EDUCATION DISRUPTED. THE IMPACT OF THE RUSSIAN INVASION ON UNIVERSITY STUDENTS' PERFORMANCE IN UKRAINE

by

Dmytro Ternovenko

A thesis submitted in partial
fulfillment of the requirements for the
degree of

MA in Economic Analysis.

Kyiv School of Economics

2026

Thesis Supervisor: _____ Dariia Mykhailyshyna, PhD in Economics

Approved by _____
Head of the KSE Defense Committee, Professor

_____

_____

_____

Date _____

Kyiv School of Economics

Abstract

EDUCATION DISRUPTED. THE IMPACT OF THE RUSSIAN INVASION ON UNIVERSITY STUDENTS' PERFORMANCE IN UKRAINE

by Dmytro Ternovenko

Thesis Supervisor: Dariia Mykhailyshyna, PhD in Economics

The full-scale russian invasion of Ukraine has introduced unprecedented disruptions to the higher education system, threatening long-term human capital accumulation. This thesis investigates the causal impact of the exposure to war on academic performance and exam attendance of university students in Ukraine by analyzing a comprehensive dataset composed of standardized Master's degree entrance exams (UEE and UPE) from 2019 and 2021-2024 as well as of documented regional war exposure (measured through duration of air raid alerts and number of kinetic events and fatalities). This thesis employs a multi-model framework that includes a TWFE Continuous DID design to estimate cumulative war penalty on academic performance, alongside short-run TWFE OLS and a Two-Stage TWFE Logistic regression to account for immediate impact of war on scores and non-random absenteeism. The study reveals a complex landscape of both cumulative and immediate war effects on academic performance, where it significantly depresses the scores for some exams, and, at the same time, students show resilience in others. The data also exposes an influx of older male registrants in 2024 which we called 'mobilization effect' – this cohort likely seeks draft deferment rather than academic education. The main finding suggests that immediate war shocks primarily disrupt higher education through forced absenteeism in standardized exams rather than affecting the scores directly.

This thesis is dedicated to my younger self, who decided to pursue an academic education in Economics, to my present self, who was patient and stubborn enough to see it through to the end, and to my future self, who might pursue a doctoral degree someday.

To all my loved ones and close friends who supported me along the way.

TABLE OF CONTENTS

Chapter 1. INTRODUCTION 1
Chapter 2. LITERATURE REVIEW 5
Chapter 3. METHODOLOGY 13
3.1 Research question and hypotheses 13
3.2 Empirical strategy 14
3.2.1 TWFE Continuous DID design 15
3.2.2 Event-Study design 16
3.2.3 TWFE OLS design 17
3.2.4 Two-Stage TWFE Logistic regression 18
Chapter 4. DATA DESCRIPTION 19
4.1 Sources of data 19
4.2 Sample composition 21
4.3 Descriptive statistics 25
Chapter 5. EMPIRICAL RESULTS 34
5.1 TWFE Continuous DID 34
5.1.1 Assessment of baseline equivalence 34
5.1.2 Parallel trends assumption 35
5.1.3 Estimation results 37
5.2 TWFE OLS 41
5.2.1 TWFE OLS estimation results for 3 main exams 41
5.2.2 TWFE OLS estimation results for other UPEs 43
5.2.3 TWFE OLS robustness check: binary treatment 43
5.3 Two-Stage TWFE Logistic regression 46
5.3.1 TWFE Logistic regression estimation results for 3 main exams 46
5.3.2 TWFE Logistic regression estimation results for other UPEs 51
5.3.3 TWFE Logistic regression robustness checks 52
5.4 Multiple hypothesis test corrections 56
5.5 Data limitations and measurement constraints 58

TABLE OF CONTENTS – Continued

Chapter 6. CONCLUSIONS AND POLICY RECOMMENDATIONS 60
WORKS CITED 65
APPENDIX A. GEOGRAPHIC CHECKS 68
APPENDIX B. SPATIAL EVOLUTIONS 70
APPENDIX C. EVENT-STUDY ESTIMATION 72
APPENDIX D. FULL TWFE CONTINUOUS DID ESTIMATIONS 74
APPENDIX E. FULL TWFE OLS ESTIMATIONS (3 MAIN EXAMS) .. 78
APPENDIX F. ESTIMATION RESULTS FOR OTHER UPEs 82
APPENDIX G. TWFE OLS ROBUSTNESS CHECK FULL ESTIMATION RESULTS 92
APPENDIX H. FULL TWFE LOGISTIC REGRESSION ESTIMATIONS (3 MAIN EXAMS) 96
APPENDIX I. FULL TWFE LOGISTIC REGRESSION ESTIMATIONS (OTHER UPEs) 106
APPENDIX J. TWFE LOGISTIC REGRESSION ROBUSTNESS CHECKS FULL ESTIMATION RESULTS 124
APPENDIX K. MULTIPLE HYPOTHESIS TEST CORRECTIONS 136

LIST OF FIGURES

Number *Page*

Figure 1. Daily Kinetic Events vs. Daily Fatalities 21

Figure 2. Age Distribution of Master's Exam Registrants by Gender (2019, 2021-2024) 29

Figure 3. Yearly Evolution of Score Distributions 31

Figure 4. Absenteeism and Failure Rates for Exams 33

Figure 5. Parallel Trends: Raw Scores by Exam 36

Figure 6. Spatial Evolution of Exam Participation (2019, 2021-2024) 70

Figure 7. Spatial Evolution of Conflict Intensity (2022-2024) 71

Figure 8. Event-Study Plot 73

Figure 9. Forest Plots for TWFE Continuous DID Estimation Results for 3 Main Exams 77

Figure 10. Forest Plots for TWFE OLS Estimation Results for 3 Main Exams 81

Figure 11. Forest Plots for TWFE OLS Estimation Results for Other UPEs 91

Figure 12. Forest Plots for TWFE Logistic regression (Absenteeism) Estimation Results for 3 Main Exams 99

Figure 13. Forest Plots for TWFE Logistic regression (Conditional Failure) Estimation Results for 3 Main Exams 104

LIST OF TABLES

Number *Page*

Table 1. Research Hypotheses 13

Table 2. Student-Level Sample Cleaning Process 24

Table 3. Descriptive Statistics of the Analytical Sample 25

Table 4. Standardized Exams by Year (2019, 2021-2024) 30

Table 5. Baseline Equivalence Check 34

Table 6. TWFE Continuous DID Estimates of Cumulative War Exposure Across All Subjects 38

Table 7. TWFE OLS Estimates for 3 Main Exams 41

Table 8. TWFE OLS Robustness Check: Binary Treatment 44

Table 9. Impact of Immediate Shocks on Exam Absenteeism (Log-Odds) 47

Table 10. Conditional Probability of Exam Failure Given Attendance (Log-Odds) 49

Table 11. TWFE Logistic Regression Robustness Check: Linear Probability Model (Absenteeism) 52

Table 12. TWFE Logistic Regression Robustness Check: Linear Probability Model (Conditional Failure) 53

Table 13. TWFE Logistic Regression Robustness Check: Binary Treatment Logistic regression (Absenteeism) 54

Table 14. TWFE Logistic Regression Robustness Check: Binary Treatment Logistic regression (Conditional Failure) 56

Table 15. Student Observations by University Region and Year 68

Table 16. Student Observations by Physical Testing Regions and Year 69

Table 17. Event-Study Estimation Results 72

Table 18. Full TWFE Continuous DID Estimation - Foreign Language 74

Table 19. Full TWFE Continuous DID Estimation - General Competence 75

Table 20. Full TWFE Continuous DID Estimation - Law 76

Table 21. Full TWFE OLS Estimation - Foreign Language 78

Table 22. Full TWFE OLS Estimation - General Competence 79

Table 23. Full TWFE OLS Estimation - Law 80

LIST OF TABLES – Continued

Number *Page*

Table 24. TWFE OLS Estimates for Other UPEs 82
Table 25. Full TWFE OLS Estimation - Psychology 84
Table 26. Full TWFE OLS Estimation - Finance 85
Table 27. Full TWFE OLS Estimation - Management 86
Table 28. Full TWFE OLS Estimation - Economics 87
Table 29. Full TWFE OLS Estimation - Political Science 88
Table 30. Full TWFE OLS Estimation - Pedagogy 89
Table 31. Full TWFE OLS Estimation - Information Technology 90
Table 32. Full TWFE OLS Estimation with Binary Treatment - Foreign Language 92
Table 33. Full TWFE OLS Estimation with Binary Treatment - General Competence 94
Table 34. Full TWFE OLS Estimation with Binary Treatment - Law 95
Table 35. Full TWFE Logistic regression (Absenteeism) Estimation - Foreign Language 96
Table 36. Full TWFE Logistic regression (Absenteeism) Estimation - General Competence 97
Table 37. Full TWFE Logistic regression (Absenteeism) Estimation - Law 98
Table 38. Average Marginal Effects for Exam Absenteeism 100
Table 39. Full TWFE Logistic regression (Conditional Failure) Estimation - Foreign Language 101
Table 40. Full TWFE Logistic regression (Conditional Failure) Estimation - General Competence 102
Table 41. Full TWFE Logistic regression (Conditional Failure) Estimation - Law 103
Table 42. Average Marginal Effects for Conditional Exam Failure 105
Table 43. TWFE Logistic regression (Absenteeism) Estimates for Other UPEs 106
Table 44. Full TWFE Logistic regression (Absenteeism) Estimation - Psychology 108

LIST OF TABLES – Continued

Number *Page*

Table 45. Full TWFE Logistic regression (Absenteeism) Estimation - Finance 109

Table 46. Full TWFE Logistic regression (Absenteeism) Estimation - Management 110

Table 47. Full TWFE Logistic regression (Absenteeism) Estimation - Economics 111

Table 48. Full TWFE Logistic regression (Absenteeism) Estimation - Political Science 112

Table 49. Full TWFE Logistic regression (Absenteeism) Estimation - Pedagogy 113

Table 50. Full TWFE Logistic regression (Absenteeism) Estimation - Information Technology 114

Table 51. TWFE Logistic regression (Conditional Failure) Estimates for Other UPEs 115

Table 52. Full TWFE Logistic regression (Conditional Failure) Estimation - Psychology 117

Table 53. Full TWFE Logistic regression (Conditional Failure) Estimation - Finance 118

Table 54. Full TWFE Logistic regression (Conditional Failure) Estimation - Management 119

Table 55. Full TWFE Logistic regression (Conditional Failure) Estimation - Economics 120

Table 56. Full TWFE Logistic regression (Conditional Failure) Estimation - Political Science 121

Table 57. Full TWFE Logistic regression (Conditional Failure) Estimation - Pedagogy 122

Table 58. Full TWFE Logistic regression (Conditional Failure) Estimation - Information Technology 123

Table 59. Full LPM (Absenteeism) Estimation - Foreign Language 124

Table 60. Full LPM (Absenteeism) Estimation - General Competence 125

Table 61. Full LPM (Absenteeism) Estimation - Law 126

Table 62. Full LPM (Conditional Failure) Estimation - Foreign Language 127

LIST OF TABLES – Continued

Number *Page*

Table 63. Full LPM (Conditional Failure) Estimation - General Competence 128

Table 64. Full LPM (Conditional Failure) Estimation - Law 129

Table 65. Full TWFE Logistic regression (Absenteeism) Estimation with Binary Treatment - Foreign Language 130

Table 66. Full TWFE Logistic regression (Absenteeism) Estimation with Binary Treatment - General Competence 131

Table 67. Full TWFE Logistic regression (Absenteeism) Estimation with Binary Treatment - Law 132

Table 68. Full TWFE Logistic regression (Conditional Failure) Estimation with Binary Treatment - Foreign Language 133

Table 69. Full TWFE Logistic regression (Conditional Failure) Estimation with Binary Treatment - General Competence 134

Table 70. Full TWFE Logistic regression (Conditional Failure) Estimation with Binary Treatment - Law 135

Table 71. MHT Correction (TWFE Continuous DID) 136

Table 72. MHT Correction (TWFE OLS) 138

Table 73. MHT Correction (TWFE Logistic regression (Absenteeism)) 140

Table 74. MHT Correction (TWFE Logistic regression (Conditional Failure)) 142

ACKNOWLEDGMENTS

The author is wholeheartedly grateful to his thesis advisor Dariia Mykhailyshyna and the Research Workshop professors for their guidance, expertise, and support throughout, as well as for pushing to think critically about this study.

AI USE DECLARATION

During the preparation of this thesis, Generative AI (Google Gemini) was used as a supplementary tool to:

- idea brainstorming on a) model specification choice for absenteeism and failure rates, b) plot choice for presenting score distribution across years (ridge plot) and estimation results (forest plots), c) robustness checks to apply;
- coding for a) data cleaning and consolidation, b) spatial evolution visuals, c) functions and loops for iterative models estimation, d) generation of almost ready-to-paste correctly styled tables.

Any suggestions produced by the AI tool were critically evaluated and substantially modified by the author before being incorporated into the thesis.

The tool was not used to generate research findings, produce literature summaries used directly in the thesis, create citations, or conduct any form of data analysis. All research questions, theoretical framing, methodology, analysis, interpretation of results, and conclusions were developed independently by the author. The author takes full responsibility for the originality, validity, and academic integrity of the work presented in this thesis.

LIST OF ABBREVIATIONS

ACLED. Armed Conflict Location & Event Data.

AME. Average marginal effects.

CI. Confidence intervals.

DID. Difference-in-difference.

FDR. False discovery rate.

FE. Fixed effects.

LPM. Linear probability model.

MA. Master of Arts.

MCT. Master's Comprehensive Test.

MHT. Multiple hypothesis test.

MTAC. Master's Test of Academic Competence.

NMT. National Multisubject Testing.

OLS. Ordinary least squares.

PISA. Programme for International Student Assessment.

PTReg. Physical testing region.

SD. Standard deviation.

TWFE. Two-way fixed effects.

UCEQA. Ukrainian Center for Educational Quality Assessment.

UEE. Unified Entry Examination.

UNESCO. United Nations Educational, Scientific and Cultural Organization.

UPE. Unified Professional Examination.

Chapter 1

INTRODUCTION

The trajectory in which education in Ukraine has been developing since independence was irrevocably changed by a series of geopolitical aggressions from russia, culminating in its full-scale invasion of Ukraine on February 24, 2022. The invasion set off a complex series of crises, broadly affecting the economy and the educational sector of Ukraine. From continual destruction of physical assets to further displacement of students and educational staff and the psychological trauma inflicted on students, the war created "a new reality" for Ukrainian higher education students. This reality is characterized by the suspension of usual academic schedules, the forced mass migration of students, and imposed changes in the learning environment defined by regular air raid alerts, airstrikes, power outages, and a permanent threat for life and health (Polovko and Glotov 2023).

Education is considered among the key drivers of human capital accumulation and long-term economic growth, or economic recovery in the case of post-conflict societies. Thus, the disruption of education due to war causes additional obstacles for the future development of the Ukrainian economy. According to a World Bank assessment, exposure to armed conflict typically forces children out of school, which in turn generates short-term learning losses that are "never fully recovered," (Caruso et al. 2023) and leads to lower educational success, diminished productivity, and reduced lifetime earnings. Ukraine entered the war with its educational system already weakened by extensive school closures related to the COVID-19 pandemic, and continuous disruptions due to military attacks have pushed estimated learning losses to more than a full year of schooling, with estimated loss in future annual earnings per student exceeding 10% (Caruso et al. 2023). These war-induced setbacks

not only depress the current economy directly through reduced activity and destruction but also undermine one of the essential parts of post-war reconstruction: a well-educated, productive workforce capable of supporting long-term growth. Another World Bank's brief on education in Ukraine gives a notion of the amount of damage dealt: "Ukraine performed at par with its regional neighbors in eastern Europe prior to the pandemic, and it outpaced them in terms of students' learning resilience; however, …learning outcomes are now below the lowest-performing countries in Europe." (World Bank 2022). However, while these assessments provide a crucial overview of the general education sector, there remains a critical gap in understanding how this "new reality" specifically affects higher education. Unlike primary and secondary school students, university students are at a decisive stage immediately preceding labor market entry. Their performance is not only a measure of current learning but a predictor of future productivity and post-war economic recovery.

This thesis aims to start filling this gap by shifting the focus from secondary schooling to the under-researched cohort of university students. They face unique economic pressures, including imminent labor market entry and, considering the current context in Ukraine, specific mobilization risks. Specifically, this study investigates how different dimensions of war (ranging from cumulative intensity to instantaneous air raid alerts on the night before the exam) impact academic performance of higher education students. Mainly, the core of this research lies in determining whether students demonstrate resilience to the 'new reality' of war or war-related shocks created an unrecoverable penalty, which further affects human capital accumulation, and if so, how this impact differs across gender and age.

To answer these questions, we base the study on methodological framework established by Mykhailyshyna (2025) for the impact of war on secondary school students and high school graduates. To make it explicit, the short-run design

that isolates the immediate impact of war on exam scores directly follows the methodology of her paper. Expanding on this foundation, the original methodological and empirical contributions of this thesis are: (a) scope extension to higher education students by using data from the UCEQA on standardized exams that are mandatory to apply for a master's degree – UEE (Unified Entry Examination – Єдиний Вступний Іспит) and UPE (Unified Professional Examination – Єдине Фахове Вступне Випробування); (b) employing a TWFE Continuous DID to quantify the academic penalty of cumulative exposure to war; and (c) introducing a Two-Stage TWFE Logistic regression model to assess how exposure to war drives non-random absenteeism and conditional failure in the exam. We use the differences in war intensity across regions[1] and time periods to isolate the causal effect of the invasion on students' performance.

This study contributes to existing literature by focusing specifically on higher education. It shifts the focus from schooling, which dominates the current discourse on conflict economics, to the under-researched demographic of university students.

The findings reveal complex effect of war on academic performance among Ukrainian students. First, the TWFE Continuous DID estimates suggest that traditional student cohorts (females and males under 24) demonstrate resilience to cumulative war exposure in the Foreign Language exam, although it does not extend to all subjects. For exams like General Competence and Law, cumulative exposure to war, on the contrary, significantly depressed academic performance. Second, the data exposes a compositional shift in exam attendees – which we call a 'mobilization effect' – driven by an influx of older male registrants (men of age 24 or older), who are likely seeking draft deferments instead of pursuing academic education. This influx skews the aggregate

[1] Here and further 'region' stands for either 1st-level administrative region of Ukraine, i.e. oblast, or cities with special status – Kyiv or Sevastopol, if not explicitly stated otherwise.

performance observed in the data. Finally, short-run TWFE OLS estimations revealed expected negative effect of overnight air raid alerts on student scores along with the unexpected statistically significant positive effects of fatalities on academic performance of traditional cohorts of students. However, the Two-Stage TWFE Logistic regression provides the context to the observed anomaly by documenting a self-selection mechanism: while war significantly increases probability of student not showing up to the exam ('filtering out' less confident or underprepared students), the conditional probability of failure is lower for students who experienced more intensive exposure to war and still attended. This indicates that immediate war shocks primarily disrupt exam attendance rather than directly affect the scores.

Based on the findings, this study presents several policy implications for the Ministry of Education and Scienced of Ukraine. These include implementing flexible exam schedules to mitigate the immediate impact of air raid alerts on the night before an exam and expanding number of exam days and variant pools to minimize absenteeism driven by physical destruction prior to the exam. We also suggest accounting for mobilization-driven cohorts when assessing academic performance of Ukrainian students.

The remainder of this thesis is organized as follows: Chapter 2 provides a theoretical and empirical review of the literature regarding human capital and conflict; Chapter 3 details the research question and hypotheses of this study, and describes the econometric methodology framework applied; Chapter 4 describes the sample composition and provides a preliminary analysis of its two main parts – student-level and war-exposure parts; Chapter 5 presents and discusses the empirical results as well as existing data limitation and measurement constraints; and Chapter 6 concludes the main findings along with finalized policy recommendations.

Chapter 2

LITERATURE REVIEW

Understanding the connection between an armed conflict and educational outcomes requires an initial look into human capital theory, cognitive psychology, and assessments of the amount of educational infrastructure damaged. These help to build a framework that illustrates the complicated mechanism through which a war – such as a russian invasion in Ukraine – can disrupt the performance of university students.

The first theoretical basis for analyzing the relationship between war and academic performance lies within human capital theory. Perception of education as an investment originates from Becker's (1964) classical model, which was broadly accepted by the economists. In the model, instructional time, study effort, and institutional quality are among the inputs that education relies on. These are disrupted tremendously by the war: students may lose access to regular instruction, face interruptions in their study activities, or experience a decline in quality of teaching due to displacements, teacher shortages, and inability of educational institutions (whether schools or universities) to operate normally as in times of peace. From this perspective, the war itself causes a negative shock in productivity of education, which in turn reduces the efficiency of human capital accumulation in the short run even though schooling may continue formally.

The next portion of necessary framework lies within psychological and cognitive consequences of exposure to violence. Stress caused by such exposure impairs memory, attention, and overall mental health – all crucial to succeeding in studies. For instance, Lupien et al. (2009) show that increased stress hormones reduce cognitive performance, particularly in tasks that require continuous concentration and memory capacities. Additionally, studies on

youth affected by armed conflicts indicate that exposure to violence induces anxiety, trauma, and emotional instability, all harming the learning process. Blattman and Annan (2010) studied young people exposed to war in Uganda and documented substantial psychological burdens that further translated into poorer performance in education. Speaking more broadly, any stressful environment reduces students' academic outcomes (Deng et al. 2022). Through this mechanism, war increases cognitive load and emotional strain, directly affecting students even in relatively safer regions of Ukraine regardless of any other direct disruptions to the education sector.

Finally, the last set of mechanisms considered in this study through which war impact education negatively is disruption of learning infrastructure – both physical and digital. Armed conflict usually damages school buildings, limits student mobility, and reduces the availability of resources essential for education. In the context of Ukraine, both schools and universities face repeated bombings, electricity shortages, and internet connection disruptions, not mentioning displacement of both students and staff. Furthermore, constant attacks on energy infrastructure affect learning possibilities both directly through disruptions in electricity and heat, and indirectly through damaging communication network infrastructure. It makes it difficult to maintain regular educational schedules both online and offline, thus reducing study time, creating inconveniences in learning, and putting even more pressure on students' mental health, which lowers the efficiency of human capital accumulation through education.

