207***	0.0231	1.446***	.0441	1.811***	0.0889
Physical activity						
Few	0.962	0.0150	0.988	0.0265	0.897**	0.0429
1 to 3 days a week	0.999	0.0177	1.028	0.0313	0.986	0.0524
4 to 7 days a week	0.609***	0.0844	1.354**	0.2256	1.174	0.3467

P-value: *p < 0,10; **p <0,05; ***p <0,01. 1- Robust standard errors. N=204145

Source: Elaborated by the author

Changes in pair effects (magnitude) with the inclusion of family variables are not as visible when family variables are added. Highlighting the relative risk of being an occasional and regular smoker, given that parents provide psychological support to their children, I reduced it to the two answers in the table, which is in line with research on social norms within families, as in Scalici and Schulz (2017).

By adding the variables that denote the background of the young people's families, the results are seen as follows below.

Table 6 - Model III - Background of each individual's family and peer effects to explain smoking (Equation 8)

Variables (Omitted values are the bases for comparison; Non-smokers represent the base response category)	Have not smoked in the last 30 days		Occasional smoker		Regular Smoker	
	Relative ratio risk	Std. error	Relative ratio risk	Std. error	Relative ratio risk	Std. error
Peer effects	1.873***	0.2825	3.510***	0.9176	3.233***	1.5092
Age						
14 years old	1.178***	0.0226	1.145***	0.0391	1.165**	0.0755
15 to 19 years old	1.470***	0.0301	1.664***	0.0590	2.148***	0.1389
Sex						
Women	0.987	0.0135	1.016***	0.0239	.854***	0.0361
Participatin in labour market	1.186***	0.0228	1.416***	0.0432	1.772***	0.0870
Physical activity						
Few	0.978	0.0153	1.012***	0.0273	.912**	0.0438
1 to 3 days a week	1.023	0.0183	1.063***	0.0325	1.001	0.0538
4 to 7 days a week	0.590***	.0819	1.291***	0.2167	1.126	0.3338
Mothers education						
Elementary school inc.	0.981	0.0297	0.978	0.0482	1.085	0.0936
Complete primary edu.	1.081***	0.0220	0.980	0.0338	1.121**	0.069
Incomplete high school	1.036	0.0306	0.978	0.0491	.897	0.0867
Complete high school	1.071**	0.0312	1.075	0.0524	1.189**	0.1049
Incomplete college	1.063***	0.0225	0.964	0.0353	1.184***	0.0779
Complete college	1.143***	0.0363	1.098	0.0596	1.213**	0.1201
I do not know	0.999	0.0238	0.965	0.0396	1.375***	0.0974
Parent attention						
Sometimes	0.904***	0.0161	0.865***	0.025	0.760***	0.0394
Always and often	0.723***	0.0129	0.655***	0.0199	0.631***	0.0329
Parents smoke	1.317***	0.0194	1.496***	0.0365	1.576***	0.0672

P-value: *p < 0,10; **p <0,05; ***p <0,01. 1- Robust standard errors. N=204145

Source: Elaborated by the author

When the control variable for schools is added (correlation effect), the results are that the relative risks are not statistically significant (at the 1%, 5% and 10% levels), except for regular cigarette consumption (existence of anti-smoking policy). It is important to note that there is not much information available to schools regarding other explicit or indirect norms, as shown

in Graph 6. The use of an institutional variable to control the peer effect on smoking is an innovation when compared to the work of Gaviria and Raphael (2001). Below, in Table 7, the estimated coefficients for the presence or absence of a smoke-free policy are summarized.

Table 7 – Model IV - Addition of an institutional variable along with peer effects that explain cigarette use (Equation 9)

Variables (Omitted values are the bases for comparison; Non-smokers represent the base response category)	Have not smoked in the last 30 days		Occasional smoker		Regular Smoker	
	Relative ratio risk	Std. error ¹	Relative ratio risk	Std. error ¹	Relative ratio risk	Std. error ¹
Peer effects	1.884***	0.2843	3.500***	0.9141	3.317**	15.507
Age						
14 years old	1.181***	.0228***	1.144***	0.0392	1.178**	0.0767
15 to 19 years old	1.473***	.0303***	1.662***	0.0593	2.168***	0.1417
Sex						
Women	0.987	0.0135	1.016	0.0239	.854***	0.0361
Participation in labour market	1.186	0.0228	1.416	0.0432	1.773***	0.0871
Physical activity						
Few	0.978	0.015	1.013	0.0273	0.909	0.0437
1 to 3 days a week	1.022	0.018	1.063***	0.0326	0.997*	0.0536
4 to 7 days a week	0.590***	0.081	1.291	0.2167	1.125	0.3336
Mothers education						
Elementary school inc.	.9763545	0.029	0.980	0.0483	1.063	0.0918
Complete primary edu.	1.075***	0.021	0.982	0.0341	1.096	0.0689
Incomplete high school	1.031819	0.030	0.980	0.0493	0.880	0.0853
Complete high school	1.066**	0.0311	1.078	0.0526	1.166*	0.1031
Incomplete college	1.058***	0.0225	0.9657	0.0354	1.163**	0.0765
Complete college	1.140***	0.0362	1.100*	0.0596	1.197*	0.1184
I do not know	0.999	0.0238	0.965	0.0396	1.375***	0.0974
Parent attention						
Sometimes	0.905***	0.0161	0.865***	.02571	0.764***	0.0395
Always and often	0.723***	0.0125	0.655***	0.0199	0.632***	0.0329
Parents who smoke	1.316***	0.0193	1.497***	0.0364	1.570***	0.0667
No anti-smoking policy	1.034***	0.0138	0.987	0.0225	1.145***	0.0463

P-value: *p < 0,10; **p < 0,05; ***p < 0,01. 1- Robust standard errors. N=204145

Source: Elaborated by the author

For the inclusion of the regressor of the consumption of complementary drugs (other drugs), the parameter estimates for the pair effect, when analyzing the relative risk of being in the group of regular smokers, presents statistical significance at the level of 5%.

