

SOME INSIGHTS REGARDING FRAUD AT THE BACCALAUREATE EXAM IN ROMANIA

Octavian CEBAN

PhD Candidate,
Bucharest University of Economic Studies, Bucharest, Romania

E-mail: octavianceban1995@gmail.com



Ionela-Roxana GLAVAN

PhD, Assistant Lecturer, Department of Statistics and Econometrics,
Bucharest University of Economic Studies, Bucharest, Romania

E-mail: roxana.glavan10@yahoo.com



Andreea MIRICA

Assistant Lecturer, PhD,
Bucharest University of Economic Studies, Bucharest, Romania

E-mail: miricaandreea89@gmail.com



Roxana-Violeta PARTAS-CIOLAN¹

PhD Candidate,
Bucharest University of Economic Studies, Bucharest, Romania

E-mail: roxana.partas@gmail.com

Abstract

Fraud as a deceptive act is observed in a multitude of scenarios, the bacculaureate exams not being bypassed. Our research aims to provide insights into fraud tentative cases at the Bacculaureate exam in Romania based on a quantitative approach using microdata from 2021. How many students attempt fraud at the bacculaureate exam grouped by academic and demographic factors and what are the main aspects that influence the likelihood of fraud attempt are the topics explored in this paper. Author based computations and several logistic regression models are constructed, indicating that males, compared to females, have a higher probability to be expelled from the exam, and rural candidates are more likely to attempt fraud. In 2021, the technical maths exam remains the subject with the highest probability to be removed from exam due to fraud.

Keywords: logistic regression; bacculaureate; exam; fraud; demographics

Introduction

Academic fraud is wide-spread among upper-secondary students, both during their studies and at external final examinations, with no differences in terms of gender or school profile (Sorgo et al 2015, 67-87). As a consequence, cheating is very hard to fight and although recognised in the media, it is rarely prosecuted (Berkhout et al 2020, 20). In ex-soviet countries, the academic fraud during final high school examinations is a major challenge, as “teachers often feel obliged to ‘help’ students” during these exams (Plikšnys et al 2009, 60). Despite a considerable amount of academic literature dedicated to understanding why academic fraud occurs, there is still a limited understanding of the topic (Walker and Holtfreter 2015).

In response to the 2010 Bacalaureate fraud, Romania introduced in 2011 harsh punishments for teachers and students attempting to cheat in the exam as well as surveillance cameras in the exam rooms (Pop 2021). The campaign was effective in reducing cheating attempts as well as punishing frauds (Borcan et al 2017). However, fraud attempts during the Bacalaureate exams in Romania still exists today. In this respect, the purpose of this paper is to provide some insights of the characteristics of students that attempt to cheat at this exam.

Research studies analysing dishonest behaviour for high school students often use a questionnaire-based approach (see for example Brown-Wright et al 2013) or qualitative techniques (see for example Karyono et al 2019). The purpose of this paper is to provide some insights regarding fraud tentative cases at the Bacalaureate exam in Romania using a quantitative approach based on microdata provided by the Romanian Ministry of Education on the 9th of July 2021 session (Ministry of Education 2021b). According to the Ministerial Order for the organisation of the 2021 Bacalaureate Exam, any fraud tentative by a candidate leads to the removal of the candidate from exam (Ministry of Education 2020). More specifically, the paper will answer the following questions:

- Q1: How many students attempt fraud at the bacalaureate exam?
- Q2: What are the main factors that influence the likelihood of fraud attempt?

Methodology

To answer the first question, we will analyse the number of candidates expelled from the exam out of 10 000 candidates present. Several breakdowns grouped in two categories are performed: academic factors (subject, profile, type of learning and year of graduation) and demographic factors (sex and area of residency).

To answer the second question, several binary variables were created (table 1).