All mechanisms described above are closely tied to each other, making up a logical and theoretical framework to perceive how dimensions of war affect students' performance directly and indirectly, in long-run as well as short-run. Through this framework, a reader can understand ***why*** the expected impact is negative.

Theoretical expectations on the adverse effects of war on education are strongly supported by broad empirical evidence from conflicts and other stressful events across the globe. Literature consistently demonstrates that armed conflict acts as a severe negative shock to education as a human capital accumulation process, creating long-term deficits that are often difficult to recover. Based on the evidence from World War II, the educational attainment of children who were in elementary schools during or immediately after the war decreased significantly with the size of educational loss to be equivalent to around 20% of a study year on average. This, in turn, seems to be a reason for a decrease in earnings, and "in Austria and Germany the magnitude of this effect can be situated in the order of 0.8% GDP" (Ichino and Winter-Ebmer 2004). Another evidence comes from the Gianmarco Leon's (2012) study on Civil Conflict in Perú, where the research showed that the timing of exposure to violence is crucial for educational outcomes. It was found that children in Peru who were exposed to violence during early life stages accumulated nearly a year less of schooling on average, with effects being irreversible in the long run for the very early childhood period violence exposure (Leon 2012). Besides, Swee (2015) identifies the war-caused effects of the Bosnian War on schooling attainment, finding that cohorts that experienced war more intensively "are less likely to complete secondary schooling but not primary schooling" (Swee 2015). Moreover, the negative effects were weaker for females compared to males, including the negative impact on mental health.

However, the losses that war causes are not limited to what we may call "quantity" measures, such as loss in years of schooling or completion rate mentioned above. They also occur in the "quality" of schooling for those who do not drop out in case of a stressful event. To measure this quality, we consider grades on standardized exams as a proxy.

Empirical research provides robust evidence that exposure to violence during the academic year directly disrupts students' performance in standardized

exams, especially if such exposure occurs right before the exam. For instance, the examination of the Israeli-Palestinian conflict shows that Palestinian students exposed to fatalities during their final year of high school were significantly less likely to pass the high-stakes exit exam (Brück et al. 2019). It is crucial to mention that the authors identified that the source for this effect was primarily the psychological burden of the conflict rather than school closures alone. Similarly, in the context of the so-called "War on Drugs" in Mexico, Michaelsen and Salardi (2020) demonstrate that the increase in local violence led to cognitive letdown, which in turn negatively affected the test scores, highlighting how exposure to violence reduces educational performance by impairing mental health and concentration capabilities. Besides, the negative impact of the presence of homicides in the week before standardized exams was much larger compared to the effect of being exposed to violence over the school year, which additionally highlights how short-term exposure around exam dates can disproportionately affect its outcomes (Michaelsen and Salardi 2020).

Although studies mentioned above (i.e., Brück et al. (2019) and Michaelsen and Salardi (2020)) isolate the direct impact of exposure to violence and psychological stress caused by it on students' performance, war also disrupts education through other channels: destruction of infrastructure and forced remote learning. To understand how these channels of education system disruptions affect students, we look at relevant studies on natural disasters and the COVID-19 pandemic, which led to destruction and online schooling, respectively.

The destruction of educational infrastructure is a frequent consequence of modern warfare, and it degrades academic outcomes even when students are not directly exposed to combat. Natural disasters are a sufficient example of such cases, and research on them demonstrates how the loss of physical infrastructure creates barriers to success in learning. For instance, Di Pietro

(2018) analyzed the impact of the 2009 L'Aquila earthquake in Italy. He finds that university students enrolled in the affected area have the probability of graduating on time less by 6.6 percentage points compared to a cohort from another university, whose schedules were not affected by any natural disaster at the time (Di Pietro 2018). Similarly, Thamtanajit (2020) used the 2011 severe flood in Thailand as a natural experiment and measured its impact on the students' scores on the national examination O-net. The results suggest statistically significant negative effects for most test scores for grades 6 and 9. The absolute value of loss in test scores attributed to the flood ranged between 0.03 and 0.11 standard deviations (Thamtanajit 2020). The findings from both papers suggest that the destruction of educational infrastructure – regardless of whether it is from an earthquake, a flooding, or a missile attack – is enough to disrupt the learning process by interrupting the continuity of learning and forcing changes in regular academic schedules.

Additionally, apart from destruction, war forces a massive shift to online learning, and the effect of this on educational performance was extensively studied during the COVID-19 pandemic. The literature on loss in learning during the pandemic provides concerns about the efficiency of online studying due to sudden urgencies. For instance, Orlov et al. (2021) assessed the effect of a sudden switch to online learning on students' performance in intermediate economics courses taught at universities in the United States. The results suggest a decline in scores by 0.2 standard deviations. Authors also find heterogeneity in the educational outcomes across the courses, and the course instructors' survey suggests that the negative effect was mitigated by previous experience in online teaching and exploitation of teaching methods that encourage students to active engagement (Orlov et al. 2021). This paper implies that switching to online learning alone has a significant negative effect on students' academic outcomes, apart from a decline in cognitive functions and psychological stress caused by war.

Despite the extensive evidence regarding disruption of education, there remains a significant gap in the related literature – the impact of armed conflict on students that are currently enrolled in higher education. While the negative effects of war on schooling are well studied (e.g., Leon (2012), Swee (2015), and Brück et al. (2019)), research on university students specifically affected by war is scarce. Although literature on the impact of exogenous shocks on university students' performance exists, it mainly limited to natural disasters (Di Pietro 2018) or COVID-19 pandemic (Orlov et al. 2021). However, the impact of war is fundamentally different from these events. To us, this seems to be a critical oversight, as university students represent a demographic with unique economic pressure that does not apply to younger children, such as labor market entry and conscription risk. Most research that comes close to higher education still focuses on the point of transition from school to university – high school exit exams (e.g., Brück et al. (2019), Thamtanajit (2020)). While studies confirm that conflict undermines access to higher education, there is no insight into the academic performance of students after they have enrolled in university.

One of the few studies that focuses on the academic performance of higher education students is by Grueso (2023), in which the author examined the impact of a sharp increase in homicide rate on college test scores in Colombia. According to his research, a 10% rise in homicide rate is linked to a decline in test scores of approximately 0.07 standard deviations, with the negative effect being larger for female and low-income students (Grueso 2023). Grueso (2023) also observed that online students experienced an even larger negative impact (0.14 standard deviations) compared to offline students, which is an important finding to consider in the context of war in Ukraine. In turn, this suggests that the flexibility offered by online learning does not necessarily mitigate the negative impact of being exposed to violence. In fact, when students are under persistent stress, the lack of a physical learning environment might exacerbate the decline in academic performance.

However, findings provided by Grueso (2023) do not capture the magnitude of disruption caused by a full-scale war, where the impact on the educational system is systematic and multifaceted rather than localized. The full-scale russian invasion of Ukraine provides a unique yet tragic context to analyze how war disrupts education through the mechanisms described above. This invasion represents a massive, disproportional country-wide shock to the educational system. According to UNESCO (2023), over 3,200 educational institutions in Ukraine have been damaged, and more than 400 have been completely destroyed as of January 2023. In turn, the physical destruction of educational infrastructure has led to a shift in the learning process – over 1.8 million students were studying online as of the same date (UNESCO 2023). Yet online schooling is different across the country – whether in its methods or overall quality. Specifically, during the COVID-19 pandemic, students in rural area tend to struggle more with a deficit of devices necessary for online education and bad internet connection compared to urban areas (DIF 2021). The possible assumption that pre-war experience with remote learning could have provided a sufficient basis for online education in Ukraine is challenged by recent empirical evidence. Karakai and Hatsko (2025) indicate that technical deficits observed during the pandemic were not resolved, and reliance on remote learning may have deepened educational inequalities. Specifically, access to digital devices remains a critical bottleneck, and "digital inequality is one of the key drivers of differences in learning outcomes" (Karakai and Hatsko 2025). Furthermore, online learning is now disrupted even more by targeted attacks on energy infrastructure, which negatively affect the network and electricity supply required for remote learning. This creates an environment where students face the double burden of 1) reliance on online education which, according to Grueso (2023), makes academic outcomes more sensitive to exposure to violence; and 2) continual disruptions of remote learning by targeted attacks on energy infrastructure.

To address this complex environment, this thesis builds upon the methodological framework established by Mykhailyshyna (2025). Her recent working paper provides a pioneering approach to measuring the causal impact of the russian invasion on educational outcomes of children exiting school by exploiting cross-country differences using PISA data and within-country variations in war intensity. Crucially, Mykhailyshyna (2025) identifies a specific mechanism of disruption. The decline in performance is caused by unpredictability of air raid alerts, which lead to acute stress and sleep deprivations if they occur on the night before an exam. The cumulative effect of missed classes due to the impact of dimensions of war – whether it is due to the occurrence of air raid alerts during the study time or the destruction of school – happened to be statistically insignificant (Mykhailyshyna 2025). While Mykhailyshyna (2025) establishes the baseline impact on secondary school students and high school graduates, this thesis extends the analysis specifically to the university graduates (bachelors) willing to apply for MA programs. To make the attribution structure of this empirical strategy explicit, the short-run design which isolates the immediate impact of air raid alerts on the night before an exam directly follows the framework established by Mykhailyshyna (2025). However, building upon it, the original methodological and empirical contributions of this thesis are: (a) extending the scope of analysis to higher education by examining UEE and UPE standardized exams; (b) introducing a Two-Stage TWFE Logistic regression model to assess how war exposure drives non-random absenteeism and conditional failure; and (c) applying a continuous TWFE DID framework to estimate the academic penalty of cumulative exposure to war.

This study fills the gap in the literature regarding higher education disruption in Ukraine due to the russian invasion.

Chapter 3

METHODOLOGY

3.1 Research question and hypotheses

The primary objective of this study is to answer: **'How do different dimensions of war – air raid alerts, missile attacks, and fatalities – affect scores on standardized exams of undergraduates willing to apply for MA?'**. To address this question, we test the hypotheses presented in Table 1.

Table 1. Research Hypotheses

Hypothesis	Description
H1	Dimensions of war – air raid alerts, missile attacks, and fatalities – reduce undergraduates' performance on standardized exams.
H2	As war goes on, students adapt to "new reality", and thus cumulative dimensions of war have smaller effect on their performance.
H3	While students may adapt to cumulative effects of war (i.e. stress, missing classes, and remote learning), an immediate impact (e.g., air raid alert or missile attack on the night before an exam) remains a significant determinant of decline in academic performance.
H4	Effects of war on undergraduates' performance are not uniform across individuals in different demographic groups. In particular, the negative impact of dimensions of war is expected to be different for females of all ages, as well as males aged 24 and older, compared to males younger than 24
H5	Dimensions of war increase the probability of student being absent in case of an immediate impact, and the probability of student failing if they were not absent.

3.2 Empirical strategy

To measure the causal impact of the russian invasion on university students' academic performance, the empirical strategy closely follows one D. Mykhailyshyna used in her paper, which focuses on the effect of war on scores of standardized exams which are taken during and after school (PISA, NMT) (Mykhailyshyna 2025). Specifically, while the short-run night-before-alert model is directly adopted from Mykhailyshyna (2025), this study introduces three novel methodological extensions: (a) applying the framework to higher-education UEE/UPE exams, (b) utilizing a two-stage absenteeism and conditional failure logistic regression, and (c) estimating cumulative exposure via a TWFE Continuous DID framework. Because Ukraine is entirely affected by the war, we do not have a perfect 'control group' of students with zero war exposure. Instead, all regions experience the war, but at different levels of intensity. This provides the rationale to employing a 'continuous treatment' approach. While using this approach, this study acknowledges the methodological warnings that were raised regarding its estimands interpretation (Callaway et al. 2024). It was demonstrated that TWFE estimands with continuous treatments can be difficult to interpret and do not yield a causal parameter right away. The validity of the TWFE Continuous DID framework in this thesis relies on 'strong parallel trends' assumption, which requires that average academic performance would have evolved identically across all regions regardless of the specific 'treatment dosage' (which is cumulative exposure to war in our case) the students experienced. Under this assumption, the linear TWFE estimand provides a quantified effect of academic penalty related to exposure to war.

The analysis is divided into three specific models to capture different types of war effects. In all models, standard errors are clustered at the physical testing region level because students taking exams in the same region share the same local conditions (Bertrand et al. 2004). The number of clusters is the number

of regions used for this study – 22. It is important to acknowledge a methodological limitation regarding this cluster count, since a small number of clusters (typical "rull of thumb" suggests fewer than 30-40) can cause standard cluster-robust standard errors to over-reject the null hypothesis (Cameron et al. 2008). While applying a wild cluster bootstrap is a standard workaround for the issue of small cluster count, this study relies on standard cluster-robust standard error. Hence, the reported statistical significance should be interpreted with caution, particularly the borderline p-values, as standard errors might be slightly downward biased.

3.2.1 TWFE Continuous DID design

The first model measures how the total, accumulated stress and destruction of the war affect test scores over time. We use both time- and region-varying cumulative shocks, such as the total hours of air raid alerts or total number of kinetic events that occurred during the academic year prior to taking the exam. The model specification is:

$$Score_{irt} = \beta_0 + \beta_1 CumulativeShock_{rt} + \gamma \boldsymbol{X}_i + \theta_t + \eta_r + \epsilon_{irt} \quad \textbf{(1)}$$

where $Score_{irt}$ is an exam raw score for a specific subject for individual i in region r, and at year t, restricted to students who attended the exam and got the raw score; $CumulativeShock_{rt}$ represents one of three metrics of cumulative war exposure – air raid alerts, kinetic events, or fatalities – in region r and at year t (the precise construction and sources are detailed in Chapter 4); $\boldsymbol{X}_i$ is a vector of individual i characteristics, specifically age and gender; θ_t and η_r are year and region fixed effects, respectively. The coefficient β_1 shows the average marginal penalty that cumulative exposure to war has on the student's

score. Equation 1 is estimated independently for each of the three core subjects (more on choice of core subjects in Chapter 4).

It is important to note on the choice of dependent variable. This study estimates the DID models using raw exam scores (before they are 'translated' to a 100-200 scale) to account for students who failed (and got 0 points in such scale), and thus to not exclude the variance in scores which is crucial to find causal effect. We do not standardize (z-scoring) scores within each exam-year cell since pooling treated and untreated students to calculate the mean and standard deviation is a biased approach. If the treatment (in our case, war exposure) systematically shifts students' scores, the post-treatment mean and variance are themselves affected by the shock. Normalizing against these shifting parameters artificially 'bakes in' the treatment effect, pulling the DID estimator toward zero (Kling et al. 2007). By utilizing raw scores, the Two-Way Fixed Effects (TWFE) estimator avoids endogenous normalization and remains unbiased. Because the models are estimated independently for each of the three core subjects, the region and year fixed effects (η and θ) fully absorb any time-invariant regional disparities and aggregate baseline differences in exam difficulty across years. This specification isolates the true impact of war exposure on academic performance.

3.2.2 Event-Study design

To make sure the DID model is valid, we must test the 'parallel trends' assumption. We also want to see how the war's impact changes year by year. To do this, we use an event-study model with the main variable of interest being an interaction between fixed measure of total regional war intensity during the academic years ($TotalShock_r$) and the year dummy:

$$Score_{irt} = \beta_0 + \sum_{k \neq 2021} \delta_k \left(TotalShock_r \times \mathrm{I}(Year = k) \right) + \gamma \boldsymbol{X}_i + \theta_t + \eta_r + \epsilon_{irt} \quad \textbf{(2)}$$

where δ_k captures the effect of 'future destiny' on the score of the exam, year 2021 being the baseline reference period ($k = 2021$ is omitted). Ideally, the pre-treatment coefficient δ_{2019} is expected to be close to zero and statistically insignificant, while the post-treatment coefficients $\delta_{2022,2023,2024}$ are expected to exhibit a statistically significant negative effect. However, it is vital to consider that since data for 2020 is not available, this event-study design relies on a single pre-treatment period (which is 2019). Hence, while a statistically insignificant pre-treatment coefficient is consistent with the parallel trend assumption, this test is underpowered. It cannot definitively rule out non-linear pre-treatment trends, thus this validation should be interpreted with caution due to the data limitations.

3.2.3 TWFE OLS design

To isolate the acute psychological and physiological disruptions (e.g., sleep deprivation) caused by the immediate effects of dimensions of war on the night before the exam, the model is specified using instantaneous shock variables:

$$Score_{irt} = \beta_0 + \beta_1 InstantShock_{rt} + \beta_2 \boldsymbol{X}_i + \theta_t + \eta_r + \epsilon_{irt} \qquad \textbf{(3)}$$

where $InstantShock_{rt}$ represents duration of alert on the night before the exam, events, or fatalities on the day before the exam. This specification identifies the localized short-term penalty of immediate effect of exposure to war, holding aggregate cumulative exposure constant through the regional and yearly fixed effects.

3.2.4 Two-Stage TWFE Logistic regression

To investigate potential self-selection bias and assess the impact of exposure to war on failure[2] of the exam or decision of student to show up for the exam, a fixed-effects logistic regression is utilized for two separate binary outcomes. This approach documents the existence of non-random attrition, providing evidence to explain any unexpected positive coefficients that may be observed in TWFE Continuous DID and TWFE OLS estimations. First, we model the student's initial decision to appear at the testing center. Second, for those who attended (1 = Attended, 0 = Absent), it models the probability of successful exam completion (1 = Passed, 0 = Failed). The specification is presented below:

$$\ln\left(\frac{P(Y_{irt}=1)}{1-P(Y_{irt}=1)}\right) = \beta_0 + \beta_1 InstantShock_{rt} + \beta_2 \boldsymbol{X}_i + \theta_t + \eta_r \quad \textbf{(4)}$$

where Y_{irt} represents binary indicators for exam absence or failure. This model estimated the log-likelihood of having no valid score due to war shocks. This model is specifically designed to account for non-random 'absenteeism' or 'failure', which rates significantly changed during the wartime years, as will be shown in Chapter 4.

To explain the implementation of these models, the following chapter details the construction of the analytical sample and the specific war-exposure metrics used.

[2] Failure is defined as a student not meeting the minimum threshold of correct answers required to receive a standardized score on the official 100-200 scale. Hence, a student may achieve a positive raw score (and thus be included in the TWFE Continuous DID and TWFE OLS models) but still be classified as 'failed' in this binary specification if their raw score falls below the passing threshold.

Chapter 4

DATA DESCRIPTION

4.1 Sources of data

Exams data (UCEQA)

For undergraduate standardized exams scores, we use data on UEE and UPE participants and their scores posted on the UCEQA[3] (Ukrainian Center for Educational Quality Assessment) website for 2019, and 2021-2024 (including data on MCT – Master's Comprehensive Test – and MTAC – Master's Test of Academic Competence – conducted in 2022 instead of UEE and UPE). For 2020, there is no student-level data available, and only statistical distributions are present. However, data for 2020 would most likely reflect the effects of the COVID-19 pandemic on students' performance, changing the baseline scores for comparison with war periods, so this data would not be included in the sample anyway.

The datasets for each year are anonymized and have data on the individual level. For each individual, there is data on year of birth (which lets me calculate age at the time of participation), gender, scores on each subject the person choose for the exam, and names of region and either city or village where the exam was taken. Observation counts range from ~36.8 thousand in 2022 to ~204 thousand in 2024. All numbers of observations are mentioned considering cleaned data (by cleaning we mean omitting rows corresponding to students who either or both 1) did not attend the exam; 2) had their results annulled; 3)

[3] The data can be found here: https://testportal.gov.ua/yevi-yefvv-zvity-dani/ (clicking on each year lets you further choose "depersonalized results database", which is not present for 2020 and years before 2019 as mentioned above).

took the exam outside of Ukraine), and for all subjects included in each year's exam.

War exposure data

For air alerts (duration and timing), the data was taken from the dataset prepared by V. Klymenko[4] (Klymenko 2025). An important limitation to mention here is that official data is present only starting from the 15th of March 2022 (the first day when siren records occur); however, the gap between the 24th of February and the 15th of March in 2022 can be filled with unofficial data available from the same source.

For fatalities, we use daily data from the Armed Conflict Location and Event Data Project (ACLED[5]). There, kinetic events are classified into different types, war events being a separate category. Additionally, each event has data on the location of its occurrence, and number of fatalities attributed to this specific event.

Regarding the kinetic events and fatalities, we want to highlight that those two dimensions of war are not correlated as it may seem. To demonstrate this, we report a scatterplot in Figure 1, plotting kinetic events and fatalities that happened on the same day. As we see, most attacks resulted in low fatalities, and the days with the most fatalities did not necessarily have the most attacks.

[4] The data can be found here: https://github.com/Vadimkin/ukrainian-air-raid-sirens-dataset/tree/main/datasets

[5] The data can be found here: https://acleddata.com/conflict-data/download-data-files

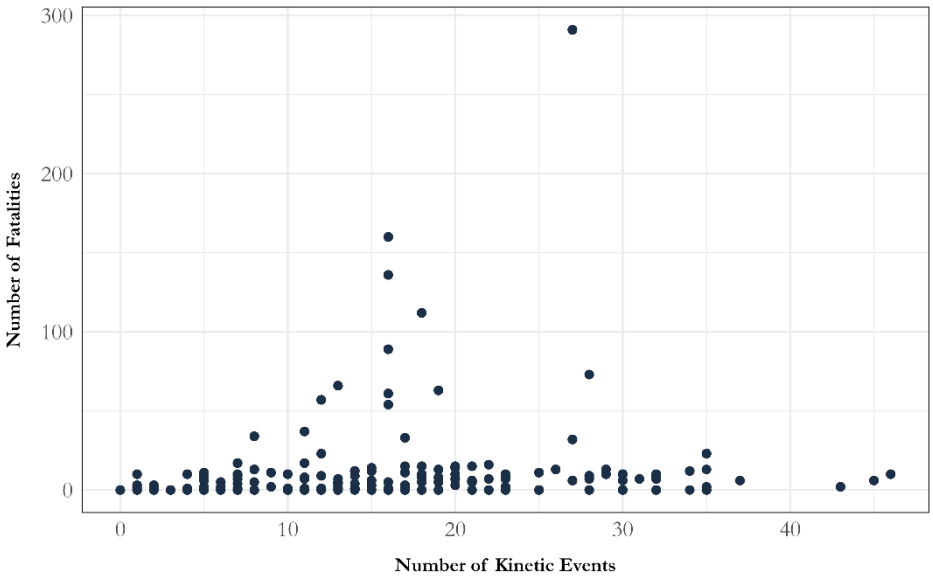


Figure 1. Daily Kinetic Events vs. Daily Fatalities

4.2 Sample composition

The initial dataset, obtained by merging standardized exam records from the UCEQA for the years 2019 and 2021-2024 with war data from V. Klymenko and ACLED, constitutes a baseline sample of 603,297 student observations. As a preliminary data quality check, the sample was evaluated for missing student identification codes. The raw state records demonstrated completeness at this stage, with zero observations dropped due to missing identifiers, maintaining the full sample of 603,297 for further geographic and exposure filtering.