Table 8 – Model V - Presence of consumption of complementary drugs (other drugs) along with other variables to explain smoking (Equation 10)

Variables (Omitted values are the bases for comparison; Non-smokers represent the base response category)	Have not smoked in the last 30 days		Occasional smoker		Regular Smoker	
	Relative ratio risk	Std. error¹	Relative ratio risk	Std. error¹	Relative ratio risk	Std. error¹
Peer effects	1.796***	.27258	3.124***	0.8249	2.664**	1.262955
Age						
14 years old	1.156***	0.022	1.091**	0.0375	1.068***	0.0699
15 to 19 years old	1.359***	0.028	1.405***	0.0508	1.565***	0.1032
Sex						
Women	0.994	0.013	1.032***	0.0247	0.884***	0.0380
Participatin in labour market	1.125***	0.021	1.272***	0.0397	1.465***	0.0734
Physical activity						
Few	0.982	0.015	1.021	0.0278	0.922***	0.0447
1 to 3 days a week	1.012	0.0182	1.039	0.032	0.949***	0.0517
4 to 7 days a week	0.622***	0.0865	1.448**	0.2487	1.404***	0.4213
Mothers education						
Elementary school inc.	0.996	0.030	1.026	0.0510	1.158*	0.1003
Complete primary edu.	1.076***	0.0220	0.984	0.0345	1.102	0.0698
Incomplete high school	1.024	0.0304	0.965	0.0489	0.857	0.0837
Complete high school	1.051*	0.0308	1.047	0.0515	1.107	0.0985
Incomplete college	1.037	0.0222	.922**	0.0340	1.062	0.0703
Complete college	1.099***	0.0352	1.016	0.0553	1.025	0.1018
I do not know	0.975	0.0234	0.910**	0.0377	1.221***	0.0872
Parent attention						
Sometimes	0.930***	0.016	0.915***	0.0275	0.841***	0.0439
Always and often	0.766***	0.013	0.738***	0.0228	0.788***	0.041
Parents smoke	1.270***	0.018	1.390***	0.0342	1.366***	0.0588
Anti-smoking policy						
No anti-smoking policy	0.934**	0.027	0.960	0.0490	0.913	0.0856
Yes anti-smoking policy	1.006	0.029	1.020	0.0516	1.208***	0.1119
Others drugs	2.666***	0.059	4.813***	0.1491	9.917***	0.4641

P-value: *p < 0,10; **p < 0,05; ***p < 0,01. 1- Robust standard errors. N=204145

Source: Elaborated by the author

The inclusion of this variable is an innovation in studies on peer effects, as in the latter, the consumption of any other type of drug is usually regressed, against the effects of social interaction. Details are in Table 8.

When adding the variables that capture the year and geographic effect (Major Brazilian Regions), the results are in most cases statistically significant for the dependent variable, as shown in the following table:

Table 9 – Model¹ VI - controlling for the effect of year and Major Brazilian Regions. Equation 11

Variables (Omitted values are the bases for comparison; Non-smokers represent the base response category)	Have not smoked in the last 30 days		Occasional smoker		Regular Smoker	
	Relative ratio risk	Std. error ²	Relative ratio risk	Std. error ²	Relative ratio risk	Std. error ²
Peer effects	1.487**	0.2256	2.501***	0.6580	2.064***	0.9616
Year	1.757272	0.1123	2.345***	0.2754	1.606**	0.3415
Region						
Norte	0.833***	0.0153	0.817***	0.0261	0.763***	0.0451
Nordeste	0.797***	0.0171	0.823***	0.0300	0.829***	0.0546
Sul	1.084***	0.0248	1.052***	0.0407	1.119*	0.0742
Sudeste	1.022	0.0229	0.990***	0.0379	1.318***	0.0836

P-value: *p < 0,10; **p < 0,05; ***p < 0,01. ¹Full results for this model are found in Appendix II

2- Robust standard errors. N=204145

Source: Elaborated by the author

Finally, the last model with the estimated parameters is one period lagged, as shown below:

Table 10 – Model¹ VII - Pseudo-panel with lagged mean social interaction effect with control for the others variables. Equation 12