Table 1. Variable names and value; Source: designed by the authors

Variable Name	Values
variable_removed_from_exam	Removed from exam = 1, otherwise = 0
Sex_F	Yes = 1; No = 0
Sex_M	Yes = 1; No = 0
Profile_Humanities	Yes = 1; No = 0
Profile_Natural_resources_and_environmental_protection	Yes = 1; No = 0
Profile_Sciences	Yes = 1; No = 0
Profile_Services	Yes = 1; No = 0

Profile Sports	Yes = 1; No = 0
Profile Technical	Yes = 1; No = 0
Track Technology	Yes = 1; No = 0
Track Theoretical	Yes = 1; No = 0
Track Vocational	Yes = 1; No = 0
Type of learning full time	Yes = 1; No = 0
Area of residency RURAL	Yes = 1; No = 0
Area of residency URBAN	Yes = 1; No = 0
Subject ea exam Romanian language humanities profile	Yes = 1; No = 0
Subject ea exam Romanian language sciences profile	Yes = 1; No = 0
Subject eb exam	Yes = 1; No = 0
Subject eb exam Hungarian language sciences profile	Yes = 1; No = 0
Modern Language English	Yes = 1; No = 0
Modern Language French	Yes = 1; No = 0
Subject ec exam History	Yes = 1; No = 0
Subject ec exam Maths mathematics informatics	Yes = 1; No = 0
Subject ec exam Maths natural sciences	Yes = 1; No = 0
Subject ec exam Maths technical	Yes = 1; No = 0
Subject ed exam Animal and vegetal biology	Yes = 1; No = 0
Subject ed exam Geography	Yes = 1; No = 0
Subject ed exam Human anatomy	Yes = 1; No = 0
Subject ed exam Informatics Mathematics Informatics C Cplusplus	Yes = 1; No = 0
Subject ed exam Logic	Yes = 1; No = 0
Subject ed exam Physics theoretical	Yes = 1; No = 0

Next, several logistic regression models were created using “variable” as dependent variable and each of the other variables as factors. However, only those variables with enough number of observations were considered factors. Moreover, those who were absent were eliminated from the computation.

The computations were performed using Python libraries: pandas (McKinney 2010), numpy (Harris et al 2020), seaborn (Waskom 2021), statmodels (Skipper and Perktold 2010). For each model, the values for the coefficients as well as the associated probabilities are reported.

Results

Figure 1 below presents the number of candidates expelled from the exam per 10 000 candidates present at the exams between 2004 and 2021. The highest value is observed in 2011, the year when surveillance cameras were introduced in the exam rooms. A second important peak is observed in 2015, followed by a constant decline until 2019. In 2020, the year when the COVID-19 pandemic was declared, a third peak occurs in the series, but of a lower intensity.

Next, the number of candidates expelled from the exam per 10 000 candidates present at the exam is analysed by subject. The subjects at the Baccalaureate exam depend either on track or on the school profile and specialization. There are three subjects for which students must take a written exam: EA (Romanian language – humanities profile; Romanian language – sciences profile); EC (History; Maths – mathematics informatics specialisation; Maths – natural sciences specialisation; Maths – technical profile; Maths – pedagogical profile; Animal and vegetal biology; Inorganic chemistry – technical profile; Inorganic chemistry – theoretical track; Economics; Geography; Human anatomy; Informatics – Natural sciences profile, Pascal; Informatics – Mathematics Informatics profile C/C++; Informatics – Mathe-

matics Informatics profile Pascal; Informatics – Natural sciences profile C/C++; Logic; Organic chemistry – technical profile; Organic chemistry – theoretical track; Philosophy; Physics – technical profile ; Physics – theoretical track; Psychology; Sociology). The highest number of candidates expelled from the exam per 10,000 candidates present is registered for the EA subjects.