To avoid the spatial complexities introduced by 'universities in exile' – institutions legally registered in frontline or heavily bombarded regions (e.g., Kharkiv, Zaporizhzhia) but operating elsewhere – exposure metrics were mapped exclusively to the student's physical examination center (PTReg) rather than their university's legal address (Uni_Reg). This approach guarantees spatial consistency since it ensures that the fixed effects of regions where a student took an exam align with the exposure to war that the student, as we assume,

experienced. At the same time, it introduces a critical measurement-error concern. The physical testing regions are themselves endogenous since students who originate from heavily attacked regions may have travelled westward to safer regions and took exam there. If students previously exposed to war the most systematically select safer testing regions, then PTReg-based exposure mapping understates the true exposure experienced by these individuals. This measurement error biases the estimates of war effect on scores toward zero (attenuation bias). Consequently, some of statistically insignificant and/or close to zero estimated effects may be caused by this measurement error rather then reflect actual academic resilience.

After the baseline integrity check, a geographic filter was applied to account for structural breaks in data collection caused by military occupation. University registration data showed that students were enrolled in institutions originating from Donetsk, Luhansk, and Kherson regions throughout 2022-2024 (Appendix A, Table 15). However, this is likely because 'universities in exile' were legally moved to government-controlled regions. Diagnostic cross-tabulations of physical examination centers confirmed severe discontinuity; domestic testing locations in Donetsk, Luhansk, Kherson, and the Crimea (including Sevastopol) were either completely absent or statistically negligible after 2021 (Appendix A, Table 16). To maintain consistency in geographical distribution of metrics of exposure to war, any observations associated with these regions were excluded. This restriction removed 21,298 observations, leading to 581,999 students in the sample.

The sample was further cleared to ensure valid treatment assignment by removing students who completed their standardized exams outside of Ukraine. Diagnostic statistics revealed a sharp increase in international testing starting in 2023, which reflects the displacement of Ukrainian students abroad. Since primary independent variables of this study (daily air raid alert durations, events, and fatalities) are spatially bound to the regions of Ukraine, students

that tested abroad did not experience any immediate war shocks during the period of examination. Excluding these individuals dropped an additional 5,532 observations, leading to a sample of 576,467 students.

The dataset was further filtered to get rid of noise caused by people not showing up and canceled exam attempts. In the exams studied, large number of students who registered for the test do not attend it. This trend got much worse during the 2022-2024 because of the war. To make sure that the analysis only uses valid observations. an 'attendance requirement' was applied, filtering out 101,934 more students lacking a recorded 'Passed' or 'Failed' status or any valid numeric score. This ensures that the primary sample reflects the effects war has on academic performance among students who did the exam.

While absent students were filtered out of the primary sample, absenteeism and failure are most likely not entirely random. To capture the war impact on decision to attend an exam and likelihood of passing it, another dataset with two separate binary variables was generated for supplementary analysis by using the fixed-effect logistic regression specified in Equation 4. This dataset contains only missed attempts (absents) and failures.

Another issue addressed was the high risk of self-selection bias in minority foreign languages exams. Students can choose between English, German, French, and Spanish for the foreign language part of the Master's entrance exams. Because English instruction is standard in the Ukrainian primary and secondary education systems, students opting for minority languages likely possess distinct educational backgrounds that introduce self-selection bias. Rather than dropping these individuals entirely – which would unnecessarily discard their valid performances on other standardized subjects such as the General Competence or Law exams – a precise filtering was applied. All students only having scores for French, German, and Spanish examinations were filtered out from the dataset. This focused method resulted in loss of 5,239 observations.

As a final cleaning step, observations missing fundamental demographic identifiers (Age and Gender) or containing biologically improbable age values likely due to data-entry errors (which we define as under 18 or over 70 years old) were removed.

The comprehensive step-by-step cleaning process of student-level data is summarized in Table 2. Following the steps of methodological restrictions mentioned above, the baseline dataset of 603,297 observations shrinked to an analytical sample of 469,291 students.

Table 2. Student-Level Sample Cleaning Process

Step	Description	Remaining (N)	Dropped (N)	Remaining (%)
0	Baseline: Total starting observations from merged UCEQA data	603,297	0	100.00%
1	Removed rows with missing Student IDs	603,297	0	100.00%
2	Excluded discontinuous regions (Donetsk, Luhansk, Kherson, Crimea, Sevastopol)	581,999	21,298	96.47%
3	Excluded students testing Abroad (no domestic war exposure)	576,467	5,532	95.55%
4	Excluded students lacking valid statuses (Absent/Annulled/NA) AND lacking any valid raw scores	474,533	101,934	78.66%
5	Omitted non-English language scores and dropped students with no remaining valid subjects	469,294	5,239	77.79%
6	Removed missing demographic variables and extreme age outliers (Age < 18 or Age > 70)	469,291	3	77.79%

4.3 Descriptive statistics

The descriptive statistics for analytical sample are summarized in Table 3, which provides the mean, standard deviation, and range for the main dependent, demographic, and war-exposure variables. The descriptive statistics are calculated at the student-exam observation level, representing a cross-sectional panel that covers both peacetime baseline years (2019 and 2021) and the wartime period (2022-2024).

Table 3. Descriptive Statistics of the Analytical Sample

Variable	**Mean**	**SD**	**Min**	**Max**
Standardized Exam Score (100-200 scale)	143.75	21.14	100.00	200.00
Raw Exam Score (number of correct answers)	22.68	17.54	0.00	128.00
Female (Indicator: 1=Yes, 0=No)	0.36	0.48	0.00	1.00
Age (Years)	25.59	6.88	18.00	70.00
Night-time Air Raid Alerts (Hours)	1.84	3.95	0.00	15.00
Regional Daily Kinetic Events	1.77	5.41	0.00	46.00
Regional Daily Fatalities	1.67	11.20	0.00	291.00
Cumulative Air Raid Alerts (Hours)	653.94	1,215.20	0.00	6,608.82
Cumulative Regional Kinetic Events	498.40	1,381.80	0.00	7,696.00
Cumulative Regional Fatalities	299.40	1,072.07	0.00	6,943.00

The main dependent variable, the raw exam score, which represents the absolute number of correct answers, has a mean of 22.68 and ranges from minimum 0 to maximum 128. Hence, it exhibits a large standard deviation of 17.54. This high variance is mainly driven by changes to the tests themselves across different years and subjects. For instance, the maximum possible raw

point for Law exams fluctuated drastically during the war, firstly shrinking in 2022, and then expanding during 2023-2024. While these varying maximums create noise in raw score data, the TWFE DID is designed to mathematically absorb these year-to-year structural changes in number of questions, allowing the unbiased causal estimation. Furthermore, models will be run on each exam separately, so that the difference in possible raw score range between the exams will not affect the results.

We present the standardized exam score mainly for informational and descriptive purposes, since it provides a scaled 'transformed' score metric, that makes it easier to visually track across years. It is important to mention that descriptive statistics for the standardized exam score are calculated based on the subset of students who successfully passed their exams, which is why minimum value is 100 and not 0 (since all failures results in 0 'translated' score, including them in calculation of descriptive statistics would drastically corrupt the calculations).

The sample is characterized by a significant gender shift over time. While the overall female share of the sample is 35%, this is heavily affected by the 2024 cohort, where female participation rate dropped to 24.5%. This reflects a sharp increase in adult men registering for the exams during the mobilization period. The average age of students is 25.59 years, ranging from 18 to 70 years, capturing both traditionally aged and older students.

The explanatory variables for war exposure are categorized into immediate daily shocks and cumulative exposure during the academic year. *Night-time Air Raid Alerts* variable represents the total duration of air raid alerts at the physical testing region between 18:00 on the day prior to the exam and 09:00 on the day of the exam (following the D. Mykhailyshyna's methodology to define the 'night' (Mykhailyshyna 2025)). The average exposure is 1.84 hours, reaching a maximum of 15.00 hours, indicating cases where testing regions were under continuous alert for the entire duration of the defined night window. Regional

daily kinetic events (explosions, missile attacks and remote violence) average 1.77 per day, with a maximum of 46 events recorded in a single region during the most intensive offensive operations. Regional daily fatalities follow a similar distribution, with an average of 1.67 and a peak of 291 reported casualties in a single day. To account for the long-term psychological and educational toll of the conflict, cumulative variables track exposure from the start of the academic year. The most heavily affected students lived in regions with over 6,600 hours of air raid alerts and nearly 7,700 kinetic events, mostly concentrated in frontline or close to the frontline regions like Kharkiv and Sumy regions. The variation in these metrics (the presence of zero-exposure observations in the pre-war years in particular) is necessary to identify the impact of conflict intensity on students' academic performance.

Figure 6 (Appendix B) illustrates how the geographic distribution of exam participants changed from 2019 to 2024. During the pre-war baseline years of 2019 and 2021, test-takers were primarily concentrated in traditional major educational centers, such as Kyiv, Kharkiv, Lviv, and Odesa. When the full-scale invasion began in 2022, there was a visible westward shift in the data, indicated by increased number of exam participants in Lviv. Participation dropped significantly in the eastern and southern regions (most noticeably in Kharkiv, Mykolaiv and Odesa regions) due to immediate threat, while western regions maintained higher numbers as internally displaced students moved there. Through 2023-2024 the geographical distribution changes completely: the 2024 map shows a massive surge in total number of exams taken, forcing the color scale's maximum to jump to 125,000 to capture the volume. This increase leads us to the conclusion of the presence of a 'mobilization effect'. Because enrollment in university grants a legal deferment from army, taking these exams became a useful option for those men who want to avoid deferment, causing participant numbers to increase.

It is crucial to mention that this documented westward shift provides an evidence of a major measurement limitation raised earlier regarding the endogeneity in physical testing regions. As internally displaced students from heavily exposed to war eastern and southern regions (e.g., Kharkiv, Sumy, Mykolaiv) migrated to safer western regions (e.g., Lviv, Ivano-Frankivsk), observations related to them are tied to the exposure intensity of safer physical testing region. That is, the war exposure for these displaced students is systematically mismeasured downward. As mentioned before, this introduces the attenuation bias in TWFE Continuous DID and TWFE OLS estimates. It is a crucial context for interpreting the estimated effects, as it may partially explain why students from traditional cohorts may appear 'resilient'.

While Figure 6 shows where the exams were taken, Figure 7 (Appendix B) illustrates the intensity of the war students experienced in different regions. To capture different types of exposure, the maps track three specific war exposure measures in years 2022-2024: total hours of air raid alerts, the number of kinetic events, and fatalities. Maps on geographical distribution of air raid alerts show that it acted as a nationwide psychological burden. In contrast, the bottom two rows of maps show that direct physical violence, which is measured by events and fatalities, stayed highly concentrated along the eastern and southern front lines. Regions like Kharkiv, Zaporizhzhia, and Dnipropetrovsk experienced most of the physical destruction throughout all three years, while all other regions were relatively less affected by direct hits. This difference is crucial for the study. Because air raid alerts became widespread while physical violence not, we can use this spatial variation in our regression models. It allows us to separate the general stress of living under constant air threats from the specific effects of localized physical destruction.

A visual analysis of age distributions by gender and year revealed a severe structural anomaly in 2024: a massive increase in older male registrants (Figure 2).

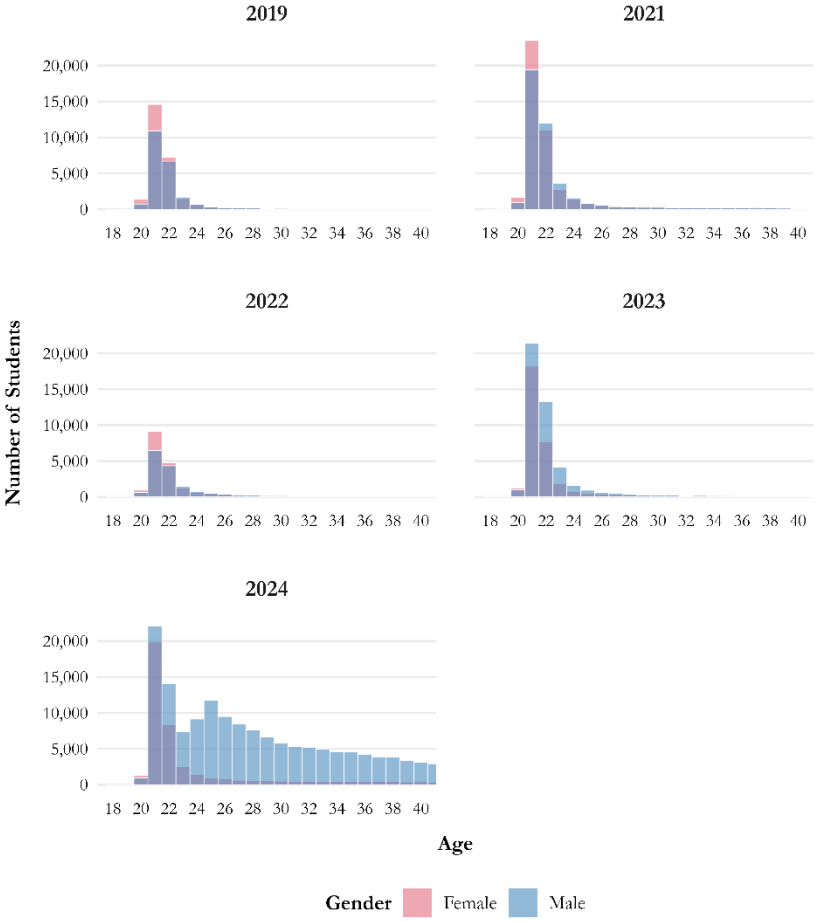


Figure 2. Age Distribution of Master's Exam Registrants by Gender (2019, 2021-2024)

Note: The x-axis is truncated at 40 years for demonstration purposes. The oldest participant in the sample is 70 years old.

The next step is to define which specific exams can be used for the study. The study requires a consistent baseline to correctly estimate the impact of the war using a Difference-in-Differences (DID) design. This means we can only analyze exams that were administered both before and during the full-scale invasion. Table 4 outlines the availability of standardized exams across years 2019 and 2021-2024.

Table 4. Standardized Exams by Year (2019, 2021-2024)

Subject / Exam Type	2019	2021	2022	2023	2024
General Competence	✔	✔	✔	✔	✔
Foreign Language (English)	✔	✔	✔	✔	✔
Law	✔	✔	✔	✔	✔
Economics	✗	✗	✗	✔	✔
Management	✗	✗	✗	✔	✔
Finance	✗	✗	✗	✔	✔
Pedagogy	✗	✗	✗	✗	✔
Psychology	✗	✗	✗	✔	✔
Information Technology	✗	✗	✗	✗	✔
Political Science	✗	✗	✗	✔	✔

Note: (✔) indicates the exam was administered. (✗) indicates the exam was not administered. The framed rows denote the three core subjects suitable for the Difference-in-Differences (DID) design as they are present in both pre-war and wartime periods.

The Ministry of Education significantly expanded the testing system by introducing several new specialized exams, such as Economics, Management, and Psychology, in years 2023-2024. However, because these subjects did not exist during the 2019 and 2021 years, they lack a pre-war baseline. Hence, to satisfy the 'parallel trends' assumption required for the Event-Study and DID designs described in Chapter 3, the sample is restricted to the three core subjects that are framed in Table 4: General Competence, Foreign Language (English), and Law. They provide the uninterrupted data necessary to accurately isolate the impact of the war from changes in test options, since these three exams existed across all five years.

Panel A of Figure 3 shows the distribution of standardized scores (on the 100-200 scale) for the three core subjects. This ridge plot reveals how the structural shocks of the war completely altered student performance by separating the data by years. We use standardized scores on the 100-200 scale for this visual

because we believe this universally scaled metric allows for coherent visual tracking of trends across years and subjects, such as the variance collapse in 2022 and the demographic shifts in 2024, which would otherwise be affected by the varying maximum point totals of raw exams. However, as detailed in Chapter 3, the formal DID estimations strictly uses the raw exam scores. While the 100-200 scale is optimal for visual comparability, the non-manipulated raw scores are required in the regression framework to avoid normalization bias as well as to allow the two-way fixed effects to properly isolate the true causal impact of the war.

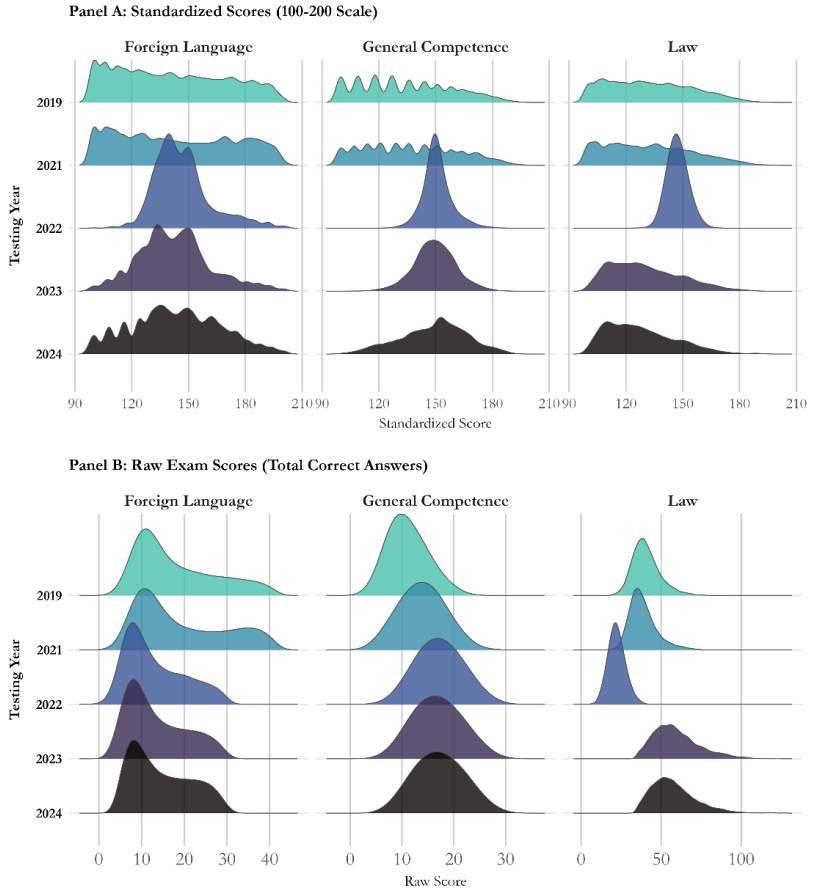


Figure 3. Yearly Evolution of Score Distributions

Note: Visual spikes in the density curves reflect the discrete scaling of the standardized scores system.

During the peacetime baseline years (2019 and 2021), the score distributions for all three subjects are wide and relatively flat, representing a normal, competitive academic environment where scores are spread across the entire 100-200 spectrum. However, the beginning of the full-scale invasion in 2022 triggered a massive 'variance collapse'. Across the three main exams the 2022 distributions spike sharply in the middle. Clustering around the median likely reflects both the emergency of the exams and a potential softening of grading to accommodate students testing under unprecedented trauma. After that, the shapes of the distributions change again, clearly reflecting the demographic shifts discussed earlier. For instance, the curves for Foreign Language and Law become significantly wider and flatter in 2024, and more importantly the lower end of the score scale (between 100 and 140) rises noticeably. This shift perfectly captures the 'mobilization effect': large number of older men entered the exams only to secure draft deferments, and their lack of recent academic preparation pulled the overall score distributions downward.

To still provide full transparency into the underlying data structure, Panel B of Figure 3 presents the distributions of the raw scores (total correct answers) for the same period. While Panel A is optimal for observing behavioral trends, Panel B visually exposes the severe structural volatility of the exams themselves during the wartime period, which were mentioned above. Because the maximum possible raw points varied year-to-year as the Ministry of Education adapted the testing format, the raw score distributions shift across the horizontal axes. This is most noticeable in Law exam: in 2022, the distribution shifts to the left, reflecting shortening of the exam due to emergency, and after this it expands massively to the right in years 2023-2024 as the exam was lengthened and thus once again restructured. If viewed in isolation, these raw score shifts could be misinterpreted as wild fluctuations in student legal aptitude. However, exposing this volatility explicitly justifies the

methodological design detailed in Chapter 3. By running the TWFE regressions independently by subject using these raw scores, year fixed effects absorb these structural changes in exams.

Figure 4 shows the absenteeism and failure rates among registered students for 2019 and 2021-2024. Both rates were relatively low in pre-war years. However, the full-scale invasion clearly disrupted these trends, and the failure rate dropped to zero in 2022. This was likely due to emergency changes in the exams, such as conducting a one-off Master's Comprehensive Test (MCT), and perhaps a general softening of grading policy to support students taking exams under severe trauma. In contrast, all war-period years show a noticeable increase in absenteeism rate.



Figure 4. Absenteeism and Failure Rates for Exams

Since these extreme shifts are driven by war rather than normal student behavior, only analyzing the test scores of students who successfully passed would ignore a huge part of the story.

Chapter 5

EMPIRICAL RESULTS

This chapter presents the findings from the TWFE DID and Event-Study estimations as described in Chapter 3.

5.1 TWFE Continuous DID

5.1.1 Assessment of baseline equivalence

Before estimating the causal effects of dimensions of war on the standardized exams scores, we find it necessary to validate the baseline equivalence assumption.