Variables (Omitted values are the bases for comparison; Non-smokers represent the base response category)	Have not smoked in the last 30 days		Occasional smoker		Regular Smoker	
	Relative ratio risk	Std. error ²	Relative ratio risk	Std. error ²	Relative ratio risk	Std. error ²
Peer effects lagged (2012)	0.725***	0.0633	0.614***	0.0950	0.493***	0.1273
Cohort³ (Birth x sex)						
1993 (2)	0.849	0.1982	1.019	0.3627	0.705	0.3089
1994 (1)	0.994	0.2061	1.315	0.4216	0.521*	0.2058
1994 (2)	1.304	0.2793	1.856***	0.6136	0.595	0.2763
1995 (1)	1.268	0.2294	1.659***	0.4679	0.998	0.3050
1995 (2)	1.153	0.2136	1.500	0.4344	0.598	0.1994
1996 (1)	1.129	0.1965	1.176	0.3229	0.625	0.1821
1996 (2)	1.185	0.2080	1.228	0.3413	0.578***	0.1750
1997 (1)	0.998	0.1717	0.903	0.2456	0.398***	0.1154
1997 (2)	1.029	0.1772	1.034	0.2822	0.399***	0.1167
1998 (1)	0.672	0.1152	0.488***	0.1329	0.192***	0.0567
1998 (2)	0.742	0.1269	0.638*	0.1730	0.186***	0.0545
1999 (1)	0.495***	0.0866	0.354***	0.0993	0.143***	0.0454
1999 (2)	0.528***	0.0914	0.494**	0.1361	0.100***	0.0320
2000 (1)	0.312***	0.0964	0.189***	0.122	0.178**	0.1279
2000 (2)	0.234***	0.0699	6.112***	18.842	0.043***	0.0456
2001 (1)	0.251	0.2854	2.615	2.701	0.260	0.3512
2001 (2)	8.361***	3.043	0.995	0.6521	0.362	0.03335

P-value: *p < 0,10; **p < 0,05; ***p < 0,01. 1- The estimations of the other variables are found in Appendix III

2- Robust standard errors.3- Number 1 indicated Men and 2 the Women's. N=204145

Source: Elaborated by the author

Given the cohort estimates in Table 10, the most statistically significant effects are for the periods whose year of birth is between 1996 and 2000 - regardless of whether they are male or female. Attention given to the fact that the relative risks of being a regular smoker are less than 1, which is interpreted as a lower propensity to smoke, in the age groups from 16 years (reference to the year 2012) to 19 years (base year of 2015).

Unlike Models I to VII, which presented relative risks greater than 1, for the effects of pairs on occasional and regular smokers, here there is a decrease in these, and the values obtained were significant at 1%, assuming by hypothesis that time-sequential social interactions are relevant to explaining behavior in the present.

5.3- Discussion

Taking into account the estimated models, in general terms, the effect size for the social interaction variable proved to be relevant, to the point that in almost all parameters, the statistical significance of the relative risk ratios is for a p value of 1%. The exceptions did not change the main result, which was the relevance of peer effects. Evidence in general supports that peer effects are positive and relevant. Becker (2017) showed (with 2SLS model) that the social interaction effect would increase the probability of smoking for 14-year-old students by 25 percentage points, and in the present research, no age division of peer effects was made.

It is important to note that all the analyzes that will come later are not directly comparable, as issues of measurement of variables, the estimator used, for example, were not the same.

In terms of detailed analysis, firstly, our reference model showed that contagion effects when increasing 1 unit in its proportion increase the relative risk of being an occasional and regular smoker by a factor of 3.44 and 3.21, respectively. This model serves only as an initial indicator of how the relative risks of the pair effects of the next estimations will behave.

In model II, the relative risk for regular and occasional smokers is 3,576 and 3,367, compared to never having smoked. This result that the effects of pairs in models estimated via maximum likelihood is somewhat similar to the work of Scalici and Schulz (2017), considering that in this research, the pairs were increased by norms against the use of cigarettes within the family; Odds ratios of greater than 24 percentage points were found for "heavy smokers".

Both the question of age, as well as the proxy of income availability, make the relative risk of using cigarettes occasionally and regularly to be increased, as shown in Table 4. This confirms our first hypothesis that these individual variables have RRR greater than 1.

Model II, referring to the estimates that show greater chances of being in classes with smokers, shows that the impact is positive for smoking behavior. This impact on smoking

behavior corroborates that found in Gaviria and Raphael (2001), McVicar (2001) and Nikaj (2017), both studies on peer effects at the class level.

As mentioned in the Introduction to this article, possible biases in the estimation of peer effects, resulting from endogenous self-selection and factors that are generally omitted in the analyzes of this literature, were mitigated through the other estimated models (II to VII). This referred to the inclusion of other variables and the verification of their effect on the responses (Y) and on the social interaction effect.

On the other hand, for model III, when the family context is inserted, there is a small increase in the relative risk of being an occasional and regular smoker, in a few decimal places.

The hypothesis that students' family bias variables affect (Y) responses is partially supported. Parents or guardians who are more attentive to their children reduce the relative risks of being occasional and regular users (RRR=0.655 and 0.631), when compared to the response category. Using the mother's level of education, the hypothesis that higher levels of education always decrease the relative risks of being a smoker was not observed. The presence of smoking parents in the family increased the relative risk only for occasional and regular users.

Moving on to Models IV to VI, when adding the presence of an anti-smoking policy, this was not significant enough to affect the relative risks of peer effects, with the exception of a p-value of at least 10% that impacts regular cigarette use. The verification of implicit norms in school homes would provide a better view of the institutional context of schools, as well as indicators of aggregate violence in the localities of the school environment.

Objectively, the inclusion of the variable consumption of other drugs and the effects of year and region confirmed that these regressors in general contribute to explain the behavior of smoking. Being a student in 2015 increases the relative risks of being an occasional and regular smoker, when compared to not smoking (RRR=2,345 and 1,606). Being a student in the Northeast region, the relative risk decreases by a factor of 0.823 and 0.829.