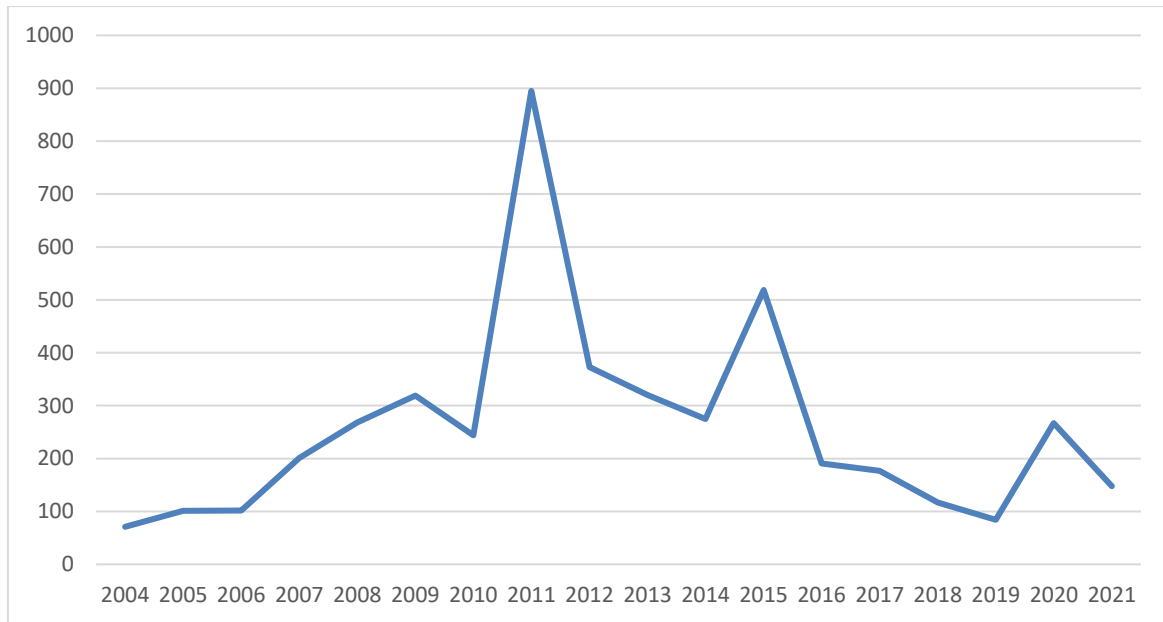


Figure 1. Number of candidates removed from the exam per 10 000 candidates present at the exam; source: designed by the authors based on data provided by the Ministry of Education (2021a)

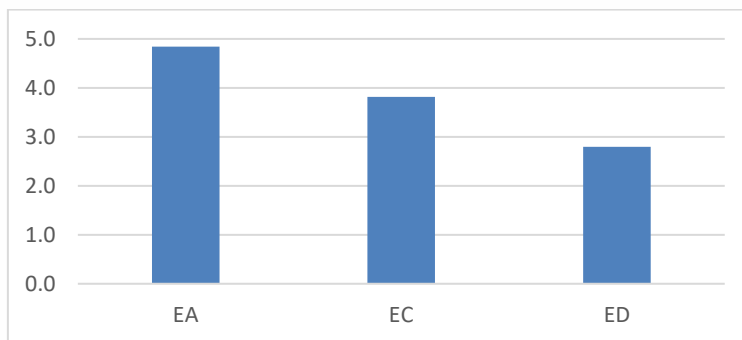


Figure 2. Number of candidates removed from the exam per 10000 candidates present at the exam by subject; source: designed by the authors based on data provided by the Ministry of Education (2021b)

Table 2 presents the number of candidates expelled from the exam per 10,000 candidates present by profile and type of courses, while figure 3 shows the share of candidates expelled from exam by profile. The highest number of candidates expelled from the exam per 10,000 candidates present is registered for the “Natural resources and environmental protection” profile for the part time student and the lowest number for both Sciences students attending evening courses and Services part time students. Moreover, 37% of the students expelled from the exam studied in vocational schools.

Table 2. Number of candidates removed from the exam per 10000 candidates present at the exam by profile and type of courses; source: designed by the authors based on data provided by the Ministry of Education (2021b)

Profile	Students attending evening courses	Full time students	Part time students
Arts	Not Available	2.817695	Not Available
Humanities	526.3158	2.942042	10.44932
Natural resources and environmental protection	47.61905	26.42706	625
Sciences	0	5.435296	30.39514
Services	69.20415	10.13069	0
Sports	Not Available	37.40065	Not Available
Technical	82.35294	35.75504	114.9425

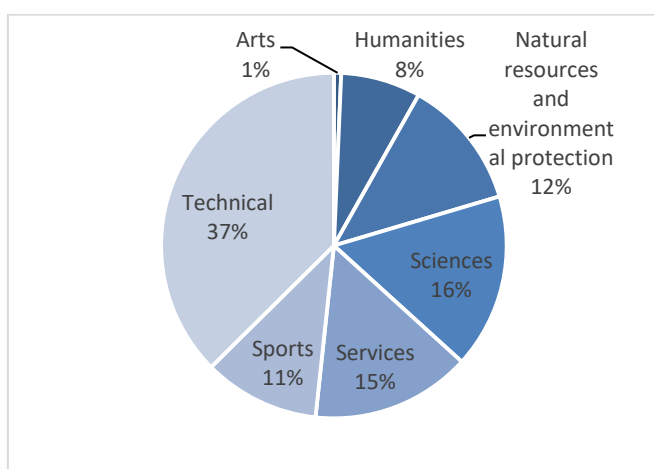


Figure 3. The share of candidates removed from exam by profile; source: designed by the authors based on data provided by the Ministry of Education (2021b)

The number of candidates expelled from the Baccalaureate exam per 10,000 candidates who graduated between 2020 – 2021 is approx. 4.5, approximately 13 times lower compared to those who graduated in other promotions. Moreover, when analysing the indicator by area of residency and sex, the highest value is observed for males from rural areas (table 3).