Table 5. Baseline Equivalence Check

Exam Subject	Future War Exposure	Pearson Correlation (r)	p-value
General Competence	Future Air Raid Alerts	-0.015	0.0154
	Future Fatalities	0.064	0.0000
	Future Kinetic Events	0.052	0.0000
Foreign Language	Future Air Raid Alerts	-0.054	0.0000
	Future Fatalities	-0.000	0.9510
	Future Kinetic Events	-0.018	0.0000
Law	Future Air Raid Alerts	-0.048	0.0000
	Future Fatalities	0.082	0.0000
	Future Kinetic Events	0.066	0.0000

Table 5 presents Pearson correlation coefficients between students' pre-war academic performance (2019 and 2021) and the eventual cumulative war exposure of their region. Due to the large sample size, standard errors approach zero, causing highly statistically significant p-values. However, the correlation coefficients for all three exposure metrics (alerts, events, and fatalities) are close to zero ($|r| \approx 0$). This confirms that regions destined for high war intensity in the future did not exhibit systematically different academic performance in pre-war years (2019, 2021), thus satisfying the baseline equivalence assumption required for the continuous DID design.

5.1.2 Parallel trends assumption

It is also necessary to visually and mathematically validate the parallel trends assumption. Because the causal impacts are estimated independently for each exam subject, the parallel trends assumption must be validated for each exam separately using the main dependent variable (raw scores) trends. Figure 5 illustrates trends of scores for the three main exams. Regions were grouped by *Future Air Raid Alerts* (a total duration of air raid alerts) into three groups (High, Moderate, and Low intensity) by simply splitting the vector of regions arranged by the *Future Air Raid Alerts* into three equally sized buckets. The reason *Future Air Raid Alerts* is used for such splitting is that its distribution among regions has the least skewness. Across all three exams, the average score for all groups evolved close to parallel between 2019 and 2021, especially regarding raw scores for General Competence and Law.

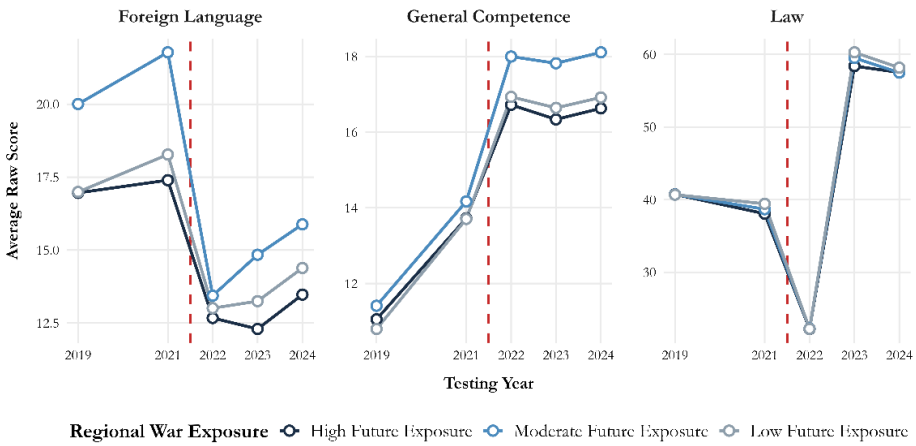


Figure 5. Parallel Trends: Raw Scores by Exam

However, a formal test is required to conclude that the parallel trends assumption is satisfied. To do so, we estimated Event-Study design models (Equation 2) independently for all three main exams, taking each of three cumulative war exposure variables separately. The comprehensive 3x3 grid of event-study plots alongside the estimation results are presented in Appendix C (Table 17 and Figure 8).

Considering coefficients for interaction terms between cumulative war exposure variables and the 2019 pre-treatment baseline, they are often statistically significant, which reveals a violation of the parallel trends assumption. Specifically, this assumption is violated completely for Foreign Language exam across all three war variables; at the same time, General Competence experiences such violation only in effects of future air raid alerts, and Law – in other two var variables, thus suggesting a partial violation. For instance, the estimated coefficient at interaction between 2019 and Future Kinetic Events for Law is 0.0001 and statistically significant at 1% level. Given the extreme maximum value for *Future Kinetic Events* (which is almost 17,500), this estimate holds its economic power, resulting in up to almost 2 more raw score on average for regions that are mostly exposed to war, ceteris paribus. Since 2021 is the omitted baseline reference year, such estimate for 2019

suggests that regions which were eventually affected by war the most during 2022-2024 already exhibited downward trend prior to the war. This could likely happen to COVID-19 pandemic and subsequent switch to online education, impact of which was likely still affecting scores in 2021. Thus, it introduces a downward bias into the TWFE Continuous DID estimates.

At the same time, such finding underlines the multi-model methodological framework used in this study. While some estimates for General Competence and Law are expected to be unbiased, we suspect that isolation of long-term impact of cumulative exposure to war might be limited. That is why we also rely on estimating the short-term impact on raw scores via Short-Run TWFE OLS (Equation 3) and on absenteeism and failure via TWFE Logistic regression (Equation 4) which we will use to isolate the immediate impact of exposure to war.

5.1.3 Estimation results

The TWFE Continuous DID estimates are presented in Table 6 below. The dataset has been split into three subsamples based on 1) gender; 2) age (defining 'traditional students' cohort for men strictly younger than 24, and all other). The estimated coefficients measure the marginal impact of accumulating one additional unit of localized war exposure on raw exam scores. Full estimation results are reported in Appendix D.

The results exhibit heterogeneity depending on the exam subject, nature of the war shock, and subsample. For Foreign Language, the cumulative effects of exposure to war are predominantly statistically insignificant for the aggregate sample as well as for the 'Women' and 'Men < 24' cohorts, suggesting no evidence of an accumulated penalty on academic scores for this subject.

Table 6. TWFE Continuous DID Estimates of Cumulative War Exposure Across All Subjects

Exposure	Aggregate	Women	Men < 24	Men 24+
Panel A: Foreign Language				
Cumulative Air Raid Alerts	0.00019 (0.00012)	0.00014 (0.00012)	0.00013 (0.00013)	0.00041*** (0.00009)
Cumulative Kinetic Events	0.00000 (0.00014)	-0.00004 (0.00015)	-0.00002 (0.00014)	0.00020*** (0.00006)
Cumulative Fatalities	-0.00010 (0.00013)	-0.00010 (0.00016)	-0.00005 (0.00013)	0.00010** (0.00005)
Panel B: General Competence				
Cumulative Air Raid Alerts	-0.00001 (0.00005)	0.00002 (0.00005)	-0.00001 (0.00004)	0.00011* (0.00006)
Cumulative Kinetic Events	-0.00011* (0.00005)	-0.00015** (0.00006)	-0.00014** (0.00005)	0.00003 (0.00004)
Cumulative Fatalities	-0.00015*** (0.00003)	-0.00016*** (0.00004)	-0.00012*** (0.00003)	-0.00001 (0.00003)
Panel C: Law				
Cumulative Air Raid Alerts	-0.00030* (0.00015)	-0.00026 (0.00018)	-0.00027* (0.00013)	-0.00049*** (0.00008)
Cumulative Kinetic Events	-0.00034*** (0.00008)	-0.00040*** (0.00008)	-0.00024*** (0.00008)	-0.00036** (0.00014)
Cumulative Fatalities	-0.00042*** (0.00014)	-0.00044*** (0.00015)	-0.00031* (0.00018)	-0.00042*** (0.00013)

Note: Continuous DID estimates (β). Standard errors clustered at the physical testing region level in parentheses.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

However, an unexpected anomaly emerges in 'Men 24+' subsample. There is a persistent positive highly statistically significant (at 1% level) effect across all cumulative war exposure variables. This implies that older men in regions that there more exposed to dimensions of war scored better relative to those in relatively safer regions. For instance, effect of 1 SD (1,215.2 hours) increase in air raid alerts total duration during the academic year prior to taking the exam translates into a raw score higher by 0.5 (0.07 SD[6]) on average, ceteris paribus.

[6] SD of Foreign Language exam raw score in 'Men 24+' subsample is 7.23.

Similarly, the positive effect of 1 SD (1,381.8) increase in kinetic events on raw score is 0.28 (0.04 SD[6]), and of 1 SD (1,072.07) increase in fatalities – 0.11 (0.02 SD[6]). For additional reference, a 1 additional raw score corresponds to 1 to 10 additional scores on a 100-200 scale[7] for Foreign Language exam. All these effects appear to be economically significant – especially the cumulative effects of air raid alerts – thus confirming that students in 'Men 24+' subsample likely have features that distinguish them from traditional student cohorts. These findings suggest rejection of **H1** and support **H2** since we do not see any statistically significant effect of dimensions of war on Foreign Language exam raw score for traditional student cohorts. At the same time, we find support for **H4** since the effects for 'Men 24+' subsample are persistently different from all other subsamples.

While there were no significant effects of dimensions of war on Foreign Language exam for traditional student cohorts, this is not the case for General Competence. Although cumulative effect of air raid alerts remains statistically insignificant, we notice a high statistical significancy for the other two war variables – kinetic events and fatalities. A 1 SD increase in kinetic events leads to expected raw score being lower by 0.21 (0.04 SD[8]) for 'Women' subsample and by 0.19 (0.04 SD[9]) for 'Men < 24' subsample on average, ceteris paribus. Similarly, the effects of 1 SD increase in fatalities are the expected raw score lower by 0.17 (0.04 SD[8]) for 'Women' and by 0.13 (0.03 SD[9]) for 'Men < 24'. A 1 additional raw score corresponds to 1 to 8 additional scores on 100-200 scale for General Competence exam. Importantly, 'Men 24+' subsample behaves differently from traditional student cohorts again, having effects of all dimensions of war statistically insignificant at any conventional significance level. In contrast, results for General Competence exam support **H1** and suggest rejection of **H2**, and similarly support **H4**.

[7] Correspondence between raw score and score on 100-200 scale is non-linear.
[8] SD of General Competence exam raw score in 'Women' subsample is 4.76.
[9] SD of General Competence exam raw score in 'Men < 24' subsample is 5.06.

In contrast to the general mandatory exams, the specialized Law exam raw scores are highly sensitive to cumulative war exposure across all subsamples, and almost across most of the dimensions of war. For instance, effects of 1 SD increase in number of kinetic events during the academic year prior to taking the exam results in expected raw score lower by (on average, holding all other effects fixed): 0.55 (0.03 SD[10]) for 'Women'; 0.33 (0.02 SD[11]) for 'Men < 24'; and 0.5 (0.03 SD[12]) for 'Men 24+'. However, these effects appear to be much less economically significant, since Law exam has more questions, thus score on 100-200 scale is less sensitive to changes in raw score. An analysis of the full regression outputs also reveals that Law exam is also different in demographic effects (effects of *Age*). While for Foreign Language and General Competence older participants are on average expected to have lower score compared to younger ones across all subsamples, the coefficient on *Age* for 'Men 24+' subsample is positive and highly statistically significant in Law exam. As we will see further, this positive effect of age holds for TWFE OLS model, too. This finding suggests that participants with more professional experience in law score more in this exam. Results for Law exam support **H1** but suggest rejection of **H2** and **H4**, since the effects are similar across all subsamples. As we can see, the exams seem to be heterogeneous per se, thus we cannot fully reject or accept any hypothesis among **H1**, **H2** and **H4** based on the results for all three main exams.

Forest plots in Figure 9 (Appendix D) visually complement the TWFE Continuous DID estimates discussed above. Forest plots were used to make it easy to compare the effects of dimensions of war across different exams and subsamples.

[10] SD of Law exam raw score in 'Women' subsample is 17.
[11] SD of Law exam raw score in 'Men < 24' subsample is 17.2.
[12] SD of Law exam raw score in 'Men 24+' subsample is 15.2.

5.2 TWFE OLS

5.2.1 TWFE OLS estimation results for 3 main exams

Next, we analyze the immediate impact of war exposure on students' raw scores. The pooled TWFE OLS estimates are presented in Table 7 below.

Table 7. TWFE OLS Estimates for 3 Main Exams

Exposure	**Aggregate**	**Women**	**Men < 24**	**Men 24+**
Panel A: Foreign Language				
Alerts (Night Before)	-0.0687*** (0.0221)	-0.0793*** (0.0216)	-0.0250 (0.0287)	-0.0825* (0.0402)
Kinetic Events (Day Before)	0.0526*** (0.0091)	0.0325 (0.0221)	0.0270* (0.0150)	0.0685*** (0.0226)
Fatalities (Day Before)	0.0012 (0.0010)	-0.0011 (0.0016)	0.0026** (0.0013)	0.0061*** (0.0007)
Panel B: General Competence				
Alerts (Night Before)	0.0109 (0.0272)	0.0026 (0.0194)	0.0207 (0.0226)	0.0771 (0.0574)
Kinetic Events (Day Before)	0.0027 (0.0088)	-0.0042 (0.0053)	-0.0122 (0.0089)	-0.0012 (0.0291)
Fatalities (Day Before)	0.0068*** (0.0015)	0.0104*** (0.0023)	0.0181*** (0.0032)	0.0073*** (0.0013)
Panel C: Law				
Alerts (Night Before)	-0.1202* (0.0595)	-0.1279* (0.0679)	-0.0907 (0.0658)	-0.1563** (0.0572)
Kinetic Events (Day Before)	0.0508 (0.0470)	0.0440 (0.0653)	0.0421 (0.0426)	0.0646* (0.0366)
Fatalities (Day Before)	0.0027 (0.0226)	0.0184 (0.0244)	-0.0132 (0.0242)	-0.0159 (0.0169)

Note: Pooled TWFE OLS estimates for the 2022-2024 wartime period. Standard errors clustered at the physical testing region level in parentheses.
Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Estimated coefficients on Kinetic Events and Fatalities suggest an immediate positive impact to students' scores on average. For instance, exposure to fatalities on the day prior to the General Competence exam has a highly

significant positive effect on scores across all subsamples. This suggests the presence of self-selection bias – exposure to this specific dimension of war causes low-performing students or those who are not confident in their potential exam results to refuse to take the exam (the analysis of absenteeism and failure probabilities is discussed below). This conclusion is also supported by other estimates for Foreign Language (coefficients for Kinetic Events and Fatalities for 'Men 24+' cohort). However, we get expected results for Air Raid Alerts effect on Foreign Language and Law exams, though not for all subsamples. For instance, for 'Women' subsample, increase of 1 SD in hours of air alert during the night before exam (between 18:00 on previous day, and 9:00 at the day of the exam, as defined above in Chapter 4) results in exam raw score lower by approximately 0.31 (0.04 SD[13]), which we consider an economically significant impact considering how this decrease translates to difference in score on 100-200 scale. This statistically significant effect on scores partially supports **H3**, demonstrating that while students may be resilient to cumulative exposure to war, the immediate impact remains highly statistically significant for Foreign Language exam. For 'Men 24+' students taking Law exam, the expected impact is almost twice as large in terms of raw scores, but this effect translates to a smaller impact on standardized score on 100-200 scale. At the same time, there is no statistically significant negative effect of air alerts on 'Men < 24' students across all 3 main exams. Hence, we find evidence of heterogenous effects of war on students of different demographic groups, which supports **H4**.

The same as for TWFE Continuous DID, full estimation results for TWFE OLS and forest plots are in Appendix E.

[13] SD of Foreign Language exam raw score in 'Women' subsample is 8.86.

5.2.2 TWFE OLS estimation results for other UPEs

Looking for additional evidence of immediate impact of war exposure on UPEs raw score results, we run TWFE OLS for other 7 subjects. The estimation results (both short and full) for other UPEs are presented in Appendix F, along with forest plots. Year fixed effects are dropped for Pedagogy and Information Technology subject since these two exams were conducted in 2024 only.

Estimation results further strengthen the highly heterogeneity of immediate impact of war exposure across various subjects. For example, students who take Political Science exam experience negative and highly statistically significant shock from air raid alerts on the night before the exam across all demographic cohorts. For this specific subject, 'Men 24+' students raw score being expected to be lower by 1.6 (0.11 SD[14]) for each additional hour of air alert on the night before the exam, which is a much more severe impact compared to what was found for Foreign Language or Law exams. Similarly, exposure to kinetic events and fatalities during the day before the exam has a statistically significant negative effect on raw score for quantitative and technical subjects like Finance and Information Technology. However, these effects are not consistent across all demographic groups. In general, the results for other UPEs confirm that the immediate impact of dimensions of war is not uniform. Rather, it varies sharply depending on the specific cognitive requirements of the exam and the demographic characteristics of students.

5.2.3 TWFE OLS robustness check: binary treatment

To verify the sensitivity of initial estimates to the measurement approach of immediate war exposure effects, we re-estimated the TWFE OLS models using dummy variants of war variables (treating any exposure as 1, and no exposure

[14] SD of Political Science exam raw score in 'Men 24+' subsample is 14.64.

as 0). Results of TWFE OLS with binary war variables are presented below in Table 8 as a robustness check for TWFE OLS with continuous instantaneous shocks.

Table 8. TWFE OLS Robustness Check: Binary Treatment

Exposure	Aggregate	Women	Men < 24	Men 24+
Panel A: Foreign Language				
Alerts Dummy (Night Before)	-0.2957 (0.1937)	-0.2202 (0.1832)	-0.5645*** (0.1756)	-0.2839 (0.3465)
Kinetic Events Dummy (Day Before)	-0.1398 (0.4350)	-0.0370 (0.3176)	0.0363 (0.4373)	-0.2866 (0.5408)
Fatalities Dummy (Day Before)	0.0948 (0.2607)	0.4231 (0.3265)	0.1181 (0.3093)	-0.3907 (0.3415)
Panel B: General Competence				
Alerts Dummy (Night Before)	0.0232 (0.1688)	-0.2487** (0.1014)	-0.2386** (0.1046)	0.2671 (0.2561)
Kinetic Events Dummy (Day Before)	0.0157 (0.3999)	0.1972 (0.3192)	0.1580 (0.3788)	-0.1507 (0.4485)
Fatalities Dummy (Day Before)	-0.3105 (0.4745)	-0.0138 (0.3464)	-0.1387 (0.4844)	-0.8622 (0.5785)
Panel C: Law				
Alerts Dummy (Night Before)	1.3490*** (0.3924)	1.1881*** (0.3864)	1.4806** (0.5436)	0.9587** (0.4092)
Kinetic Events Dummy (Day Before)	0.5914 (0.6284)	1.1191 (0.6520)	-0.1163 (0.7264)	0.1992 (0.4199)
Fatalities Dummy (Day Before)	-0.2208 (0.8033)	0.0622 (0.9628)	-0.5359 (0.7746)	-0.8681 (0.6770)

Note: Pooled TWFE OLS estimates (β) for 2022-2024 using binary dummy variables (1 if shock occurred, 0 otherwise). Standard errors clustered at the physical testing region level in parentheses.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

The estimate coefficients suggest that collapsing the intensity of kinetic events and fatalities on the day before the exam into dummy variables leads to no statistically significant effect of these dimensions of war across all three exams. This indicates that the self-selection bias driven by absenteeism is triggered

strictly by the intensity of war exposure, rather than the factual occurrence of violence. Effects air raid alerts on the night before the exam experience noticeable changes, too. Comparison of affected and unaffected by alerts students in 'Men < 24' subsample reveals that former perform by more than 0.56 (0.06 SD[15]) less correct answers in Foreign Language exam on average. At the same time, data does not suggest any statistically significant effects on 'Women' students, which we concluded from the initial TWFE OLS model. Same for General Competence exam, we now notice that the mere presence of air alert results in lower raw score for both traditional students cohorts – 'Women' and 'Men < 24', which was not present in the initial TWFE OLS. However, the magnitude of the effects for this specific model are not of economic significance. Considering the discrete nature of raw scores, an academic penalty of approximately -0.25 (0.05 SD[8]) raw score in General Competence exam for 'Women' students implies that the effect we found in the initial model is now diluted across all female test-takers who experienced air raid alerts on the night before the exam. At the same time, the Law exam exhibits a paradoxical result of highly significant positive impact of air raid alerts on students' raw scores.

The reason behind such 'flip' in all the coefficients are the different estimation approaches. The initial TWFE OLS considers the different intensity of dimensions of war and thus distinguishes between students who experienced 2 minutes and 15 hours of air alerts on the night before the exam (the least value after 0 and the maximum present in the dataset, respectively); the same logic holds for kinetic events and fatalities on the day before the exam. Meanwhile, the dummy variant TWFE OLS model compares averages for just the two groups – those, who experienced no exposure, and those, who experienced at least some portion of it, even if a minuscule one. Results of the latter model underscore that the academic penalty depends mainly on the intensity of war

[15] SD of Foreign Language exam raw score in 'Men < 24' subsample is 8.78.

exposure that results in different: a) lost hours of sleep in case of air alerts; and b) psychological shock in case of kinetic events and fatalities. And these are the exact mechanisms through which war has an immediate impact on academic performance of undergraduate students.

5.3 Two-Stage TWFE Logistic regression

5.3.1 TWFE Logistic regression estimation results for 3 main exams

To analyze how dimensions of war affect students' decisions to show up for the exam, we run TWFE Logistic regression model. The estimation results are presented in Table 9 below.

The estimation results reveal that exposure to kinetic events on the day before the exam has the most persistent effect on absenteeism across Foreign Language and General Competence exam, providing the mechanism behind the self-selection bias observed in the TWFE OLS estimates. To interpret log-odds coefficients, we calculate Average Marginal Effects (AME), reported in Table 38.

Specifically, for General Competence exam, the log-odds coefficients are positive and statistically significant at 1% significance level across all demographic subsamples. Aggregate coefficient shows an average effect on all students in the sample, and the estimated effect suggests that each additional kinetic event on a day before the exam experienced by students increases probability of them not showing up to take the exam by 0.35 percentage points on average. This may seem a small effect at first glance, but for regions that experience heavy attacks, this effect translates into high absenteeism rates.