Research with multi-year data showed that effects for years are relevant (Alli and Dyer, 2009) and this article found statistical significance, although stacked data that did not cover the same individuals were used.

As for the hypothesis that the consumption of complementary drugs is also associated with the consumption of industrialized cigarettes, it was confirmed, since in the results presented, the use of other drugs presented an increased relative risk for being an occasional and regular smoker. Similar to the evidence seen in Hallal et al. (2017), who found adjusted odds ratios of 7.6 and 4.4 for the complementarity of tobacco and alcohol to the use of other tobacco products, the present research strongly supports the hypothesis of a strong association of interrelated substance consumption. The data used came from PeNSE (2012).

Not least, the mother's education is a strong predictor of the individual being an occasional smoker and regular will depend on certain levels of mother's education and not for all scales, a hypothesis not supported by the estimates. Participation in the labor market was an important predictor of smoking behavior, in addition to the consumption of other drugs (alcohol, marijuana, etc.), which corroborates the hypothesis that the proxy for income and "complementary goods" exists in the data.

Finally, the lag of the effects of peers as a corrector of one of the forms of endogeneity presented statistically significant relative risks, which corroborates those sequential behaviors over time are present, as indicated in the relative risks of 0.614 and 0.943 for occasional and regular smokers. The interpretation made from the RRRs is that the presence of smoking peers in 2012 does not interfere with smoking behavior in the present (2015). Clark and Lohéac (2007), using panel data (probit model) from Add Health and assuming that smoking behavior is stable over one year, found marginal effects at the statistical level of 10% of the social interaction variable, for groups of boys in schools (effect of 0.253).

Another important study by Robalino and Macy (2018), using an IV probit model in a panel data structure from the American AddHealth, found that the probability of adolescents smoking increases with successive time lags by 4.5 percentage points; given the differences in the results of the aforementioned research, our study also differs from this one, with the use of time and sex cohorts, a methodological innovation.

Regarding the effects of year of birth and sex cohorts, the relative risks of being born in the late 1990s - irrespective of sex - are less than 1, which indicates that a considerable group of young people in those years are at lower risk. of being regular smokers.

5– Final Remarks

This research sought to bring to light a theme that is still little explored in the Brazilian context, which is the effect of peers on tobacco consumption, using theoretical and methodological issues as a contribution to analyzing behavioral data from Brazilian students at the same time. Questions related to information limitation and modeling of peer effects in networks are necessary to assess health outcomes for young people in Brazil, which this research cannot cover so robustly. The availability of large volume samples and data makes a more complete task difficult.

Furthermore, for the very analytical cut chosen for this research, which was the impact of the peer effect at the class level, a limitation, as this level of social interaction may not be the most adequate for the data in the Brazilian sample. Sociological factors that can be useful as a preliminary step in research may contribute to future research designs.

Also recurrent in the current state of the art of literature on social interactions, research that measures how genetic factors at the individual level interrelate and correlate with environmental factors can show the peculiarities of behavior in relation to health in general. Sociological factors that can be useful as a preliminary step in research have already been found, for example, in the review of

It is important for future Brazilian research in health economics to incorporate the references mentioned above, converging economic sciences with what is done in other sciences.

For the present case of this research, it appears that the effects of pairs are significant as a whole, and in this case, we find robust evidence that is in agreement with classic works and Brazilian research on the subject.

The use of variables to control for peer effects on smoking was used here as an innovation, whereas in general cases the instrumental variables methodology is widely used. As a result of this, temporal effects (year) were explored in order to also add to the estimates, the contribution of this regressor being very important to explain the use of cigarettes, in this case a change of a 3-year period from 2012 to 2015. No less importantly, a scale of concern from parents to young people over two periods of time, also proved to be an important psychological level variable on smoking.

Unlike academic research that takes into account the comparative geographic effects by countries around the world, here a strong result was found in favor of some Large Brazilian Regions (intra-local heterogeneity) operating as moderators of peer effects, which differentiated if of the above. Because there is no sample data from 2012 and 2015 on drug use for the regions analyzed, taking into account socioeconomic, cultural and genetic aspects, no more than what was used can be explored in this article.

In terms of implications for public policies and the Brazilian context, due to the lack of a robust set of applied research already carried out, it is difficult to draw conclusions from the results presented here. However, bringing an analysis based on the Theory of Public Choice (Leeson and Thompson, 2021) on health issues in other countries, drug and food control and regulation policies generally tend towards private rather than public purposes; the allocation of health expenditures reflect private interests and social welfare gains do not occur, but regulatory and interest group capture of policy-making bodies.

Outlined both the above, together with applied research that identifies the transmission of normative values from parents to children and effects of social contagion to classmates for adolescents, whose behavior can be changed, this research agenda in the Brazilian context is open and demanded in the future.