Table 3. Number of candidates removed from the exam per 10000 candidates present at the exam by area of residency and sex; source: designed by the authors based on data provided by the Ministry of Education (2021b)

	RURAL	URBAN	Total
F	7.42	3.34	4.71
M	24.48	17.26	19.45
Total	14.91	9.99	11.57

Table 4 below presents the results of the logistic regression models using Variable “removed from exam”, where 1 is removed and 0 is not removed, as a dependent variable

and each of the other variables as factors. All the selected variables proved to be significant at 5% level, with the following exceptions:

- Profile_Services
- Track_Vocational
- Subject_eb_exam_Hungarian_language__sciences_profile_
- Modern_Language_English
- Modern_Language_French
- Subject_ed_exam_Geography
- Subject_ed_exam_Human_anatomy
- Subject_ed_exam_Physics_theoretical_

Table 4. Logistic regression models results; source: designed by the authors based on data provided by the Ministry of Education (2021b)

Dependent Variable: Variable removed from exam				
Independent variable	intercept	p_value _intercept	Independent variable estimator	p_value _estimator
Sex F	-6.24	0.00	-1.42	1.3E-12
Sex M	-7.66	0.00	1.42	1.3E-12
Profile Humanities	-6.55	0.00	-1.41	6.9E-06
Profile Natural_resources_and_environmental_protection	-6.84	0.00	1.05	3.0E-05
Profile Sciences	-6.53	0.00	-0.95	1.9E-05
Profile Services	-6.75	0.00	-0.07	7.7E-01
Profile Sports	-6.84	0.00	1.26	2.2E-06
Profile Technical	-7.11	0.00	1.57	3.5E-20
Track Technology	-7.42	0.00	1.38	1.5E-15
Track Theoretical	-6.15	0.00	-1.51	7.3E-15
Track Vocational	-6.78	0.00	0.17	5.1E-01
Type_of_learning_full_time	-5.20	0.00	-1.66	4.3E-10
Area_of_residency_RURAL	-6.91	0.00	0.40	1.7E-02
Area_of_residency_URBAN	-6.51	0.00	-0.40	1.7E-02
Subject_ea_exam_Romanian_language_humanities_profile	-6.53	0.00	-1.50	1.6E-06
Subject_ea_exam_Romanian_language_sciences_profile	-8.03	0.00	1.50	1.6E-06
Subject_eb_exam	-8.76	0.00	2.05	4.2E-02
Subject_eb_exam_Hungarian_language_sciences_profile	-6.73	0.00	-1.64	1.0E-01
Modern_Language_English	-6.99	0.00	0.25	3.8E-01
Modern_Language_French	-6.76	0.00	0.03	9.1E-01
Subject_ec_exam_History	-6.57	0.00	-0.73	5.3E-04
Subject_ec_exam_Maths_mathematics_informatics	-6.63	0.00	-0.98	1.1E-03
Subject_ec_exam_Maths_natural_sciences	-6.69	0.00	-0.67	2.6E-02
Subject_ec_exam_Maths_technical	-7.42	0.00	1.38	1.5E-15
Subject_ed_exam_Animal_and_vegetal_biology	-7.15	0.00	1.50	4.2E-19
Subject_ed_exam_Geography	-6.72	0.00	-0.12	5.0E-01
Subject_ed_exam_Human_anatomy	-6.75	0.00	-0.07	7.2E-01
Subject_ed_exam_Informatics_Mathematics_Informatics_C_Cplusplus	-6.71	0.00	-2.15	3.2E-02
Subject_ed_exam_Logic	-6.66	0.00	-1.10	2.5E-03
Subject_ed_exam_Physics_theoretical	-6.71	0.00	-13.14	9.6E-01

Note: Table 4 Logistic regression models results; source: designed by the authors based on data provided by the Ministry of Education (2021b)

The probability to be removed from exam computed for valid logistic regression models are presented in Figure 4. Males have a higher probability to be expelled from the exam compared to females. Also, rural candidates are more likely to attempt fraud compared to urban ones. When considering the profile, technical candidates are the most likely to be expelled from the exam, compared to candidates from other profiles. Also, non-full-time students are more susceptible to fraud compared to the ones enrolled full time. Regard-

ing the exam subject, candidates taking animal and vegetal biology followed by those taking technical maths have the highest probabilities to be removed from exam.