Table 9. Impact of Immediate Shocks on Exam Absenteeism (Log-Odds)

Exposure	Aggregate	Women	Men < 24	Men 24+
Panel A: Foreign Language				
Alerts (Night Before)	0.0180 (0.0139)	0.0226** (0.0091)	0.0057 (0.0089)	0.0313 (0.0205)
Kinetic Events (Day Before)	0.0153*** (0.0058)	0.0118 (0.0119)	0.0087** (0.0044)	0.0230*** (0.0059)
Fatalities (Day Before)	-0.0012*** (0.0003)	0.0025*** (0.0006)	0.0005 (0.0005)	-0.0022*** (0.0001)
Panel B: General Competence				
Alerts (Night Before)	0.0009 (0.0185)	0.0089 (0.0122)	-0.0132 (0.0134)	0.0179 (0.0248)
Kinetic Events (Day Before)	0.0233*** (0.0035)	0.0214*** (0.0042)	0.0194*** (0.0067)	0.0265*** (0.0052)
Fatalities (Day Before)	-0.0024*** (0.0002)	0.0005 (0.0006)	-0.0012*** (0.0004)	-0.0026*** (0.0001)
Panel C: Law				
Alerts (Night Before)	0.0142 (0.0163)	-0.0044 (0.0241)	0.0480*** (0.0172)	0.0318** (0.0156)
Kinetic Events (Day Before)	-0.0057 (0.0131)	-0.0012 (0.0142)	-0.0115 (0.0088)	-0.0198 (0.0173)
Fatalities (Day Before)	-0.0080** (0.0038)	-0.0084* (0.0048)	-0.0036 (0.0060)	-0.0129*** (0.0021)

Note: Pooled TWFE Logistic regression estimates (Log-Odds) for 2022-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Specifically, prior to exam on 08.09.2022, Kharkiv Oblast experienced 46 kinetic events (according to ACLED data). For this specific region, this led to higher probability of student being absent of 16.1 percentage points on average according to the estimated effects. At the same time, looking at coefficients for Foreign Language, kinetic events do not have a statistically significant effect on probability of being absent only for 'Women' subsample, but the effect still holds for other two subsamples, as well as for the whole dataset. Looking at the effects for 'Men 24+' subsample, each additional kinetic event leads to probability of being absent to be higher by 0.43 percentage points on average.

Similarly to General Competence exam, we look at the possible maximum effect present in the sample. 35 kinetic events occurred prior to the Foreign Language exam on 07.09.2022 in Kharkiv region, and estimated effects suggest that for 'Men 24+' students, at this time and at this region, the probability of not showing up for this exam was higher by 15.05 percentage points on average. These robust findings strongly support the first assumption of **H5**, confirming that the immediate impact of dimensions of war increase the probability of a student being absent.

On the contrary, air raid alerts on the night before the exam generally do not affect absenteeism rates, with statistically significant effects being isolated to specific high-stakes contexts and specific subsamples, such as the Law exam for both male cohorts. As for fatalities on the day before the exam, unexpectedly, coefficients for this dimension of war are predominantly negative and highly statistically significant. This may be again caused by self-selection bias. As discussed previously in Chapter 4, kinetic events and fatalities are not significantly correlated, so it is expected for the estimated effects of these to be different.

As the second stage of TWFE Logistic regression estimation, we run the same model, but now estimating the conditional probability of failing the exam given the student attended the exam. The estimation results are presented in Table 10. Similarly to previous stage on absenteeism, the AME was calculated and reported in Table 42.

The very first striking finding is that across all three main exams and three dimensions of war, the estimated coefficients are exclusively negative. This inherently implies that exposure to war decreases the probability of failure for students who attend the exam. Seemingly counterintuitive at first, this finding suggests the self-selection bias, which we previously hypothesized in other models. Since war – primarily through kinetic events – filters out students that are likely underprepared or have low confidence in their potential scores (as

suggested by the first stage of TWFE Logistic regression estimates on absenteeism), the exam takers who actually attended the exam are likely high performing and resilient. This in turn results in average scores for attendees skewing higher in regions with higher war intensity and correspondingly higher absenteeism. Such finding leads us to rejection of second assumption of **H5** for the three main exams. Immediate exposure to war does not universally increase the probability of failure for those students who attend the exam but rather decrease it.

Table 10. Conditional Probability of Exam Failure Given Attendance (Log-Odds)

Exposure	**Aggregate**	**Women**	**Men < 24**	**Men 24+**
Panel A: Foreign Language				
Alerts (Night Before)	-0.0356**	-0.0177	-0.0252	-0.0620***
	(0.0147)	(0.0220)	(0.0169)	(0.0206)
Kinetic Events (Day Before)	-0.0117**	-0.0177***	-0.0042	-0.0123**
	(0.0047)	(0.0064)	(0.0118)	(0.0053)
Fatalities (Day Before)	-0.0030***	0.0012***	-0.0097***	-0.0035***
	(0.0004)	(0.0004)	(0.0014)	(0.0003)
Panel B: General Competence				
Alerts (Night Before)	-0.0123	-0.0040	0.0029	-0.0049
	(0.0371)	(0.0500)	(0.0304)	(0.0427)
Kinetic Events (Day Before)	-0.0231	-0.0386	-0.0035	-0.0355
	(0.0371)	(0.0261)	(0.0271)	(0.0503)
Fatalities (Day Before)	-0.0050***	0.0002	-0.0278***	0.0002
	(0.0009)	(0.0012)	(0.0034)	(0.0021)
Panel C: Law				
Alerts (Night Before)	-0.0094	-0.0173	-0.0159	0.0731
	(0.0348)	(0.0413)	(0.0499)	(0.0664)
Kinetic Events (Day Before)	-0.0152	0.0206	-0.0391	-0.0420
	(0.0151)	(0.0362)	(0.0324)	(0.0272)
Fatalities (Day Before)	-0.0051	-0.0017	-0.0067	-0.0101
	(0.0039)	(0.0069)	(0.0059)	(0.0081)

Note: Pooled TWFE Logistic regression estimates (Log-Odds) for 2022-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

We must note that other mechanisms – such as possible unobserved persistency of 2022 leniency policy during 2023-2024; or a survivorship bias driven by students who were displaced westward but still attended, representing both high-resource and high-motivation cohorts regardless of the original exposure to war – can mathematically contribute to estimated negative effects of war variables on probability of failure. However, when paired with the first-stage finding of non-random absenteeism, the data suggests the self-selection as the primary driver of this phenomenon.

This primary self-selection described above is mostly noticeable in Foreign Language exam, where, for 'Men 24+' and 'Women' subsamples, the AME for kinetic events on the day before the exam suggest that each additional kinetic event decreases the probability of failing (or increases the probability of passing) by 0.05 percentage points on average. Using the same reference 35-events day as in the previous stage, this effect translates into failure probability reduced by 1.75 percentage points on average for corresponding students.

For the General Competence and Law exams, the conditional probability effects are either not statistically significant at any conventional significance level or are negligible in their magnitude. This indicates that, for these exams, attending students' failure probability is inelastic to recent exposure to war.

This two-stage estimation of TWFE Logistic regression documents the mechanism behind the anomalies observed in the TWFE OLS estimates. The data suggests that the immediate impact on academic performance due to exposure to war is primarily concentrated on attendance rather than scores of the resilient students that actually take the exam.

Same as for all other models, the full estimation results and forest plots for both stages of TWFE Logistic regression estimation are reported in Appendix H.

5.3.2 TWFE Logistic regression estimation results for other UPEs

To check whether data on other UPEs suggests same findings, we run the same Two-Stage TWFE Logistic regression estimation for these exams. The estimation results are reported in Appendix I.

The estimated effects for other UPEs are highly heterogeneous, with data on several exams suggesting conclusions that contradict the findings observed in the three main exams. Regarding absenteeism, while kinetic events on the day before the exam hold their positive impact on absenteeism rate for such exams like Psychology and Pedagogy, students who take Economics or Political Science exams are vulnerable to alerts on the night before the exam – this effect was largely absent in the three main exams. Furthermore, students who take Political Science exams are also prone to be absent due to the fatalities on the day before the exam.

As for the failure rates, the estimation results suggest much deeper contradictions. Unlike the three main exams, where the self-selection consistently decreased failure rates, other exams such as Political Science, Finance, and Management have statistically significant positive coefficients in case of kinetic events or fatalities on the day before the exam, meaning that each additional occurrence of these dimensions of war effectively increases the failure rate. This inversion indicates that, for certain subjects, the previously observed self-selection bias does not hold. The immediate impact of war exposure not only impacts the decision to show up for these exams, but also reaches students who attend them, increasing the probability of failure.

5.3.3 TWFE Logistic regression robustness checks

Full estimation results for all robustness checks are reported in Appendix J.

LPM estimation

As the first robustness check of TWFE Logistic regression, we run an LPM estimation on the same data. The estimation results for absenteeism LPM are reported in Table 11 below, and for conditional failure – in Table 12 (p.53).

Table 11. TWFE Logistic Regression Robustness Check: Linear Probability Model (Absenteeism)

Exposure	**Aggregate**	**Women**	**Men < 24**	**Men 24+**
Panel A: Foreign Language				
Alerts (Night Before)	0.0023 (0.0020)	0.0029** (0.0012)	0.0006 (0.0009)	0.0059 (0.0040)
Kinetic Events (Day Before)	0.0021** (0.0008)	0.0018 (0.0018)	0.0008 (0.0005)	0.0044*** (0.0012)
Fatalities (Day Before)	-0.0002*** (0.0000)	0.0004*** (0.0001)	0.0001 (0.0000)	-0.0004*** (0.0000)
Panel B: General Competence				
Alerts (Night Before)	0.0002 (0.0027)	0.0011 (0.0018)	-0.0015 (0.0015)	0.0034 (0.0048)
Kinetic Events (Day Before)	0.0036*** (0.0006)	0.0035*** (0.0007)	0.0024** (0.0009)	0.0052*** (0.0011)
Fatalities (Day Before)	-0.0004*** (0.0000)	0.0001 (0.0001)	-0.0001*** (0.0000)	-0.0005*** (0.0000)
Panel C: Law				
Alerts (Night Before)	0.0015 (0.0022)	-0.0006 (0.0025)	0.0049** (0.0018)	0.0058* (0.0033)
Kinetic Events (Day Before)	-0.0005 (0.0015)	-0.0001 (0.0015)	-0.0010 (0.0007)	-0.0037 (0.0036)
Fatalities (Day Before)	-0.0009** (0.0005)	-0.0010* (0.0005)	-0.0003 (0.0005)	-0.0026*** (0.0006)

Note: LPM estimates for 2022-2024 representing percentage point changes in absenteeism probability.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 12. TWFE Logistic Regression Robustness Check: Linear Probability Model (Conditional Failure)

Exposure	Aggregate	Women	Men < 24	Men 24+
Panel A: Foreign Language				
Alerts (Night Before)	-0.0003	0.0002	-0.0002	-0.0017**
	(0.0005)	(0.0006)	(0.0004)	(0.0008)
Kinetic Events (Day Before)	-0.0002	-0.0004	0.0000	-0.0004**
	(0.0001)	(0.0003)	(0.0002)	(0.0002)
Fatalities (Day Before)	-0.0001***	0.0001**	-0.0002***	-0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Panel B: General Competence				
Alerts (Night Before)	-0.0000	0.0000	0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Kinetic Events (Day Before)	-0.0000	-0.0001**	-0.0000	-0.0001
	(0.0001)	(0.0000)	(0.0000)	(0.0001)
Fatalities (Day Before)	-0.0000***	0.0000**	-0.0000***	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Panel C: Law				
Alerts (Night Before)	0.0000	0.0002	-0.0003	0.0007
	(0.0003)	(0.0004)	(0.0004)	(0.0010)
Kinetic Events (Day Before)	-0.0002	-0.0001	-0.0002	-0.0006
	(0.0002)	(0.0002)	(0.0004)	(0.0004)
Fatalities (Day Before)	-0.0001**	-0.0000	-0.0002*	-0.0003*
	(0.0001)	(0.0001)	(0.0001)	(0.0001)

Note: LPM estimates for 2022-2024 representing percentage point changes in conditional failure probability.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

The LPM estimates for absenteeism almost perfectly mirror direction, magnitude, and statistical significance of the previously calculated AME of TWFE Logistic Regression. Specifically, kinetic events on the day before the exam maintain the same impact on absenteeism for General Competence and Foreign Language exams; the same holds for air raid alerts on the night before the exam effects in both male subsamples for Law exam. Besides, the estimated negative effects of fatalities on the day before the exam remain similar.

As mentioned before, we do the same robustness check for failure rates. The estimation results for conditional failures LPM are reported in Table 12 on p.53. The same mirroring holds in the results for conditional failure. In general, this almost perfect alignment between the two estimators confirms that the same mechanism behind self-selection bias holds.

Switching to Binary Treatment

As we did for TWFE OLS, we also test the sensitivity of TWFE Logistic regression estimates to the measurement approach by running TWFE Logistic regression model with binary treatment. The estimation results for absenteeism are reported in Table 13 below.

Table 13. TWFE Logistic Regression Robustness Check: Binary Treatment Logistic regression (Absenteeism)

Exposure	Aggregate	Women	Men < 24	Men 24+
Panel A: Foreign Language				
Alerts (Night Before)	0.0721***	0.0297	0.0842**	0.0618**
	(0.0257)	(0.0529)	(0.0412)	(0.0308)
Kinetic Events (Day Before)	0.2454**	0.2057***	0.1396	0.3196**
	(0.1173)	(0.0662)	(0.1227)	(0.1313)
Fatalities (Day Before)	0.1483***	0.1593	0.1395**	0.1507
	(0.0445)	(0.0984)	(0.0630)	(0.0980)
Panel B: General Competence				
Alerts (Night Before)	0.0837***	0.0558*	0.0940**	0.0698**
	(0.0244)	(0.0295)	(0.0385)	(0.0308)
Kinetic Events (Day Before)	0.2414**	0.1860**	0.1521	0.3197**
	(0.1187)	(0.0835)	(0.1185)	(0.1306)
Fatalities (Day Before)	0.1459***	0.1157	0.1315**	0.1495
	(0.0494)	(0.0785)	(0.0637)	(0.0989)
Panel C: Law				
Alerts (Night Before)	-0.0311	0.0583	-0.1070	-0.1062
	(0.0499)	(0.0721)	(0.0769)	(0.0739)
Kinetic Events (Day Before)	-0.3426***	-0.4136***	-0.2892**	-0.4160***
	(0.0766)	(0.0913)	(0.1201)	(0.0742)

Table 13 – Continued

Exposure	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.1661 (0.1677)	-0.1665 (0.2019)	0.0063 (0.2435)	-0.3356*** (0.1249)

Note: Logistic regression estimates (log-odds) for 2022-2024 measuring the impact of shock presence (1) vs absence (0) on absenteeism.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

The estimated coefficients deviate largely from the initial TWFE Logistic regression with continuous treatment, with the most remarkable examples being (a) inverted coefficients for fatalities on the day before the exam in General Competence exam; and (b) highly statistically significant negative coefficients for kinetic events on the day before the exam in Law exam. Exactly as observed in TWFE OLS robustness check of the same approach, collapsing the variation in treatment to binary indicators results in distorted estimated effects, which in turn conceals the true dimensions of war impact. This happens because binary variable treats isolated attack the same as mass bombardment, and thus averages out the actual effects of the extremes in treatment intensity.

We do the same robustness check for failure rates. The estimation results for conditional failures TWFE Logistic regression with binary treatment are reported in Table 14 below. Mirroring the binary absenteeism results, the estimated coefficients are highly inconsistent with the initial model. These inversions in estimated effects re-confirm the methodological validation for our continuous treatment specification (firstly confirmed in TWFE OLS robustness check): capturing the true impact of dimensions of war on students' academic performance requires measuring the intensity of exposure rather than merely its occurrence.

Table 14. TWFE Logistic Regression Robustness Check: Binary Treatment Logistic regression (Conditional Failure)

Exposure	Aggregate	Women	Men < 24	Men 24+
Panel A: Foreign Language				
Alerts (Night Before)	-0.0334 (0.0581)	-0.0155 (0.0907)	0.0080 (0.0682)	-0.0635 (0.0969)
Kinetic Events (Day Before)	0.0057 (0.1031)	-0.0954 (0.1023)	-0.1022 (0.0860)	0.0895 (0.1357)
Fatalities (Day Before)	0.1711 (0.1632)	0.2532*** (0.0791)	0.2214* (0.1245)	0.1042 (0.2309)
Panel B: General Competence				
Alerts (Night Before)	0.1057 (0.1338)	0.0584 (0.3112)	0.3967** (0.1852)	0.0350 (0.1552)
Kinetic Events (Day Before)	-0.4595** (0.2255)	-0.6026 (0.6005)	-0.7758** (0.3617)	-0.3049 (0.4583)
Fatalities (Day Before)	0.8385* (0.4300)	0.2742 (0.2788)	0.1069 (0.1067)	1.6429*** (0.5877)
Panel C: Law				
Alerts (Night Before)	-0.6894*** (0.1538)	-0.6195*** (0.1476)	-0.8964*** (0.2021)	-0.2363 (0.3252)
Kinetic Events (Day Before)	-0.1806* (0.1083)	-0.0828 (0.1895)	-0.2548 (0.2840)	-0.3395 (0.2209)
Fatalities (Day Before)	-0.0872 (0.1365)	-0.0697 (0.2234)	-0.0194 (0.2515)	-0.2768 (0.2178)

Note: Logistic regression estimates (log-odds) for 2022-2024 measuring the impact of shock presence (1) vs absence (0) on conditional failure.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

5.4 Multiple hypothesis test corrections

To ensure that our findings are not driven by false positives arising from the large number of models estimated across exam subjects, demographic subsamples, and dimensions of war, we apply the Benjamini-Hochberg (1995) False Discovery Rate (FDR) procedure. We use exactly this procedure to balance the trade-offs between the 'multiple hypothesis' penalty and

maintaining the statistical power. The p-values are adjusted within each model's respective hypothesis family. The full results are reported in Appendix K.

For the TWFE Continuous DID models (Table 71), the MHT correction successfully filters out isolated coefficients. While a few aggregate effects with low statistical significancy lose it completely, the severe and highly significant effects survive the MHT correction across Foreign Language for 'Men 24+' subsample, General Competence for traditional student cohorts represented by 'Women' and 'Men < 24' subsamples, and Law for all subsamples. This confirms the results of this model discussed in section 5.1.3 above.

The Short-Run FE OLS models (Table 72) estimates are exceptionally robust to the FDR correction. While isolated effects at the 5% and 10% significance levels naturally fall above the adjusted thresholds, the core self-selection evidence – most notably the impact of fatalities on the day before the exam on General Competence and the impact of kinetic events on the day before the exam on Foreign Language – maintain high statistical significance after the correction. This alignment validates our initial conclusion on presence of self-selection mechanism revealed in unexpected direction of TWFE OLS estimates.

The Absenteeism Logistic regression models (Table 73) are the most resilient to the MHT correction in our study. The FDR correction barely alters the coefficients' statistical significance. The positive effects of kinetic events and fatalities on the day before the exam on absenteeism across the Foreign Language and General Competence exams survive with 1% and 5% adjusted significance across almost all estimated coefficients. This overwhelming results stability definitively proves that the 'filtering' effect of dimensions of war on attendance is undeniable. Besides, the foundation of self-selection mechanism is confirmed.

Finally, the Conditional Failure Logistic regression models (Table 74) also survive the MHT correction. The negative coefficients that translate into

exposure to war decreasing the likelihood of failure for attending students hold high statistical significance for Foreign Language and specific demographic subsamples in General Competence. This provides the final statistical validation for the self-selection mechanism: the immediate impact of dimensions of war forces a 'filtering' process, leading low-performing students to not show up for the exams, leaving more resilient ones among the actual test-takers, whose conditional passing rates skew higher.

5.5 Data limitations and measurement constraints

The empirical strategy of this thesis provides critical insights into the impact of the russian invasion on Ukrainian students' academic performance. Even though the estimation results of the main models used in this study have successfully passed multiple robustness checks, there are several data limitations and measurement constraints that must be acknowledged.

To ensure spatial consistency in the data between regional fixed effects and exposure to war, war metrics were mapped to the physical testing regions rather than to universities that registered students. However, this approach only leads to a proxy for true exposure. Internally displaced students that moved out of highly affected by war eastern and southern regions to relatively safer western regions have their exposure to war biased downward. This in turn introduces the attenuation bias in the cumulative exposure coefficients and potentially masks true academic penalties among students of traditional cohorts. Besides,

Another limitation is introduced through the exclusion of the most exposed to war regions – Donetsk, Luhansk, and Kherson – due to impossibility of administering UEE and UPE there. This leads to losing observations on precisely the most exposed students, and estimated coefficients to likely under-represent the upper bound of true impact of the russian invasion.

The geographic aggregation of the data yields 22 physical regions in which the exam was administered. Econometric literature demonstrates that inference with such a small number of clusters can lead to standard cluster-robust standard errors over-rejecting the null hypothesis (Cameron et al. 2008). Because a wild bootstrapping was not applied, and the aggregation was not done at a finer level (raions or cities), the reported statistical significance that lies at the borderline should be interpreted with caution.

At the same time, the demographic composition of the dataset shifted dramatically during the war – most notably due to the "mobilization effect" observed in 2024. The influx of older male individuals likely seeking university enrollment to grant draft deferment alters the baseline characteristics of the sample, make it difficult to directly compare student cohorts in pre-war and wartime periods.

The unavailability of 2020 exam data as well as earlier data restricts the pre-war baseline to only two periods – 2019 and 2021. Hence, the event-study design with 2021 omitted as a reference year relies on a single pre-treatment coefficient for 2019 to validate parallel trends assumption, which makes the event-study design in this thesis an underpowered test.

The two-stage logistic regression successfully documents non-random absenteeism and conditional failure, thus providing suggestive evidence of the self-selection mechanism. However, this process does not correct the bias present in TWFE Continuous DID and TWFE OLS estimations. To deliver effects on scores corrected for this bias, it is necessary to run a Heckman-style selection model with a valid exclusion restriction, which is beyond the scope of data available for this study.

Finally, while this study finds evidence of war negatively affecting academic performance of Ukrainian students, we cannot identify separately the channels (sleep deprivation, acute stress, infrastructure damages, student displacements, etc.) that drive the observed effects.