REFERENCES

- AHO, A. et al. The relationship between peer relations, self-rated health and smoking behaviour in secondary vocational schools. **Nursing Open**, vol. 6, n. 3, pag. 754-764, 2019.
- ALEXANDER C. et al. Peers, schools, and adolescent cigarette smoking. **Journal of Adolescent Health**, vol. 29, n. 1, pag. 22-30, 2001.
- ALLI, M. M. DYER, D. S. Estimating Peer Effects in Adolescent Smoking Behavior: A Longitudinal Analysis. **Journal of Adolescent Health**, vol. 45, n. 4, pag. 402-408, 2009.
- ALLEN, J. P. et al. Predictors of susceptibility to peer influence regarding substance use in adolescence. **Children Development**, vol. 83, n. 1, pag. 337-350, 2012.
- ALVES, R. F. PRECIOSO J, BECOÑA E. Smoking behavior and secondhand smoke exposure among university students in northern Portugal: Relations with knowledge on tobacco use and attitudes toward smoking. **Pulmonology**, maio-19, 2020.
- ARROW, K. Methodological Individualism and Social Knowledge. **The American Economic Review**, vol. 84, n. 2, 1994.
- BANZER, R. et al. Factors associated with different smoking status in European adolescents: results of the SEYLE study. **European Society of Child and Adolescent Psychiatry**, vol. 26, n. 11, 2017.
- BASTOS, Francisco Inácio Pinkusfeld Monteiro et al. (Org.). III Levantamento Nacional sobre o uso de drogas pela população brasileira. Rio de Janeiro: **FIOCRUZ/ICICT**, 2017. 528 p.
- BECKER, G. A theory of social interactions. **Journal of Political Economy**, vol. 82, n. 6, pag. 1063-1093, 1974.
- BECKER, G. MURPHY, K, Social Economics, Cambridge: Harvard University Press, 2000.
- BECKER, K. L. O efeito da interação social entre os jovens nas decisões de consumo de álcool, cigarros e outras drogas ilícitas. **Estudos Econômicos**, vol. 47, n. 1, pag. 65-92, 2017.
- BERRY, D. et al. Understanding Health Decision Making: An Exploration of Homophily. **Social Science Medicine**, vol. 14, pag. 118-124, 2018.

BLUME, L. DURLAUF, S. Identifying social interactions: a review, **Wisconsin Madison – Social Systems**, Working Papers 12, fev. 2005.

BOUNDREAUX, D. MEINERS, R. Externality: Origins and Classifications. **Natural Resources Journal**, vol. 59, n. 1, pag. 1319-1329, 2019.

BRAMOULLÉ, Y. DJEBBARI, H. FORTIN, B. Peer Effects in Networks: A Survey. **Annual Review of Economics**, vol. 12, pag. 603-629, 2020

BRITO, E. S. et al. A Cross-Sectional Evaluation of Cigarette Smoking in the Brazilian Youth Population. **Front. Public Health**, nov. 2021.

BROCK, W. DURLAUF, S. Discrete Choice with Social Interactions, **Review of Economic Studies**, vol. 68, n. 2, pag. 235-260, 2001.

_____. Multinomial choice with social interactions. NBER Technical Working Paper Series 288, mar. 2003. **NATIONAL BUREAU of ECONOMIC RESEARCH**. Disponível em <https://www.nber.org/papers/t0288> Acesso em: 10 of out. de 2021.

CARD, D. GIULIANO. L. Peer Effects and Multiple Equilibria in the Risky Behavior of Friends. **The Review of Economic and Statistics**, vol. 95, n. 4, pag. 1130-1149, 2013.

CAWLEY, J. et al. Testing for Peer Effects Using Genetic Data. **Health Economics**, vol. 28, n. 7, pag. 937-952, 2019

CHITEJI, N. Time preference, noncognitive skills and well-being across the life course: Do noncognitive skills encourage healthy behavior? **American Economic Review**, vol. 100, n.2, pag. 200-204, 2010.

COHEN, G.L. PRINSTEIN, M. J. Peer contagion of aggression and health risk behavior among adolescent males: an experimental investigation of effects on public conduct and private attitudes. **Children Development**, vol. 77, n. 4. pag. 967-983, 2006.

CLARK, A. E. LOHÉAC, Y. "It wasn't me, it was them!" social influence in risky behavior by adolescents. **Journal of Health Economics**, vol. 26, n. 4. pag. 763-784, 2007

DOMINGUE, B. et al. The social genome of friends and schoolmates in the National Longitudinal Study of Adolescent to Adult Health. **PNAS**, vol. 115, n. 4, pag. 702-707, 2018.