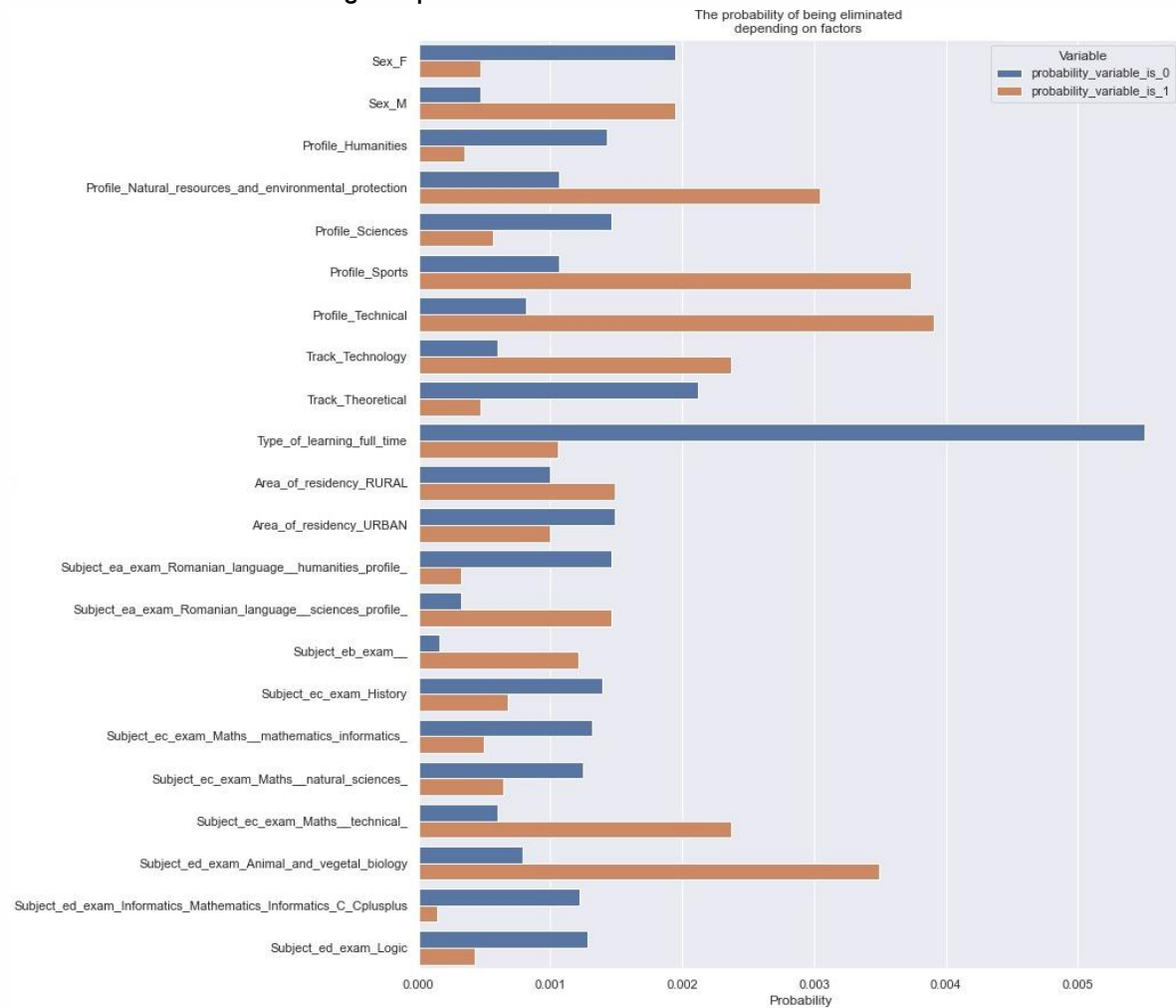


Figure 4. The probability to be removed from exam computed for valid logistic regression models; source: designed by the authors based on data provided by the Ministry of Education (2021b)

Conclusions

Researchers observe that exam fraud is on the rise during in recent years, when education has shifted to the online environment, but its scale is still difficult to quantify (Bilen and Matros, 2021). Moreover, academic studies have been conducted during covid-19 pandemic to assess the level of exam fraud.