Chapter 6

CONCLUSIONS AND POLICY RECOMMENDATIONS

The primary objective of this thesis was to quantify the causal impact of the russian invasion on the academic performance of Ukrainian university students, specifically focusing on standardized exams required for Master's degree admission. The analysis centered on the many-sided dimensions of war – air raid alerts, kinetic events, and fatalities – and their immediate and cumulative effects on academic outcomes. By leveraging a multi-model econometric framework, we evaluated the impact across different subjects, accounting for demographic shifts, regional disparities, and exam administration adapting to war environment. We used TWFE Continuous DID, TWFE OLS and Two-Stage TWFE Logistic regression models, complemented by rigorous robustness checks, to isolate the true causal effect of exposure to war on students willing to continue the academic education in Ukraine after getting Bachelor's degree. Foreign Language, General Competence, and Law constitute the main exams for this analysis, while other UPEs were analyzed complementarily in order to evaluate the heterogenous immediate impacts of dimensions of war on newly introduced specialized exams that lacked the pre-war baseline to apply the TWFE Continuous DID design.

To evaluate cumulative effect of dimensions of war, we utilized a TWFE Continuous DID design. Surprisingly, for traditional student cohorts (females and males under 24), cumulative war exposure metrics – total hours of air raid alerts, or the total number of kinetic events and fatalities over the final academic year – had no statistically significant effects. This suggests a high level of academic adaptation and resilience among traditional university students to the prolonged exposure to the war environment. However, the TWFE Continuous DID estimations did reveal a persistent, statistically significant anomalies for the older male demographic (males 24 and older); by anomalies we mean the

presence of *positive* effects of cumulative war exposure metrics on the scores of this specific demographic group. This finding empirically confirmed a severe compositional distortion: a massive 'mobilization effect' in 2024, when an influx of non-traditional registrants likely seeking military deferment rather than pursuing a higher degree skewed aggregate academic performance metrics.

While traditional students appeared to be resilient to the *cumulative* exposure to war, immediate impacts revealed a completely different picture. The initial TWFE OLS estimations conceived seemingly counterintuitive results. Specifically, the immediate impact of kinetic events and fatalities on the day before the exam frequently appeared to have either an economically negligible or even positive effect on academic outcomes across the three main exams. However, the subsequent estimation of Two-Stage TWFE Logistic regression models documents the mechanism behind the anomalies found in two other models' results, uncovering an underlying self-selection mechanism.

The first stage of the TWFE Logistic regression estimation suggested that immediate effect of exposure to war, particularly the kinetic events, triggers absenteeism dependent on the intensity of such exposure. Through this 'filtering' process, those students who do show up to the testing centers to take the exam appear not as a representative sample of fresh Bachelors, but rather a self-selected and academically resilient cohort; and the results of the first stage estimation set a foundation for such 'filtering'. The second stage of the TWFE Logistic regression confirms the hypothesized mechanism, revealing its impact on average scores. For this self-selected cohort of students, likelihood of conditional failing (given they showed up for the exam) actually skewed downward. Hence, the 'positive' effects of dimensions of war observed in the FE OLS were a statistical artifact caused by the self-selection 'filter', that left all vulnerable or low performing students without attending the exam.

It is crucial to mention that the data suggests that those effects are not uniform across all existing exams, but are rather mostly observed in two exams that are

mandatory to everyone willing to apply for Master's academic programs – Foreign Language and General Competence. For instance, when analyzing specific UPEs (such as Political Science and Finance), this self-selection mechanism was not present. It appears that the dimensions of war negatively impacted not only the attendance, but also the scores of those who did take the exam, resulting in higher conditional failure rates along with higher absenteeism.

To confirm that our findings about self-selection and causal effect were reliable, we performed several robustness checks. Firstly, we applied LPM model, which closely followed the results suggested by Average Marginal Effects of the TWFE Logistic regression. This confirmed the robustness of estimated effects on absenteeism and failure rates across different functional forms.

Furthermore, replacing the continuous war variables with binary variants (1 if any exposure to war occurred, and 0 otherwise) resulted in highly distorted results for both TWFE OLS and TWFE Logistic regression models. This demonstrates that the negative impact on academic performance is determined by the intensity of exposure rather than the mere fact of occurrence of such exposure.

Finally, to address concerns of false positive arising from numerous estimations, we applied the Multiple Hypothesis Testing correction, specifically the Benjamini-Hochberg False Discovery Rate (FDR) procedure. The core findings suggesting the self-selection mechanism remained statistically significant even after adjusting the p-value thresholds, confirming the reliability of our conclusions.

The empirical findings of this thesis not only contribute to the academic literature on human capital accumulation during armed conflict but also propose implications for corresponding policy on conducting standardized exams in Ukraine. The evidence suggests that a uniform, peacetime-approach to high-stakes exams is no longer equitable under wartime conditions. Based

on our findings, we propose three specific policy recommendations for the Ministry of Education and Science of Ukraine.

1. Adjustment of Exam Timing

Empirical evidence demonstrates that air raid alerts on the night before the exam can significantly disrupt the baseline cognitive capacity of students. To mitigate this immediate impact, the Ministry should consider implementing a flexible testing window with both morning and afternoon administration. A dynamic scheduling approach could automatically assign students to an afternoon slot if their physical testing region experienced an overnight air raid alert. This targeted adjustment would provide the most affected students with a time buffer, allowing for partial physiological and psychological recovery from sleep deprivation before entering the test centers.

2. Expansion of Exam Schedules and Variant Pools

Because immediate impact of kinetic events on the day before the exam affects whether the student will decide to attend the exam, relying on narrowly defined testing window systematically narrows possibilities for students to actually take the exam. The exam schedule should be expanded to include a wider list of available examination dates, allowing students to take exam later in case of a massive attack on their region. Crucially, to preserve the academic integrity of standardized exams and eliminate risks of tasks being leaked, the expansion must come along with a corresponding increase in exam variants. This ensures flexibility without hurting the fair evaluation.

3. Accounting for Mobilization-Induced Demographic Distortions

The TWFE Continuous DID estimates revealed a severe compositional distortion caused by an influx of older male registrants in 2024 likely seeking mobilization deferment guaranteed to students taking a degree of specific level for the first time. Because this specific cohort artificially skews aggregate academic performance downward, the Ministry must account for this demographic shift when diagnosing the overall state of the higher education system. It is important to maintain the academic reputation of Ukrainian universities and ensure equitable state funding allocations. Policymakers should systematically disaggregate the results of this non-traditional demographic from the core student body during administrative evaluations and institutional rankings. However, it is important to consider that among these older male registrants there are still those who actually seeks academic education.

WORKS CITED

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education, First Edition*. NBER. https://www.nber.org/books-and-chapters/human-capital-theoretical-and-empirical-analysis-special-reference-education-first-edition.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)* 57 (1): 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1): 249–75.

Blattman, Christopher, and Jeannie Annan. 2010. "The Consequences of Child Soldiering." *Review of Economics and Statistics* 92 (4): 882–98. https://doi.org/10.1162/REST_a_00036.

Brück, Tilman, Michele Di Maio, and Sami H. Miaari. 2019. "Learning The Hard Way: The Effect of Violent Conflict on Student Academic Achievement." *Journal of the European Economic Association* 17 (5): 1502–37. https://doi.org/10.1093/jeea/jvy051.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2024. "Difference-in-Differences with a Continuous Treatment." *National Bureau of Economic Research*, NBER Working Paper Series, February.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *The Review of Economics and Statistics* 90 (3): 414–27.

Caruso, German, Iffath Sharif, Victoria Strokova, et al. 2023. *The War in Ukraine and Human Capital Losses*. Brief (BRI). World Bank Group, Human Capital Project. https://thedocs.worldbank.org/en/doc/5c9598063f47f2c96935621be315173b-0140062023/the-war-in-ukraine-and-human-capital-losses.

Deng, Yuwei, Jacob Cherian, Noor Un Nisa Khan, et al. 2022. "Family and Academic Stress and Their Impact on Students' Depression Level and Academic Performance." *Frontiers in Psychiatry* 13 (June): 869337. https://doi.org/10.3389/fpsyt.2022.869337.

Di Pietro, Giorgio. 2018. "The Academic Impact of Natural Disasters: Evidence from L'Aquila Earthquake." *Education Economics* 26 (1): 62–77.

DIF (Ilko Kucheriv Democratic Initiatives Charitable Foundation). 2021. "How Ukrainians evaluate education during pandemic." Survey. July 19. https://dif.org.ua/article/how-ukrainians-evaluate-education-during-pandemic.

Grueso, Hernando. 2023. "The Impact of Armed Conflict on College Students." *EdWorkingPaper* 23–733. https://doi.org/10.26300/A0DV-7761.

Ichino, Andrea, and Rudolf Winter-Ebmer. 2004. "The Long-Run Educational Cost of World War II." *Journal of Labor Economics* 22 (1): 57–87. https://doi.org/10.1086/380403.

Karakai, Antonii, and Valentyn Hatsko (VoxUkraine). 2025. "Online Learning in Ukraine: How It Works and Whether Knowledge Levels Are Declining (Part 2)." September 25. https://voxukraine.org/en/online-learning-in-ukraine-how-it-works-and-whether-knowledge-levels-are-declining-part-2.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Klymenko, Vadym. 2025. "Dataset of Air Raid Sirens in the Ukraine-Russia War." GitHub. https://github.com/Vadimkin/ukrainian-air-raid-sirens-dataset.

Leon, Gianmarco. 2012. "Civil Conflict and Human Capital Accumulation: The Long-Term Effects of Political Violence in Perú." *The Journal of Human Resources* (Madison) 47 (4): 991–1022. https://doi.org/10.1353/jhr.2012.0036.

Lupien, Sonia J., Bruce S. McEwen, Megan R. Gunnar, and Christine Heim. 2009. "Effects of Stress throughout the Lifespan on the Brain, Behaviour and Cognition." *Nature Reviews. Neuroscience* 10 (6): 434–45. https://doi.org/10.1038/nrn2639.

Michaelsen, Maren M., and Paola Salardi. 2020. "Violence, Psychological Stress and Educational Performance during the 'War on Drugs' in Mexico." *Journal of Development Economics* 143 (March): 102387. https://doi.org/10.1016/j.jdeveco.2019.102387.

Mykhailyshyna, Dariia. 2025. "Education Under Fire: The Short-Run Effect of War on Academic Achievement." SSRN Scholarly Paper No. 5336845. Social Science Research Network, July 3. https://doi.org/10.2139/ssrn.5336845.

Orlov, George, Douglas McKee, James Berry, et al. 2021. "Learning during the COVID-19 Pandemic: It Is Not Who You Teach, but How You Teach." *Economics Letters* 202 (May): 109812. https://doi.org/10.1016/j.econlet.2021.109812.

Polovko, Olena, and Sergei Glotov. 2023. "The educational process of Ukrainian university students following the full-scale Russian invasion." *London Review of Education* 21 (1). https://doi.org/10.14324/LRE.21.1.31.

Swee, Eik Leong. 2015. "On War Intensity and Schooling Attainment: The Case of Bosnia and Herzegovina." *European Journal of Political Economy* 40 (December): 158–72. https://doi.org/10.1016/j.ejpoleco.2015.08.001.

Thamtanajit, Kawin. 2020. "The Impacts Of Natural Disaster On Student Achievement: Evidence From Severe Floods in Thailand." *Journal of Developing Areas* 54 (4): 129–43.

UNESCO. 2023. *In the Face of War, a Year of Action in Ukraine*. United Nations. https://doi.org/10.18356/9789210025881.

World Bank. 2022. *Education : Impact of the War in Ukraine*. Brief No. 172731. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099945306202211104.

APPENDIX A. GEOGRAPHIC CHECKS

Table 15. Student Observations by University Region and Year

Registration University	**2019**	**2021**	**2022**	**2023**	**2024**
Cherkasy Oblast	1,752	3,256	1,211	2,790	7,106
Chernihiv Oblast	878	1,469	543	1,358	3,583
Chernivtsi Oblast	1,648	2,110	1,147	1,903	3,935
Dnipropetrovsk Oblast	4,304	7,289	3,083	6,277	18,997
Donetsk Oblast	856	2,284	283	953	2,133
Ivano-Frankivsk Oblast	1,111	2,586	1,055	2,546	7,455
Kharkiv Oblast	8,052	14,229	4,141	10,549	26,717
Kherson Oblast	662	2,232	70	826	2,122
Khmelnytskyi Oblast	1,169	2,283	1,144	2,371	5,495
Kirovohrad Oblast	440	1,501	440	1,092	3,159
Kyiv City	16,940	27,352	11,859	24,763	71,322
Kyiv Oblast	1,366	1,500	1,346	1,848	5,620
Luhansk Oblast	645	1,894	396	669	2,526
Lviv Oblast	5,683	9,829	4,778	10,280	30,279
Mykolaiv Oblast	1,026	2,177	443	1,520	3,785
Odesa Oblast	5,206	8,027	3,076	5,967	14,237
Poltava Oblast	1,903	3,135	1,476	2,651	9,154
Rivne Oblast	1,222	2,912	892	2,417	6,399
Sumy Oblast	1,333	2,686	945	2,371	5,691
Ternopil Oblast	2,159	3,803	1,793	3,373	10,524
Vinnytsia Oblast	1,968	3,453	1,560	3,347	8,157
Volyn Oblast	937	2,047	951	2,168	6,221
Zakarpattia Oblast	948	1,555	845	1,758	4,712
Zaporizhzhia Oblast	2,198	4,825	1,328	3,136	11,392
Zhytomyr Oblast	1,008	2,406	1,070	2,275	5,239

Table 16. Student Observations by Physical Testing Regions and Year

PTReg	2019	2021	2022	2023	2024
Abroad	0	0	0	2,270	3,771
Cherkasy Oblast	1,366	3,129	1,211	2,930	6,173
Chernihiv Oblast	890	1,591	543	1,382	3,409
Chernivtsi Oblast	1,781	2,342	1,147	2,190	4,663
Dnipropetrovsk Oblast	4,490	8,439	3,083	8,228	22,545
Donetsk Oblast	1,371	3,755	283	0	0
Ivano-Frankivsk Oblast	1,573	3,204	1,055	3,319	9,315
Kharkiv Oblast	7,479	12,417	4,141	5,688	15,476
Kherson Oblast	866	2,331	70	0	0
Khmelnytskyi Oblast	1,200	2,335	1,144	2,833	6,225
Kirovohrad Oblast	645	1,419	440	1,521	3,615
Kyiv City	16,767	22,849	11,859	21,585	71,956
Kyiv Oblast	668	3,519	1,346	2,732	10,783
Luhansk Oblast	693	1,742	396	0	0
Lviv Oblast	5,555	9,360	4,778	10,479	32,984
Mykolaiv Oblast	1,194	2,427	443	1,742	3,835
Odesa Oblast	4,512	7,686	3,076	5,684	13,601
Poltava Oblast	1,751	3,368	1,476	3,707	10,864
Rivne Oblast	1,352	3,182	892	2,880	6,618
Sumy Oblast	1,340	2,717	945	2,539	5,429
Ternopil Oblast	1,846	3,255	1,793	3,125	8,657
Vinnytsia Oblast	2,508	4,185	1,560	4,203	10,112
Volyn Oblast	1,078	2,488	951	2,734	7,036
Zakarpattia Oblast	1,068	1,991	845	2,231	5,422
Zaporizhzhia Oblast	2,176	4,355	1,328	2,504	7,882
Zhytomyr Oblast	1,245	2,754	1,070	2,702	5,589

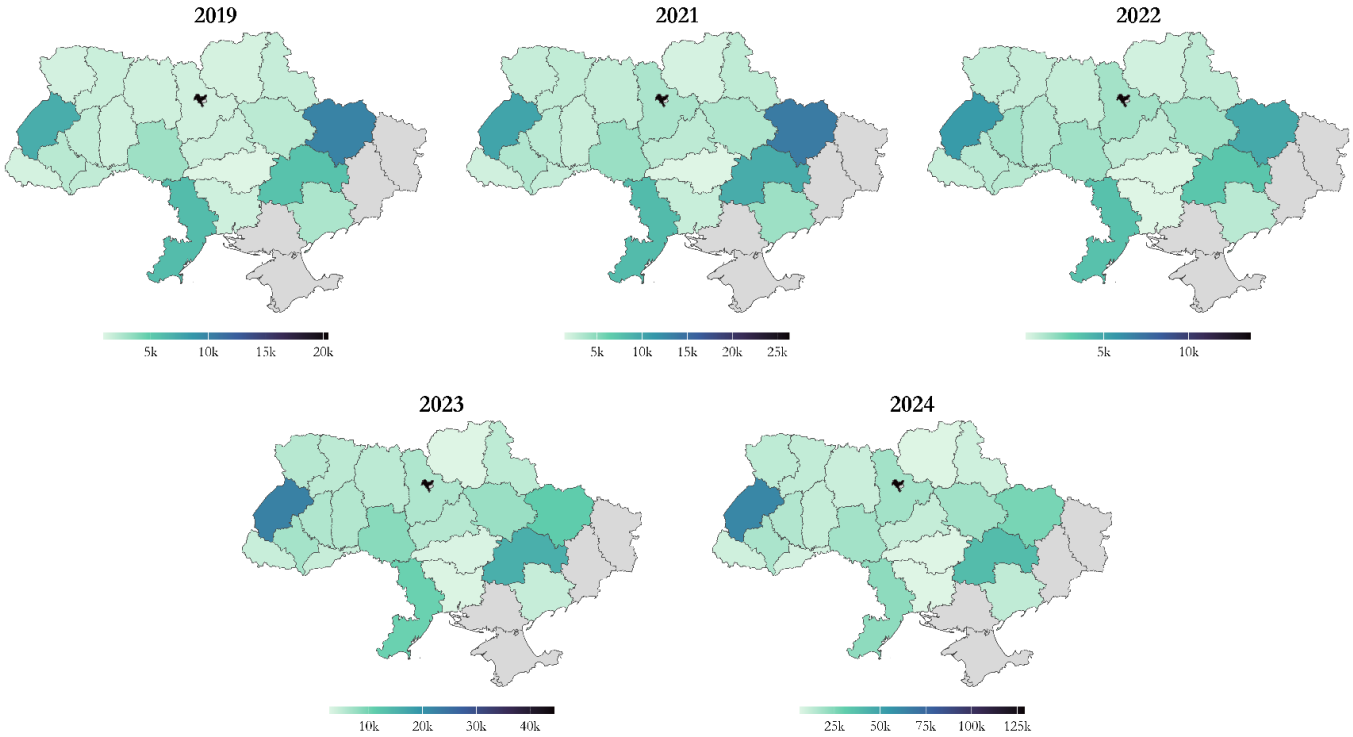


Figure 6. Spatial Evolution of Exam Participation (2019, 2021-2024)

Note: Color scales are calculated independently per year. Grey regions indicate regions with zero exams. 'k' represents thousands.

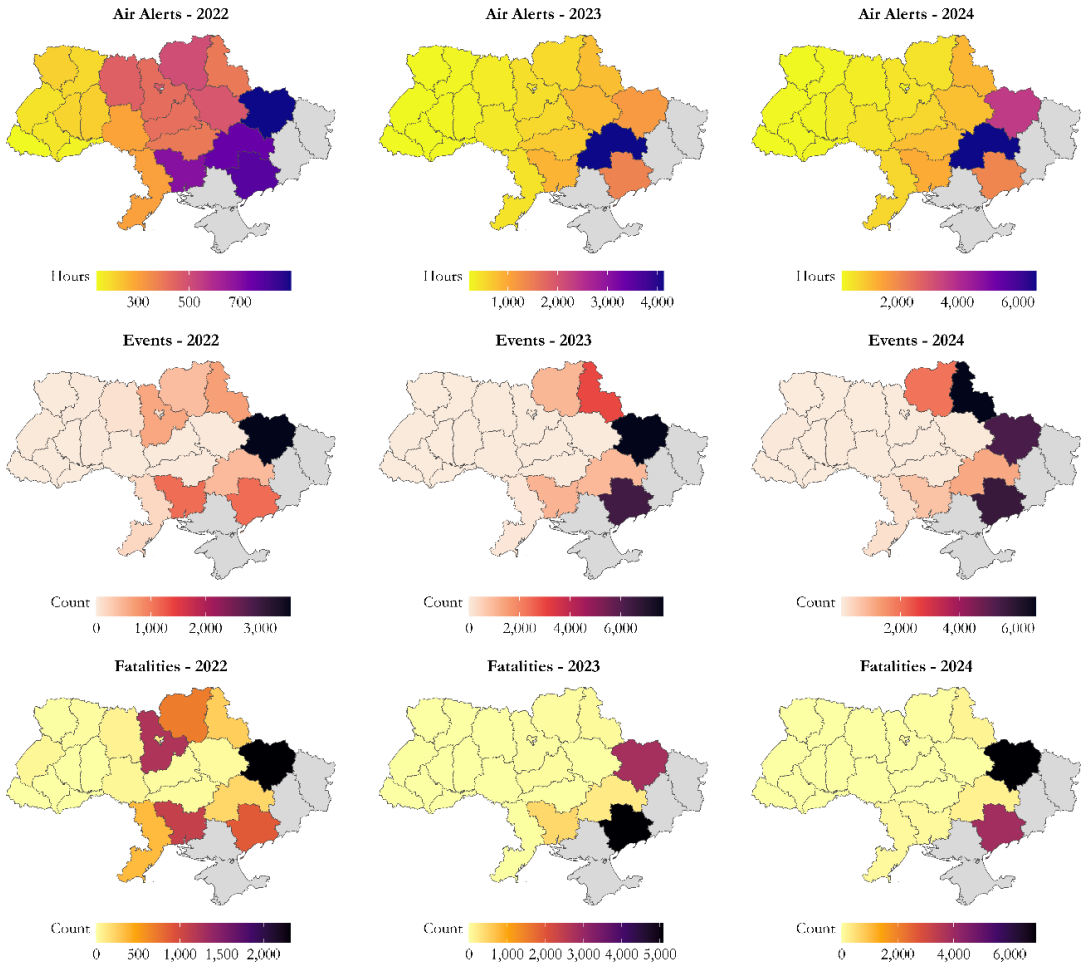


Figure 7. Spatial Evolution of Conflict Intensity (2022-2024)

Note: Color scales are calculated independently for each map to reveal relative local intensity within specific academic year in which the exam was taken. Grey regions indicate regions with zero exams.