- FERREIRA, V. R. et al. Smoking, alcohol consumption and mental health: Data from the Brazilian study of Cardiovascular Risks in Adolescents (ERICA). **Addictive Behaviors Reports**, vol. 9. junho de 2019.
- FLETHCER, J. M. Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects. **Health Economics**, vol, 19, n. 4, pag. 466-484. 2010.
- GAVIRIA, A. RAPHAEL, S. School-based peer effects and juvenile behavior. **The review of economics and statistics**, vol. 83, n. 2, pag. 257-268, 2001.
- GLAESER, E. L. SACERDOTE, B. I. SCHEINKMAN, J. A. The Social Multiplier. **Journal of the European Economic Association**, vol. 1, n. 2/3, pag. 345/353, 2003
- GRAHAM, B. HAHN, J. Identification and estimation of the linear-in-means model of social interactions. **Economics Letters**, vol. 88, pag. 1-6, 2005.
- GREEN, M. P. et al. A Closer Look at Smoking Among Young Adults: Where Tobacco Control Should Focus Its Attention. **American Journal of Public Health**, vol. 97, n. 8, pag. 1427-1433, 2007.
- GRIFFIN et al. Differential Susceptibility: The Genetic Moderation of Peer Pressure on Alcohol Use. **Journal Youth Adolescents**, vol. 44, n.10, pag1841-1853, 2015.
- HALLAL, ALLC. et al. Uso de outros produtos do tabaco entre escolares brasileiros (PeNSE 2012). **Cadernos Saúde Pública**. (2017).
- HSIEH, C. S. KIPPERSLUIS, H. Van. Smoking initiation: Peers and personality. **Quantitative Economics**, vol. 9, n.2, pag. 825-863, 2018.
- HSIEH, C. S. LIN, X. Social interactions and social preferences in social networks. **Journal of Applied Economics**, vol. 36, n. 2, pah. 165-189, 2021.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Pesquisa Nacional de Saúde do Escolar – PeNSE 2012. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/educacao/9134-pesquisa-nacional-de-saude-do-escolar.html?edicao=17047>. Acesso em: 21 of out de 2021.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Pesquisa Nacional de Saúde do Escolar – PeNSE 2015. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/educacao/9134-pesquisa-nacional-de-saude-do-escolar.html?=&t=o-que-e> Acesso em; 20 of out. de 2021.

INSTITUTO NACIONAL DO CÂNCER - INCA. Causas e prevenção. Crianças, adolescentes e jovens. Disponível em: <https://www.inca.gov.br/tabagismo/criancas-adolescentes-jovens> Acesso em: 25 de mar. de 2021.

LEESON, P. T. THOMPSON. H. A. Public Choice and Health Public. **Public Choice**. <https://doi.org/10.1007/s11127-021-00900-2>. 2021.

LI. Y. GUANG. G. Peer influence on obesity: Evidence from a natural experiment of a gene-environment interaction. **Social Science Research**, vol. 83, n. 10, 2021.

LILLARD, D. R. The Economics of Nicotine Consumption. NBER Working Paper Series 26912, março de 2020. LIU, J. et al. Social Engagement and Elderly Health in China: Evidence from the China Health and Retirement Longitudinal Survey (CHARLS). **International Journal of Environmental Research and Public Health**, vol. 16, n. 2, 2019.

MANSKI, C. Identification of Endogenous Social Effects: The Reflection Problem. **Review of Economic Studies**, vol. 60, n. 3, pag. 531-542, 1993

_____. Economic Analysis of Social Interactions. **Journal of Economics Perspectives**, vol. 14, n. 3, pag. 115-136, 2000.

MATSON, E. W. Satisfaction in action: Hume's endogenous theory of preferences and the virtues of commerce. **Journal of Economic Behavior & Organization**, vol. 183, pag. 849-860, 2020.

McVICAR, D. Estimates of peer effects in adolescent smoking across twenty-six European Countries. **Social Science and Medicine**, vol.73, n.8, pag. 1186-1193, 2011.

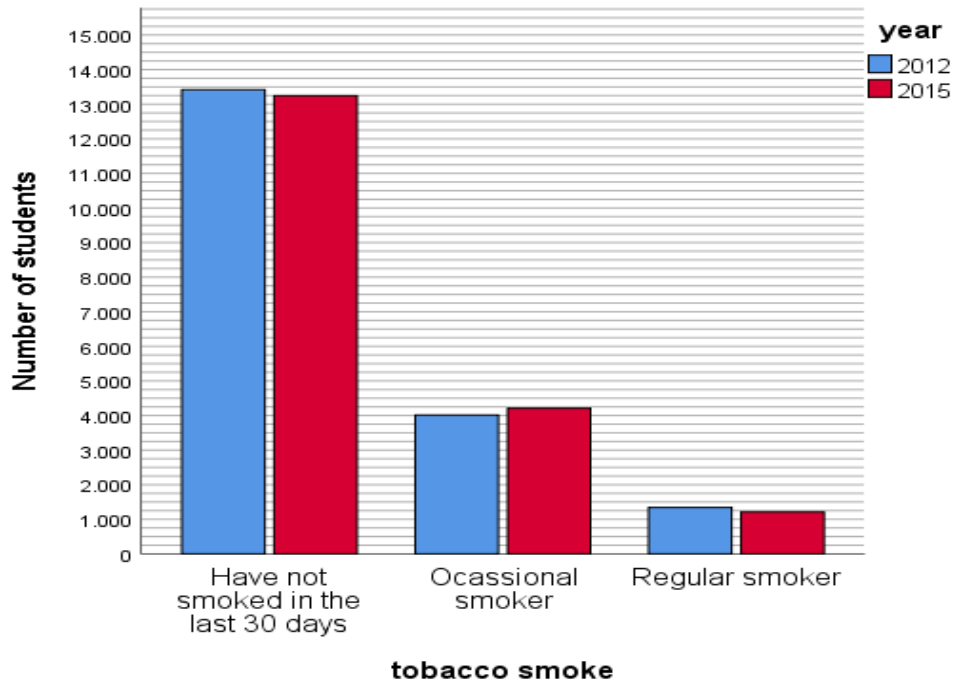
McVICAR, D. POLANSKI, A. Peer Effects in UK Adolescent Substance Use: Never Mind the Classmates? **Oxford Bulletin of Economic and Statistics**, vol. 76, n. 4, pag. 589-604, 2014.