Most important, the data presented in this study has shown an increase in exam fraud in Romania in 2021 Baccalaureate exam. The results indicate that males have a higher probability to be expelled from the exam in contrast to females. Also, it is observed that rural candidates are more likely to attempt fraud than urban candidates. Looking at the number of student fraud attempts at the baccalaureate exam grouped by academic factors, non-full-time students are more susceptible to fraud compared to the ones enrolled full time. Furthermore, our results indicate that the highest probability for students to be removed from

exam due to fraud, is observed when taking animal and vegetal biology followed by technical maths exam.

References

1. Berkhout, E., Pradhan, M., Rahmawati, D. S., & Swarnata, A. **From Cheating to Learning: An Evaluation of Fraud Prevention on National Exams in Indonesia.** RISE Working Paper 20/046. Oxford: Research on Improving Systems of Education. 2020
2. Bilen, E. & Matros, A. **Online cheating amid COVID-19.** *Journal of Economic Behavior & Organization*, Volume 182, 2021, pp 196–211, <https://doi.org/10.1016/j.jebo.2020.12.004>.
3. Borcan, O., Lindahl, M. & Mitrut, A. **Fighting Corruption in Education: What Works and Who Benefits?.** *American Economic Journal: Economic Policy*, Vol.9, No. 1, 2017, pp.180-209.
4. Brown-Wright, L., Tyler, K. M., Stevens-Watkins, D., Thomas, D., Mulder, S., Hughes, T., ... & Smith, L. T. **Investigating the link between home-school dissonance and academic cheating among high school students.** *Urban Education*, Vol. 48, No. 2, 2013, pp. 314-334.
5. Harris, C.R., Millman, K.J., van der Walt, S.J. et al. **Array programming with NumPy.** *Nature*, Vol. 585, 2020, pp. 357–362, DOI: 10.1038/s41586-020-2649-2.
6. Karyono, K., Suyahmo, S., & Utomo, C. B. **Implementation of Character Education for Creating Integrity Schools A Case Study At Public Junior High School 2 Pekalongan in 2015.** *Journal of Educational Social Studies*, Vol. 8, No. 1, 2019, pp. 111-119.
7. McKinney. **Data structures for statistical computing in python.** in *Proceedings of the 9th Python in Science Conference*. Vol. 445, 2010
8. Ministry of Education. **Ministerial Order regarding the organisation of the 2021 Bacalaureate Exam.** 2020, Accessed February 25, 2022. https://www.edu.ro/sites/default/files/_fi%C8%99iere/Legislatie/2020/ordin%205.453_2020_organizare_desfasurare_examen_bacalaureat%202021.pdf
9. Ministry of Education. **Report on the results of the 2021 Bacalaureate Exam.** 2021a, Accessed February 25, 2022. <https://www.edu.ro/sites/default/files/210709%20BAC%20DOSAR%20v1.pdf>
10. Ministry of Education. **Bacalaureate 2021 – Open data.** 2021b, Accessed February 25, 2022. <https://data.gov.ro/dataset/rezultate-bacalaureat-2021-sesiunea-i/resource/2db91f1c-77a6-44bd-a5d7-e1c384c49275>
11. Plikšnys, A., Kopincka, S., Hrynevych, L., & Palicarsky, C. **Transparency in education in Eastern Europe.** Open Society Institute. 2009.
12. Pop Zece ani de la primul Bacalaureat cu camere de supraveghere. 2021, Accessed February 25, 2022. <https://scoala9.ro/editorial-zece-ani-de-la-primul-bacalaureat-cu-camere-de-supraveghere/932/>
13. Skipper, S. & Perktold, J. *Statsmodels: Econometric and statistical modeling with python.* Proceedings of the 9th Python in Science Conference. 2010

14. Sorgo, A., Vavdi, M., Cigler, U., & Kralj, M. **Opportunity makes the cheater: High school students and academic dishonesty.** *CEPS Journal*, Vol. 5, No. 4, 2015, pp. 67-87.
15. Walker, N., & Holtfreter, K. **Applying criminological theory to academic fraud.** *Journal of Financial Crime*. 2015
16. Waskom, M. L. **Seaborn: statistical data visualization.** *Journal of Open-Source Software*, Vol. 6, No. 60, 2021, pp. 3021, <https://doi.org/10.21105/joss.03021>

¹ To whom correspondence should be addressed.