APPENDIX C. EVENT-STUDY ESTIMATION

Table 17. Event-Study Estimation Results

Panel	Interaction Term	General Competence	Foreign Language	Law
Panel A: Future Air Raid Alerts	Exposure × I(Year = 2019)	0.0000***	0.0002***	0.0001
		(0.0000)	(0.0000)	(0.0001)
	Exposure × I(Year = 2022)	0.0000	0.0002	0.0001*
		(0.0001)	(0.0001)	(0.0001)
	Exposure × I(Year = 2023)	-0.0000	0.0001	-0.0000
		(0.0001)	(0.0001)	(0.0001)
	Exposure × I(Year = 2024)	-0.0000	0.0001	-0.0001
		(0.0001)	(0.0001)	(0.0001)
Panel B: Future Kinetic Events	Exposure × I(Year = 2019)	0.0000	0.0000**	0.0001***
		(0.0000)	(0.0000)	(0.0000)
	Exposure × I(Year = 2022)	-0.0000	0.0001**	-0.0000
		(0.0000)	(0.0000)	(0.0000)
	Exposure × I(Year = 2023)	-0.0001***	-0.0000	-0.0001***
		(0.0000)	(0.0001)	(0.0000)
	Exposure × I(Year = 2024)	-0.0001***	-0.0000	-0.0001***
		(0.0000)	(0.0001)	(0.0000)
Panel C: Future Fatalities	Exposure × I(Year = 2019)	0.0000	0.0000**	0.0001***
		(0.0000)	(0.0000)	(0.0000)
	Exposure × I(Year = 2022)	-0.0000*	0.0001**	-0.0000
		(0.0000)	(0.0001)	(0.0000)
	Exposure × I(Year = 2023)	-0.0001***	-0.0000	-0.0001***
		(0.0000)	(0.0001)	(0.0000)
	Exposure × I(Year = 2024)	-0.0001***	-0.0000	-0.0002***
		(0.0000)	(0.0001)	(0.0000)

Note: The table presents the marginal event-study coefficients (δ) estimating the impact of regional war exposure on raw exam scores. Standard errors clustered at the physical testing region level are reported in parentheses below the estimates. The models include Age and Gender controls, as well as PTReg and Year two-way fixed effects. 2021 serves as the omitted baseline reference year.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

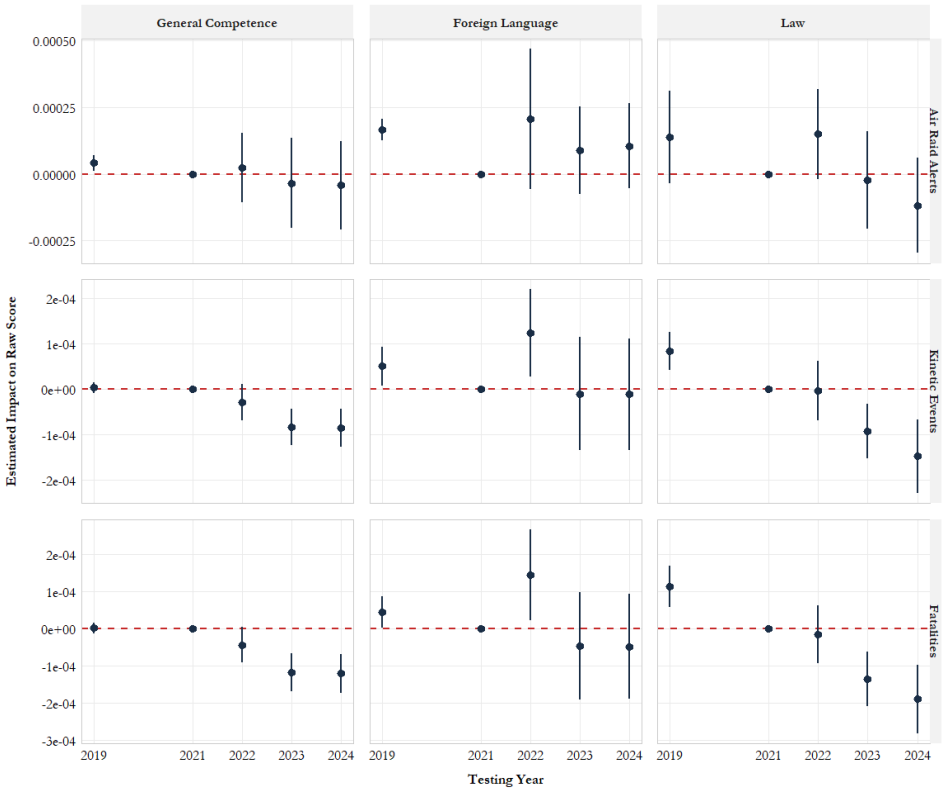


Figure 8. Event-Study Plot

Note: Points represent coefficient estimates; lines represent 95% confidence intervals.

APPENDIX D. FULL TWFE CONTINUOUS DID ESTIMATIONS

Table 18. Full TWFE Continuous DID Estimation - Foreign Language

Panel A: Air Raid Alerts (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts	0.00019	0.00014	0.00013	0.00041***
	(0.00012)	(0.00012)	(0.00013)	(0.00009)
Age	-0.12428***	-0.19108***	-0.46431***	-0.08937***
	(0.00735)	(0.00954)	(0.07715)	(0.00441)
Female	-0.80737***			
	(0.08930)			
Num.Obs.	424130	154436	140897	128797
R2	0.123	0.149	0.130	0.064
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events	0.00000	-0.00004	-0.00002	0.00020***
	(0.00014)	(0.00015)	(0.00014)	(0.00006)
Age	-0.12432***	-0.19097***	-0.46634***	-0.08938***
	(0.00726)	(0.00947)	(0.07743)	(0.00440)
Female	-0.80672***			
	(0.08979)			
Num.Obs.	424130	154436	140897	128797
R2	0.122	0.149	0.130	0.064
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities	-0.00010	-0.00010	-0.00005	0.00010**
	(0.00013)	(0.00016)	(0.00013)	(0.00005)
Age	-0.12418***	-0.19082***	-0.46634***	-0.08941***
	(0.00713)	(0.00941)	(0.07727)	(0.00441)
Female	-0.80667***			
	(0.08983)			
Num.Obs.	424130	154436	140897	128797
R2	0.123	0.149	0.130	0.064
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Continuous DID estimates. Standard errors clustered at the physical testing region level in parentheses.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 19. Full TWFE Continuous DID Estimation - General Competence

Panel A: Air Raid Alerts (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts	-0.00001	0.00002	-0.00001	0.00011*
	(0.00005)	(0.00005)	(0.00004)	(0.00006)
Age	-0.02615***	-0.06846***	-0.04422	-0.02809***
	(0.00355)	(0.00425)	(0.04685)	(0.00320)
Female	-0.76722***			
	(0.03758)			
Num.Obs.	344098	108968	102923	132207
R2	0.103	0.129	0.106	0.065
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events	-0.00011*	-0.00015**	-0.00014**	0.00003
	(0.00005)	(0.00006)	(0.00005)	(0.00004)
Age	-0.02623***	-0.06840***	-0.04225	-0.02812***
	(0.00358)	(0.00428)	(0.04616)	(0.00320)
Female	-0.76818***			
	(0.03743)			
Num.Obs.	344098	108968	102923	132207
R2	0.104	0.129	0.107	0.065
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities	-0.00015***	-0.00016***	-0.00012***	-0.00001
	(0.00003)	(0.00004)	(0.00003)	(0.00003)
Age	-0.02605***	-0.06815***	-0.04303	-0.02813***
	(0.00349)	(0.00417)	(0.04652)	(0.00319)
Female	-0.76747***			
	(0.03753)			
Num.Obs.	344098	108968	102923	132207
R2	0.104	0.129	0.107	0.065
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Continuous DID estimates. Standard errors clustered at the physical testing region level in parentheses.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 20. Full TWFE Continuous DID Estimation - Law

Panel A: Air Raid Alerts (Law)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts	-0.00030*	-0.00026	-0.00027*	-0.00049***
	(0.00015)	(0.00018)	(0.00013)	(0.00008)
Age	-0.13475***	-0.18605***	-0.60592***	0.15963***
	(0.02546)	(0.03007)	(0.11687)	(0.01728)
Female	0.66564***			
	(0.14362)			
Num.Obs.	59714	32168	21475	6071
R2	0.625	0.637	0.615	0.645
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Law)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events	-0.00034***	-0.00040***	-0.00024***	-0.00036**
	(0.00008)	(0.00008)	(0.00008)	(0.00014)
Age	-0.13437***	-0.18527***	-0.60287***	0.15982***
	(0.02496)	(0.02959)	(0.11776)	(0.01710)
Female	0.66675***			
	(0.14366)			
Num.Obs.	59714	32168	21475	6071
R2	0.625	0.637	0.615	0.645
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Law)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities	-0.00042***	-0.00044***	-0.00031*	-0.00042***
	(0.00014)	(0.00015)	(0.00018)	(0.00013)
Age	-0.13420***	-0.18495***	-0.60503***	0.15939***
	(0.02497)	(0.02939)	(0.11716)	(0.01715)
Female	0.66855***			
	(0.14398)			
Num.Obs.	59714	32168	21475	6071
R2	0.625	0.637	0.615	0.645
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Continuous DID estimates. Standard errors clustered at the physical testing region level in parentheses.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

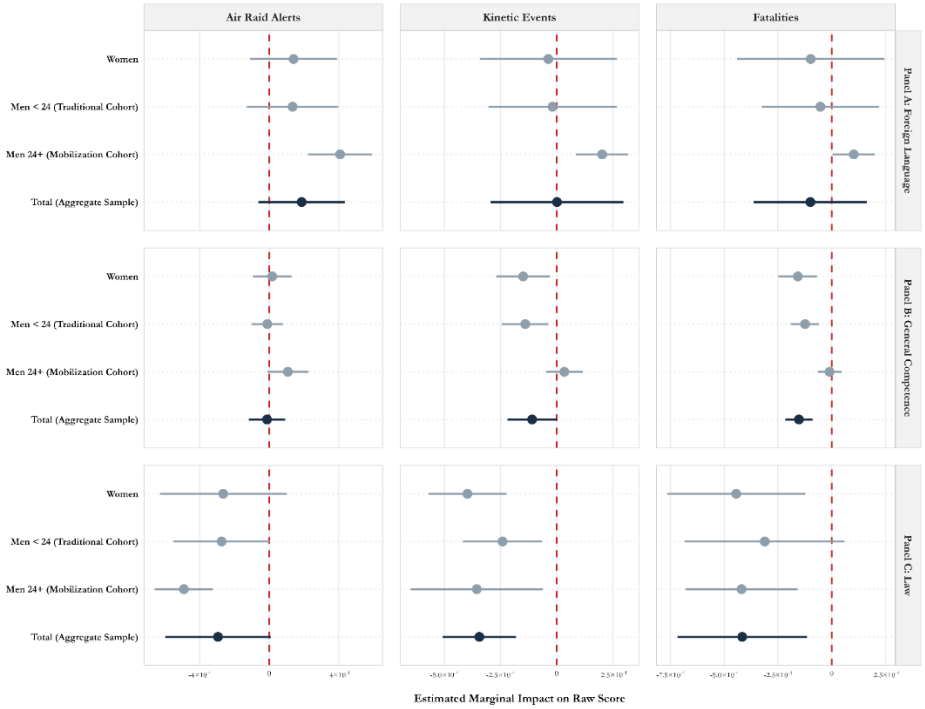


Figure 9. Forest Plots for TWFE Continuous DID Estimation Results for 3 Main Exams

Note: TWFE Continuous DID estimates and 95% CI across all main subjects and all war variables.

APPENDIX E. FULL TWFE OLS ESTIMATIONS (3 MAIN EXAMS)

Table 21. Full TWFE OLS Estimation - Foreign Language

Panel A: Alerts Night Before (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.06867***	-0.07932***	-0.02500	-0.08247*
	(0.02207)	(0.02158)	(0.02870)	(0.04017)
Age	-0.08813***	-0.13492***	-0.04679	-0.08758***
	(0.00514)	(0.00565)	(0.07994)	(0.00479)
Female	-0.88627***			
	(0.08084)			
Num.Obs.	284885	79883	85377	119625
R2	0.061	0.071	0.046	0.062
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.05256***	0.03250	0.02700*	0.06849***
	(0.00911)	(0.02205)	(0.01500)	(0.02258)
Age	-0.08820***	-0.13547***	-0.04674	-0.08752***
	(0.00489)	(0.00538)	(0.07961)	(0.00473)
Female	-0.88735***			
	(0.08137)			
Num.Obs.	284885	79883	85377	119625
R2	0.061	0.070	0.046	0.062
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.00122	-0.00107	0.00263**	0.00613***
	(0.00099)	(0.00161)	(0.00125)	(0.00074)
Age	-0.08865***	-0.13572***	-0.04718	-0.08800***
	(0.00535)	(0.00566)	(0.08022)	(0.00493)
Female	-0.88746***			
	(0.08145)			
Num.Obs.	284885	79883	85377	119625
R2	0.061	0.070	0.046	0.062
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates for 2022-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 22. Full TWFE OLS Estimation - General Competence

Panel A: Alerts Night Before (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.01086	0.00264	0.02065	0.07709
	(0.02724)	(0.01939)	(0.02256)	(0.05738)
Age	-0.02464***	-0.06531***	-0.01525	-0.02852***
	(0.00318)	(0.00373)	(0.05225)	(0.00326)
Female	-0.87049***			
	(0.04500)			
Num.Obs.	317315	94223	93306	129786
R2	0.040	0.043	0.032	0.039
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.00267	-0.00425	-0.01221	-0.00116
	(0.00883)	(0.00532)	(0.00894)	(0.02915)
Age	-0.02457***	-0.06533***	-0.01538	-0.02837***
	(0.00327)	(0.00377)	(0.05233)	(0.00332)
Female	-0.87044***			
	(0.04495)			
Num.Obs.	317315	94223	93306	129786
R2	0.040	0.043	0.032	0.038
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.00682***	0.01040***	0.01806***	0.00729***
	(0.00150)	(0.00227)	(0.00316)	(0.00129)
Age	-0.02506***	-0.06567***	-0.01719	-0.02860***
	(0.00360)	(0.00392)	(0.05325)	(0.00332)
Female	-0.86918***			
	(0.04497)			
Num.Obs.	317315	94223	93306	129786
R2	0.040	0.043	0.032	0.038
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates for 2022-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 23. Full TWFE OLS Estimation - Law

Panel A: Alerts Night Before (Law)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.12022*	-0.12788*	-0.09068	-0.15629**
	(0.05950)	(0.06787)	(0.06585)	(0.05718)
Age	-0.19017***	-0.22593***	-0.65578***	0.14090***
	(0.02319)	(0.02905)	(0.17954)	(0.02404)
Female	0.67415***			
	(0.16061)			
Num.Obs.	32932	17424	11858	3650
R2	0.695	0.711	0.683	0.689
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Law)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.05075	0.04401	0.04214	0.06463*
	(0.04698)	(0.06528)	(0.04258)	(0.03657)
Age	-0.19040***	-0.22632***	-0.65960***	0.14021***
	(0.02348)	(0.02948)	(0.18016)	(0.02414)
Female	0.67669***			
	(0.16148)			
Num.Obs.	32932	17424	11858	3650
R2	0.695	0.711	0.683	0.689
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Law)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.00273	0.01837	-0.01324	-0.01590
	(0.02259)	(0.02439)	(0.02421)	(0.01687)
Age	-0.19115***	-0.22751***	-0.65630***	0.14063***
	(0.02352)	(0.03019)	(0.17935)	(0.02381)
Female	0.67743***			
	(0.16116)			
Num.Obs.	32932	17424	11858	3650
R2	0.695	0.711	0.683	0.689
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates for 2022-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

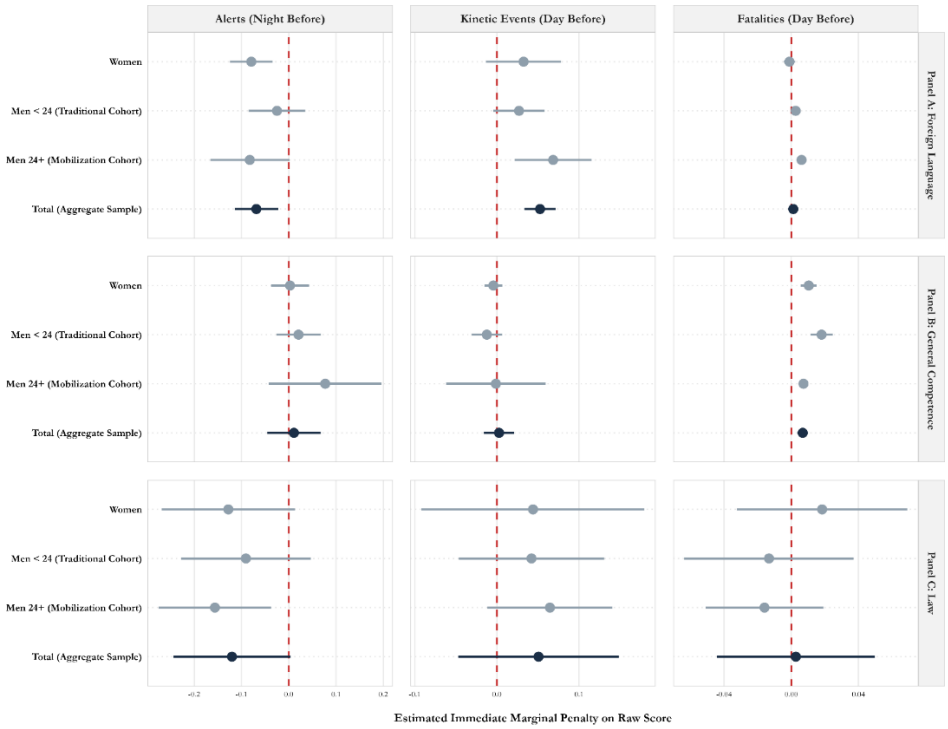


Figure 10. Forest Plots for TWFE OLS Estimation Results for 3 Main Exams

Note: TWFE OLS estimates and 95% CI across all main subjects and all war variables.

APPENDIX F. ESTIMATION RESULTS FOR OTHER UPEs

Table 24. TWFE OLS Estimates for Other UPEs

Exposure	Aggregate	Women	Men < 24	Men 24+
Psychology				
Alerts (Night Before)	-0.1367***	-0.1554***	-0.1083	-0.1991
	(0.0374)	(0.0530)	(0.1659)	(0.1962)
Kinetic Events (Day Before)	0.0033	-0.0538	0.0864	0.1934
	(0.0481)	(0.0437)	(0.1227)	(0.2596)
Fatalities (Day Before)	0.0475	-0.0053	-0.0575	0.0361
	(0.2764)	(0.3371)	(0.2377)	(0.6447)
Finance				
Alerts (Night Before)	0.1471	0.0106	0.4042**	0.0179
	(0.1113)	(0.1228)	(0.1923)	(0.3053)
Kinetic Events (Day Before)	-0.1538***	-0.2102***	-0.0519	-0.4199***
	(0.0274)	(0.0565)	(0.0736)	(0.0766)
Fatalities (Day Before)	-0.1956**	-0.4037***	-0.1476***	0.0926
	(0.0793)	(0.0599)	(0.0330)	(0.4688)
Management				
Alerts (Night Before)	0.0860**	0.1171**	0.1062***	0.0486
	(0.0409)	(0.0558)	(0.0369)	(0.1313)
Kinetic Events (Day Before)	-0.0465	-0.0485	0.0179	-0.0834
	(0.0787)	(0.0761)	(0.0688)	(0.1335)
Fatalities (Day Before)	-0.2017	-0.1177	-0.1024	-0.5601***
	(0.1265)	(0.1615)	(0.1083)	(0.1304)
Economics				
Alerts (Night Before)	-0.0182	-0.3077***	-0.0368	0.4198**
	(0.1171)	(0.1049)	(0.2077)	(0.1867)
Kinetic Events (Day Before)	0.0244	-0.0463	-0.2105*	0.6227***
	(0.0689)	(0.1382)	(0.1069)	(0.1502)
Fatalities (Day Before)	0.0049	-0.0012	0.0019	0.0245***
	(0.0032)	(0.0021)	(0.0045)	(0.0076)
Political Science				
Alerts (Night Before)	-0.6186***	-0.8160***	-0.4320***	-1.6138***
	(0.0981)	(0.1217)	(0.1418)	(0.3020)
Kinetic Events (Day Before)	0.0786	-0.0292	0.4109***	-0.8461***
	(0.0716)	(0.1591)	(0.1125)	(0.2597)
Fatalities (Day Before)	-0.0420***	-0.0607***	-0.0391***	-0.0486***
	(0.0027)	(0.0050)	(0.0048)	(0.0092)

Table 24 – Continued

Exposure	Aggregate	Women	Men < 24	Men 24+
Pedagogy				
Alerts (Night Before)	-0.0694 (0.2257)	0.3169 (0.2928)	-0.6328 (0.4272)	-0.5039 (0.3622)
Kinetic Events (Day Before)	0.1094 (0.1507)	0.3462* (0.1720)	-0.3930 (0.2330)	0.0055 (0.3605)
Fatalities (Day Before)	-0.0235 (0.0761)	-0.0906 (0.0946)	-0.9559*** (0.2363)	0.4019*** (0.0512)
Information Technology				
Alerts (Night Before)	-0.0974 (0.2272)	-0.6793 (1.5522)	-0.3112 (0.4804)	0.3587 (0.3051)
Kinetic Events (Day Before)	-0.1415 (0.1177)	0.5327** (0.1908)	-0.4665*** (0.1196)	0.1583 (0.1145)
Fatalities (Day Before)	-0.4222*** (0.0994)	-0.2496 (0.5464)	-0.5405* (0.2958)	-0.1646** (0.0787)

Note: Pooled TWFE OLS estimates (β) for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 25. Full TWFE OLS Estimation - Psychology