MOFFIT, R. Identification and estimation of dynamic models with a time series of repeated cross-sections. **Journal of Econometrics**, vol. 59, n. 1-2, pag. 99-123, 1993.

- NIKAJ, S. Peer Effects and Youth Smoking in the European Global Youth Tobacco Survey. **Review of Economics Perspectives**, vol. 17, n.2, pag. 219-238, 2017.
- NGUYEN, N. et al. Using Peer Crowd Affiliation to Address Dual Use of Cigarettes and E-Cigarettes among San Francisco Bay Area Young Adults: A Cross Sectional Study. **International Journal of Environmental Research and Public Health**, vol.17, n. 20, 2020.
- PIÑA, J. A. et al. Chapter One - Social Influences on Nicotine-Related Behaviors. **International Review of Neurobiology**, vol. 140, pag. 1-32, 2018.
- POWELL, L. M. CHALOUPKA, F. J. Parents, public policy, and youth smoking. **Journal of Policy Analysis and Management**, vol. 24, n. 2, pag. 93-112, 2005.
- REYNOLDS, C. L. The Effect of Dormitory Residence during College on Student Outcomes. **Journal of Human Capital**, vol. 14, n. 2, 2020.
- ROBALINO, J. D. MACY, M. Peer effects on adolescent smoking: Are popular teens more influential? **Plos One**, vol. 13(7), 2018.
- SANCHEZ, Z. L. et al. Trends in alcohol and tobacco use among Brazilian students: 1989 to 2010. **Revista de Saúde Pública**, vol. 49, n. 70, 2015.
- SCALILI, F. SCHULZ, P. J. Parents' and peers' normative influence on adolescents' smoking: results from a Swiss-Italian sample of middle schools students. **Substance Abuse Treatment, Prevention, and Policy**, vol. 12, n.2, 2017.
- SOTOUDEH, R. MULLAN HARRIS, K. CONLEY, D. Effects of the peer metagenomic environment on smoking behavior. **PNAS**, vol. 116, n. 33, pag. 16302-16307, 2019.
- TANKSLEY, P.T. et al. The Genome-Wide Study of Human Social Behavior and Its Application in Sociology. **Frontiers in Sociology**, vol. 26, January, 2019
- Van de BRAND, F. A. et al. The Association of Peer Smoking Behavior and Social Support with Quit Success in Employees Who Participated in a Smoking Cessation Intervention at the Workplace. **International Journal of Environmental Research and Public Health**, vol. 16, n. 16, pag. 2019.

APPENDIX I

Graph A – Tobacco smokers' due year



Source: elaborated by the author

APPENDIX II

Table A- Analyzing multicollinearity in Model I.

Model	Tolerance	VIF	Model	Tolerance	VIF
peereffects	0,972	1,029	others	0,946	1,057
age	0,92	1,087	drugs		
sex	0,926	1,08	parents att.	0,994	1,006
			mother's	0,946	1,057
partlabourmark	0,955	1,047	educ		
physical act	0,942	1,061	Parents	0,98	1,021
			smoke		
region	0,965	1,037	tobacco	0,19	5,256
			policy		
			year	0,19	5,262

Source: elaborated by the author

Table B – Model VI (controlling for the effect of year and Major Brazilian Regions)

Variables (Omitted values are the bases for comparison; Non-smokers represent the base response category)	Have not smoked in the last 30 days		Occasional smoker		Regular Smoker	
	Relative ratio risk	Std. error¹	Relative ratio risk	Std. error¹	Relative ratio risk	Std. error¹
Peer effects	1.487**	.2256	2.501***	.6580	2.064***	.9616
Age						
14 years old	1.163***	.0226	1.093***	.0377	1.068***	.0701
15 to 19 years old	1.375***	.0287	1.416***	.0515	1.602***	.1064
Sex						
Women	.995	.0138	1.034	.0247	.885***	.0381
Participatin in labour market	1.104***	.02169	1.253***	.0393	1.410***	.0707
Physical activity						
Few	.977	.0154	1.017***	.0277	.918*	.0446
1 to 3 days a week	1.004	.0181	1.034***	.0320	.942	.0513
4 to 7 days a week	.622***	.0865	1.452***	.2496	1.385	.4172
Mothers education						
Elementary school inc.	1.002	.3068	1.040	.0519	1.179*	.1026
Complete primary edu.	1.078***	.0221	.990	.0348	1.108	.0703
Incomplete high school	1.062	.0305	.970	.0492	.863	.0841
Complete high school	1.049**	.0308	1.048	.0516	1.104	.0983
Incomplete college	1.040**	.0223	.925***	.0342	1.068	.0707
Complete college	1.091***	.0350	1.009	.0550	1.009	.1003
I do not know	.969	.0233	.905***	.0375	1.209	.0864
Parent attention						
Sometimes	.964***	.0168	.917***	.0276	.848	.0443
Always and often	.770***	.0140	.741***	.0229	.793	.0419
Parents smoke	1.259**	.0187	1.379***	.0341	1.345	.0580
Anti-smoking policy						
No anti-smoking policy	.916	.0270	.941***	.0483	.872	.0823
Yes anti-smoking policy	1.705	.1173	2.296***	.2895	1.835***	.4182
Others drugs	2.612	.0587	4.716***	.1477	9.625***	.4566
Year	1.757272	.1123	2.345***	.2754	1.606**	.34153
Region						

Norte	.833***	.0153	.817***	.0261	.763***	.0451
Nordeste	.797***	.0171	.823***	.0300	.829***	.0546
Sul	1.084***	.0248	1.052***	.0407	1.119*	.0742
Sudeste	1.022	.0229	.990***	.0379	1.318***	.0836

P-value: *p < 0,10; **p < 0,05; ***p < 0,01. 1- Robust standard errors

Source: Elaborated by the author

APPENDIX III

Table A – Descriptive statistics for cohorts (birth x sex¹)

Cohort's dummies (birth x sex)	Freq.	Percent
1993 (1)	598	0.29
1993 (2)	664	0.33
1994 (1)	1,036	0.51
1994 (2)	850	0.42
1995 (1)	3,792	1.86
1995 (2)	2,654	1.30
1996 (1)	9,856	4.83
1996 (2)	7,020	3.44
1997 (1)	21,720	10.64
1997 (2)	19,046	9.33
1998 (1)	45,058	22.07
1998 (2)	53,396	26.16
1999 (1)	15,250	7.47
1999 (2)	21,868	10.71
2000 (1)	500	0.24
2000 (2)	780	0.38
2001 (1)	38	0.02
2001 (2)	19	0.01

1- Number 1 indicated Men and 2 the Women's.

Source: Elaborated by the author

Table B - Model VII - Pseudo-panel with lagged mean social interaction effect (others variables). Equation 12

Variables (Omitted values are the bases for comparison; Non-smokers represent the base response category)	Have not smoked in the last 30 days		Occasional smoker		Regular Smoker	
	Relative ratio risk	Std. error	Relative ratio risk	Std. error	Relative ratio risk	Std. error
Participation in labour market	.763***	0.0213	0.589***	0.0259	0.472***	0.0322
Physical activity						
Few	.941	.0305	1.102*	.0619	.838*	0.0841
1 to 3 days a week	.925**	.0237846	.998	.0456	.757***	0.0604
4 to 7 days a week	1.052	.0276322	1.069	.0502	.915	0.0723
Mothers education						
Elementary school inc.	1.162***	0.0472	0.867	0.0581	.856	.0948
Complete primary edu.	1.082	0.0553	0.862*	0.074	.6455***	.1011
Incomplete high school	1.127**	0.0576	0.937	0.0795	.921	.1327
Complete high school	1.123***	0.0483	0.886*	0.0631	.927	.1093
Incomplete college	1.195***	0.0672	0.962	0.0897	.960	.1495
Complete college	1.723***	1.724	1.987***	1.997	3.537***	3.612
I do not know	.978	0.0408	0.851	.0582	0.876	0.0998
Parent attention						
Sometimes	.906***	0.0232	.789***	0.0347	0.663	0.0487
Always and often	.564***	0.0145	.485***	0.0220	0.451	0.0342
Parents who smoke	1.704***	.0357	2.043***	0.0732	1.976	0.1214
No anti-smoking policy	.5589956	.04345	.376***	0.0642	.560	0.1539
Others drugs	14.59764	0.5943	38.183***	1.840	109.310	7.573
Region						
Norte	.784***	0.0229	0.633***	0.033	0.572***	0.0571
Nordeste	.776***	0.0253	0.823***	0.0466	0.801**	0.0831
Sul	1.459***	0.0469	1.438***	0.0792	1.924***	0.1841
Sudeste	1.280***	0.0435	1.354***	0.0800	2.201***	0.2234

P-value: *p < 0,10; **p < 0,05; ***p < 0,01. 1- Robust standard errors

Source: Elaborated by the author

Table C - Original questions used in the research¹

Questions used from PeNSE (2012 and 2015)	Variables of empirical models
NOS ÚLTIMOS 30 DIAS, em quantos dias você fumou cigarros?	Y
Qual é a sua idade?	Age
Qual é o seu sexo?	
Você tem algum trabalho, emprego ou negócio atualmente?	Participation in labour market
NOS ÚLTIMOS 30 DIAS, em quantos dias você usou outros produtos de tabaco: cigarros de palha ou enrolados a mão, charuto, cachimbo, cigarrilha, cigarro indiano ou bali, narguilé, rapé, fumo de mascar etc? (não incluir cigarro comum)	
NOS ÚLTIMOS 30 DIAS, em quantos dias você tomou pelo menos um copo ou uma dose de bebida alcoólica? (Uma dose equivale a uma lata de cerveja ou uma taça de vinho ou uma dose de cachaça ou uísque etc)	Others drugs
NOS ÚLTIMOS 30 DIAS, quantos dias você usou droga como maconha, cocaína, crack, cola, loló, lança-perfume, ecstasy, oxy, etc?	
NOS ÚLTIMOS 7 DIAS, sem contar as aulas de educação física da escola, em quantos dias você praticou alguma atividade física, como esportes, dança, ginástica, musculação, lutas ou outra atividade?	Physical activity
Qual nível de ensino (grau) sua mãe estudou ou estuda?	Mother's education
NOS ÚLTIMOS 30 DIAS, com que frequência seus pais ou responsáveis entenderam seus problemas e preocupações?	Parent attention
Algum de seus pais ou responsáveis fuma?	Parents smoke
A escola tem alguma política, norma ou regra escrita que proíba o uso do tabaco nas suas dependências?	Anti-smoking policy
Região geográfica	Region

1- The questions used were recoded and the answers applied in the empirical model