Panel A: Alerts Night Before (Psychology)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.13670***	-0.15541***	-0.10834	-0.19912
	(0.03743)	(0.05297)	(0.16586)	(0.19616)
Age	0.07996***	0.07319***	1.69460***	0.12434***
	(0.01831)	(0.01696)	(0.43178)	(0.04191)
Female	2.15556***			
	(0.54623)			
Num.Obs.	10855	7678	1605	1572
R2	0.105	0.117	0.102	0.091
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Psychology)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.00328	-0.05384	0.08640	0.19335
	(0.04811)	(0.04365)	(0.12269)	(0.25959)
Age	0.07982***	0.07278***	1.69285***	0.12369***
	(0.01831)	(0.01716)	(0.43106)	(0.04154)
Female	2.15014***			
	(0.55069)			
Num.Obs.	10855	7678	1605	1572
R2	0.105	0.117	0.102	0.092
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Psychology)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.04746	-0.00532	-0.05746	0.03614
	(0.27643)	(0.33708)	(0.23769)	(0.64472)
Age	0.07992***	0.07277***	1.68965***	0.12483***
	(0.01852)	(0.01741)	(0.43151)	(0.04171)
Female	2.14868***			
	(0.54937)			
Num.Obs.	10855	7678	1605	1572
R2	0.105	0.117	0.102	0.091
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 26. Full TWFE OLS Estimation - Finance

Panel A: Alerts Night Before (Finance)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.14710	0.01057	0.40417**	0.01795
	(0.11130)	(0.12278)	(0.19230)	(0.30528)
Age	-0.02917	-0.09254***	-0.86270**	0.26996***
	(0.02931)	(0.03041)	(0.37337)	(0.06479)
Female	0.51456			
	(0.55151)			
Num.Obs.	9398	4402	2868	2128
R2	0.065	0.069	0.090	0.082
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Finance)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.15382***	-0.21020***	-0.05190	-0.41991***
	(0.02744)	(0.05652)	(0.07357)	(0.07662)
Age	-0.02949	-0.09209***	-0.85800**	0.27144***
	(0.02915)	(0.03064)	(0.36807)	(0.06459)
Female	0.51325			
	(0.55063)			
Num.Obs.	9398	4402	2868	2128
R2	0.065	0.070	0.088	0.084
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Finance)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.19564**	-0.40366***	-0.14761***	0.09256
	(0.07932)	(0.05990)	(0.03304)	(0.46885)
Age	-0.02998	-0.09160***	-0.85776**	0.26973***
	(0.02923)	(0.03049)	(0.36850)	(0.06471)
Female	0.52144			
	(0.55126)			
Num.Obs.	9398	4402	2868	2128
R2	0.065	0.070	0.088	0.082
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 27. Full TWFE OLS Estimation - Management

Panel A: Alerts Night Before (Management)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.08600**	0.11711**	0.10623***	0.04860
	(0.04095)	(0.05583)	(0.03691)	(0.13132)
Age	-0.04158	-0.17573***	0.02028	0.00631
	(0.03251)	(0.04155)	(0.22908)	(0.02271)
Female	0.49556			
	(0.37192)			
Num.Obs.	31013	11139	10423	9451
R2	0.032	0.041	0.021	0.049
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Management)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.04651	-0.04846	0.01790	-0.08337
	(0.07874)	(0.07606)	(0.06879)	(0.13349)
Age	-0.04159	-0.17476***	0.01931	0.00664
	(0.03225)	(0.04107)	(0.22902)	(0.02281)
Female	0.49866			
	(0.37141)			
Num.Obs.	31013	11139	10423	9451
R2	0.032	0.040	0.021	0.049
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Management)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.20165	-0.11773	-0.10243	-0.56006***
	(0.12648)	(0.16153)	(0.10832)	(0.13038)
Age	-0.04174	-0.17490***	0.01658	0.00654
	(0.03216)	(0.04110)	(0.22909)	(0.02245)
Female	0.49745			
	(0.37187)			
Num.Obs.	31013	11139	10423	9451
R2	0.032	0.041	0.021	0.051
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 28. Full TWFE OLS Estimation - Economics

Panel A: Alerts Night Before (Economics)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.01815	-0.30773***	-0.03678	0.41980**
	(0.11714)	(0.10493)	(0.20767)	(0.18675)
Age	-0.03469	-0.13588***	-0.13492	0.14097***
	(0.02349)	(0.04307)	(0.23928)	(0.03794)
Female	-0.42424			
	(0.62117)			
Num.Obs.	10363	3621	3967	2775
R2	0.073	0.089	0.075	0.089
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Economics)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.02443	-0.04634	-0.21049*	0.62270***
	(0.06890)	(0.13824)	(0.10693)	(0.15017)
Age	-0.03495	-0.13934***	-0.12548	0.14117***
	(0.02341)	(0.04452)	(0.23806)	(0.03781)
Female	-0.42386			
	(0.62091)			
Num.Obs.	10363	3621	3967	2775
R2	0.073	0.087	0.075	0.091
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Economics)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.00486	-0.00116	0.00194	0.02447***
	(0.00322)	(0.00206)	(0.00454)	(0.00758)
Age	-0.03476	-0.13951***	-0.13561	0.14259***
	(0.02336)	(0.04431)	(0.23977)	(0.03831)
Female	-0.41967			
	(0.62079)			
Num.Obs.	10363	3621	3967	2775
R2	0.073	0.087	0.075	0.088
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 29. Full TWFE OLS Estimation - Political Science

Panel A: Alerts Night Before (Political Science)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.61864***	-0.81600***	-0.43202***	-1.61376***
	(0.09813)	(0.12167)	(0.14184)	(0.30197)
Age	-0.27947***	-0.01032	0.96043*	-0.13823*
	(0.05335)	(0.09605)	(0.48371)	(0.06869)
Female	-3.24167***			
	(0.63667)			
Num.Obs.	4180	1963	1568	648
R2	0.085	0.102	0.094	0.077
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Political Science)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.07857	-0.02916	0.41093***	-0.84611***
	(0.07163)	(0.15907)	(0.11249)	(0.25966)
Age	-0.28006***	-0.02302	0.97681*	-0.14965**
	(0.05412)	(0.10113)	(0.48581)	(0.07071)
Female	-3.19618***			
	(0.62348)			
Num.Obs.	4180	1963	1568	648
R2	0.082	0.097	0.094	0.067
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Political Science)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.04205***	-0.06074***	-0.03914***	-0.04862***
	(0.00267)	(0.00500)	(0.00476)	(0.00920)
Age	-0.28008***	-0.01330	0.94971*	-0.13580*
	(0.05424)	(0.09654)	(0.48452)	(0.07160)
Female	-3.22495***			
	(0.63065)			
Num.Obs.	4180	1963	1568	648
R2	0.084	0.100	0.094	0.065
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates for 2023-2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because 1 "Men 24+" observation was dropped due to a singleton fixed effect.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 30. Full TWFE OLS Estimation - Pedagogy

Panel A: Alerts Night Before (Pedagogy)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.06943	0.31693	-0.63285	-0.50387
	(0.22567)	(0.29284)	(0.42718)	(0.36222)
Age	0.00957	0.02282	-0.23438	0.08303
	(0.01894)	(0.01805)	(0.34282)	(0.05244)
Female	4.00899***			
	(0.37846)			
Num.Obs.	15688	9880	2591	3217
R2	0.042	0.027	0.020	0.032
FE: PTReg	X	X	X	X
Panel B: Events Day Before (Pedagogy)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.10938	0.34617*	-0.39297	0.00548
	(0.15066)	(0.17204)	(0.23298)	(0.36048)
Age	0.00960	0.02184	-0.22454	0.08386
	(0.01910)	(0.01840)	(0.34577)	(0.05242)
Female	4.00963***			
	(0.37706)			
Num.Obs.	15688	9880	2591	3217
R2	0.042	0.027	0.020	0.031
FE: PTReg	X	X	X	X
Panel C: Fatalities Day Before (Pedagogy)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.02349	-0.09056	-0.95594***	0.40189***
	(0.07608)	(0.09457)	(0.23631)	(0.05117)
Age	0.00967	0.02172	-0.22243	0.08430
	(0.01897)	(0.01829)	(0.34430)	(0.05241)
Female	4.00806***			
	(0.37782)			
Num.Obs.	15688	9880	2591	3217
R2	0.042	0.027	0.021	0.032
FE: PTReg	X	X	X	X

Note: Pooled TWFE OLS estimates for 2024. Standard errors clustered by physical testing region. Year fixed effects are absent since the exam was conducted in 2024 only.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 31. Full TWFE OLS Estimation - Information Technology

Panel A: Alerts Night Before (Information Technology)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.09745	-0.67931	-0.31121	0.35866
	(0.22718)	(1.55222)	(0.48044)	(0.30511)
Age	-0.04745	-0.38477***	-0.80816**	0.07065*
	(0.06551)	(0.08560)	(0.38562)	(0.03764)
Female	-3.28504***			
	(0.44873)			
Num.Obs.	13438	1206	7626	4606
R2	0.034	0.064	0.028	0.045
FE: PTReg	X	X	X	X
Panel B: Events Day Before (Information Technology)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.14150	0.53268**	-0.46653***	0.15829
	(0.11772)	(0.19076)	(0.11963)	(0.11448)
Age	-0.04676	-0.38613***	-0.79900*	0.07088*
	(0.06598)	(0.08603)	(0.38588)	(0.03749)
Female	-3.28485***			
	(0.44890)			
Num.Obs.	13438	1206	7626	4606
R2	0.034	0.065	0.028	0.045
FE: PTReg	X	X	X	X
Panel C: Fatalities Day Before (Information Technology)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.42217***	-0.24962	-0.54055*	-0.16461**
	(0.09944)	(0.54636)	(0.29580)	(0.07873)
Age	-0.04484	-0.38211***	-0.80128*	0.07246*
	(0.06465)	(0.08561)	(0.38591)	(0.03771)
Female	-3.28402***			
	(0.44809)			
Num.Obs.	13438	1206	7626	4606
R2	0.034	0.064	0.028	0.045
FE: PTReg	X	X	X	X

Note: Pooled TWFE OLS estimates for 2024. Standard errors clustered by physical testing region. Year fixed effects are absent since the exam was conducted in 2024 only.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

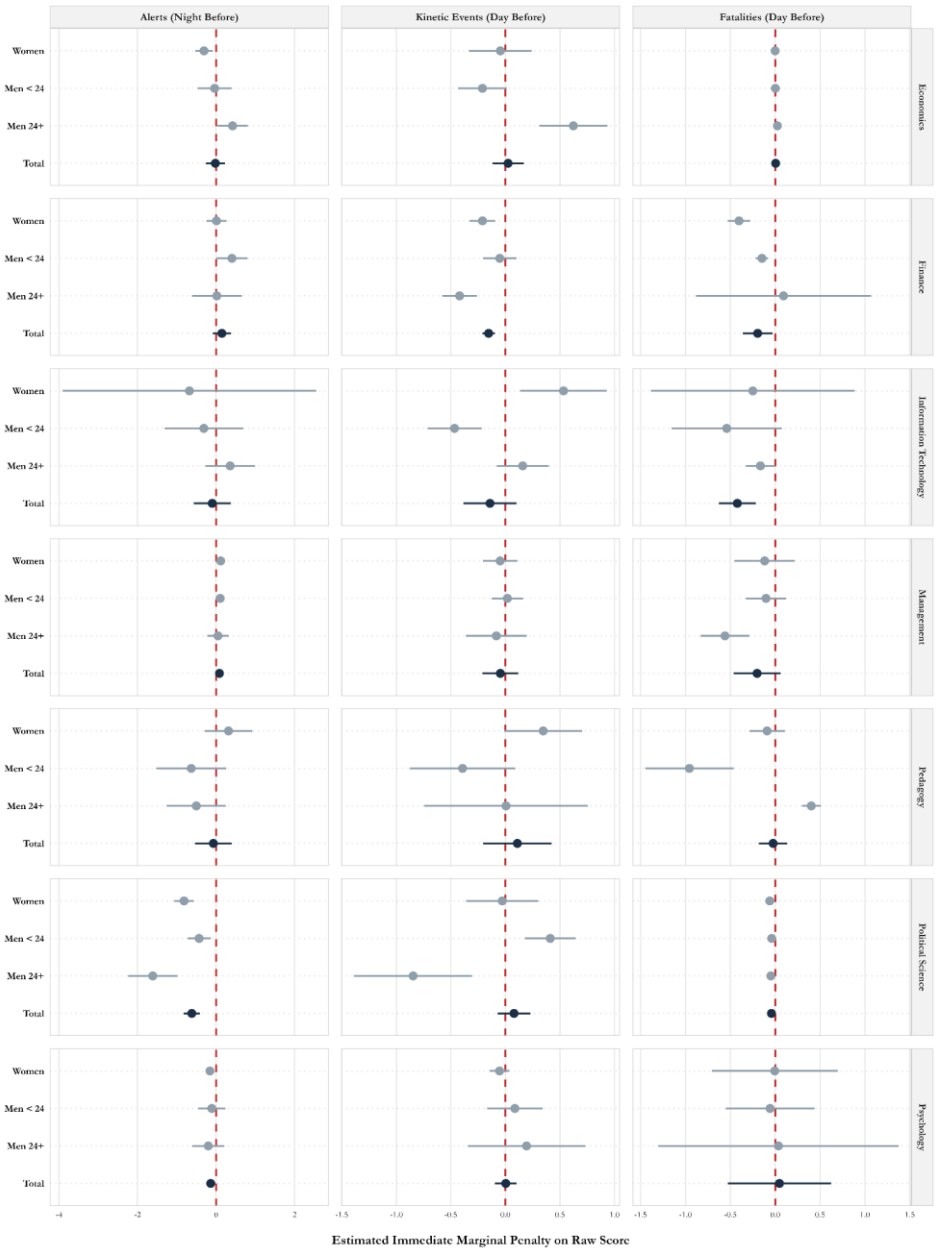


Figure 11. Forest Plots for TWFE OLS Estimation Results for Other UPEs

Note: TWFE OLS estimates and 95% CI across all other subjects and all war variables.

APPENDIX G. TWFE OLS ROBUSTNESS CHECK FULL ESTIMATION RESULTS

Table 32. Full TWFE OLS Estimation with Binary Treatment - Foreign Language

Panel A: Alerts Binary Shock (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts Dummy (Night Before)	-0.29573	-0.22017	-0.56448***	-0.28392
	(0.19372)	(0.18319)	(0.17562)	(0.34647)
Age	-0.08789***	-0.13539***	-0.04162	-0.08746***
	(0.00541)	(0.00592)	(0.07972)	(0.00475)
Female	-0.89269***			
	(0.07929)			
Num.Obs.	284885	79883	85377	119625
R2	0.061	0.070	0.047	0.062
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events Binary Shock (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events Dummy (Day Before)	-0.13981	-0.03702	0.03631	-0.28656
	(0.43499)	(0.31765)	(0.43727)	(0.54076)
Age	-0.08881***	-0.13582***	-0.04672	-0.08809***
	(0.00493)	(0.00552)	(0.07884)	(0.00500)
Female	-0.88650***			
	(0.08361)			
Num.Obs.	284885	79883	85377	119625
R2	0.061	0.070	0.046	0.062
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Table 32 – Continued

Panel C: Fatalities Binary Shock (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities Dummy (Day Before)	0.09477	0.42309	0.11812	-0.39069
	(0.26071)	(0.32654)	(0.30929)	(0.34148)
Age	-0.08850***	-0.13551***	-0.04680	-0.08791***
	(0.00514)	(0.00550)	(0.07979)	(0.00478)
Female	-0.88773***			
	(0.08161)			
Num.Obs.	284885	79883	85377	119625
R2	0.061	0.070	0.046	0.062
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates using binary treatment dummies. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 33. Full TWFE OLS Estimation with Binary Treatment - General Competence

Panel A: Alerts Binary Shock (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts Dummy (Night Before)	0.02319	-0.24869**	-0.23857**	0.26706
	(0.16881)	(0.10142)	(0.10459)	(0.25612)
Age	-0.02464***	-0.06481***	-0.01260	-0.02867***
	(0.00313)	(0.00377)	(0.05292)	(0.00308)
Female	-0.87002***			
	(0.04597)			
Num.Obs.	317315	94223	93306	129786
R2	0.040	0.044	0.032	0.039
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events Binary Shock (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events Dummy (Day Before)	0.01569	0.19719	0.15799	-0.15066
	(0.39993)	(0.31916)	(0.37885)	(0.44847)
Age	-0.02456***	-0.06506***	-0.01472	-0.02851***
	(0.00329)	(0.00376)	(0.05172)	(0.00367)
Female	-0.87056***			
	(0.04496)			
Num.Obs.	317315	94223	93306	129786
R2	0.040	0.043	0.032	0.038
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Binary Shock (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities Dummy (Day Before)	-0.31052	-0.01378	-0.13874	-0.86216
	(0.47445)	(0.34644)	(0.48441)	(0.57850)
Age	-0.02475***	-0.06529***	-0.01540	-0.02858***
	(0.00332)	(0.00378)	(0.05228)	(0.00326)
Female	-0.86999***			
	(0.04491)			
Num.Obs.	317315	94223	93306	129786
R2	0.040	0.043	0.032	0.039
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates using binary treatment dummies. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 34. Full TWFE OLS Estimation with Binary Treatment - Law

Panel A: Alerts Binary Shock (Law)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts Dummy (Night Before)	1.34896***	1.18810***	1.48060**	0.95871**
	(0.39240)	(0.38642)	(0.54359)	(0.40918)
Age	-0.19176***	-0.22796***	-0.69143***	0.13933***
	(0.02362)	(0.02972)	(0.18245)	(0.02402)
Female	0.68210***			
	(0.16379)			
Num.Obs.	32932	17424	11858	3650
R2	0.696	0.711	0.683	0.689
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events Binary Shock (Law)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events Dummy (Day Before)	0.59139	1.11913	-0.11632	0.19921
	(0.62841)	(0.65204)	(0.72641)	(0.41995)
Age	-0.19098***	-0.22669***	-0.65777***	0.14050***
	(0.02390)	(0.03028)	(0.18003)	(0.02399)
Female	0.67861***			
	(0.16137)			
Num.Obs.	32932	17424	11858	3650
R2	0.695	0.711	0.683	0.689
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Binary Shock (Law)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities Dummy (Day Before)	-0.22080	0.06218	-0.53585	-0.86814
	(0.80331)	(0.96280)	(0.77464)	(0.67705)
Age	-0.19102***	-0.22765***	-0.65638***	0.14102***
	(0.02361)	(0.03010)	(0.17965)	(0.02379)
Female	0.67614***			
	(0.16111)			
Num.Obs.	32932	17424	11858	3650
R2	0.695	0.711	0.683	0.689
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE OLS estimates using binary treatment dummies. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

APPENDIX H. FULL TWFE LOGISTIC REGRESSION ESTIMATIONS (3 MAIN EXAMS)

Table 35. Full TWFE Logistic regression (Absenteeism) Estimation - Foreign Language

Panel A: Alerts Night Before (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.01796	0.02263**	0.00574	0.03133
	(0.01385)	(0.00906)	(0.00886)	(0.02053)
Age	0.06079***	0.04838***	0.24924***	0.04828***
	(0.00160)	(0.00256)	(0.03220)	(0.00126)
Female	-0.02544			
	(0.03082)			
Num.Obs.	373427	101665	98469	173293
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.01531***	0.01180	0.00867**	0.02296***
	(0.00577)	(0.01194)	(0.00442)	(0.00586)
Age	0.06093***	0.04863***	0.24918***	0.04834***
	(0.00160)	(0.00261)	(0.03224)	(0.00126)
Female	-0.02462			
	(0.03098)			
Num.Obs.	373427	101665	98469	173293
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00121***	0.00248***	0.00054	-0.00220***
	(0.00025)	(0.00064)	(0.00047)	(0.00007)
Age	0.06093***	0.04844***	0.24916***	0.04837***
	(0.00161)	(0.00262)	(0.03223)	(0.00124)
Female	-0.02508			
	(0.03109)			
Num.Obs.	373427	101665	98469	173293
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE Logistic regression estimates for 2022-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 36. Full TWFE Logistic regression (Absenteeism) Estimation - General Competence

Panel A: Alerts Night Before (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.00086	0.00892	-0.01319	0.01791
	(0.01848)	(0.01222)	(0.01340)	(0.02484)
Age	0.05847***	0.04333***	0.23362***	0.04764***
	(0.00166)	(0.00272)	(0.02996)	(0.00126)
Female	-0.02114			
	(0.02904)			
Num.Obs.	395646	113844	105092	176710
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.02332***	0.02137***	0.01944***	0.02650***
	(0.00352)	(0.00416)	(0.00666)	(0.00521)
Age	0.05863***	0.04363***	0.23367***	0.04771***
	(0.00165)	(0.00272)	(0.03006)	(0.00127)
Female	-0.02084			
	(0.02912)			
Num.Obs.	395646	113844	105092	176710
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00239***	0.00052	-0.00117***	-0.00262***
	(0.00018)	(0.00055)	(0.00041)	(0.00008)
Age	0.05862***	0.04338***	0.23376***	0.04774***
	(0.00164)	(0.00275)	(0.02995)	(0.00124)
Female	-0.02152			
	(0.02916)			
Num.Obs.	395646	113844	105092	176710
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE Logistic regression estimates for 2022-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 37. Full TWFE Logistic regression (Absenteeism) Estimation - Law

Panel A: Alerts Night Before (Law)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.01425	-0.00438	0.04802***	0.03185**
	(0.01631)	(0.02415)	(0.01717)	(0.01561)
Age	0.08451***	0.07238***	0.37208***	0.03424***
	(0.00334)	(0.00528)	(0.06919)	(0.00449)
Female	-0.19460***			
	(0.02306)			
Num.Obs.	39128	20137	13354	5637
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Law)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.00571	-0.00121	-0.01153	-0.01984
	(0.01305)	(0.01419)	(0.00877)	(0.01732)
Age	0.08453***	0.07230***	0.37342***	0.03430***
	(0.00326)	(0.00514)	(0.06921)	(0.00451)
Female	-0.19495***			
	(0.02315)			
Num.Obs.	39128	20137	13354	5637
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Law)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00801**	-0.00844*	-0.00359	-0.01288***
	(0.00379)	(0.00480)	(0.00600)	(0.00211)
Age	0.08460***	0.07222***	0.37388***	0.03431***
	(0.00326)	(0.00519)	(0.06951)	(0.00457)
Female	-0.19667***			
	(0.02331)			
Num.Obs.	39128	20137	13354	5637
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE Logistic regression estimates for 2022-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

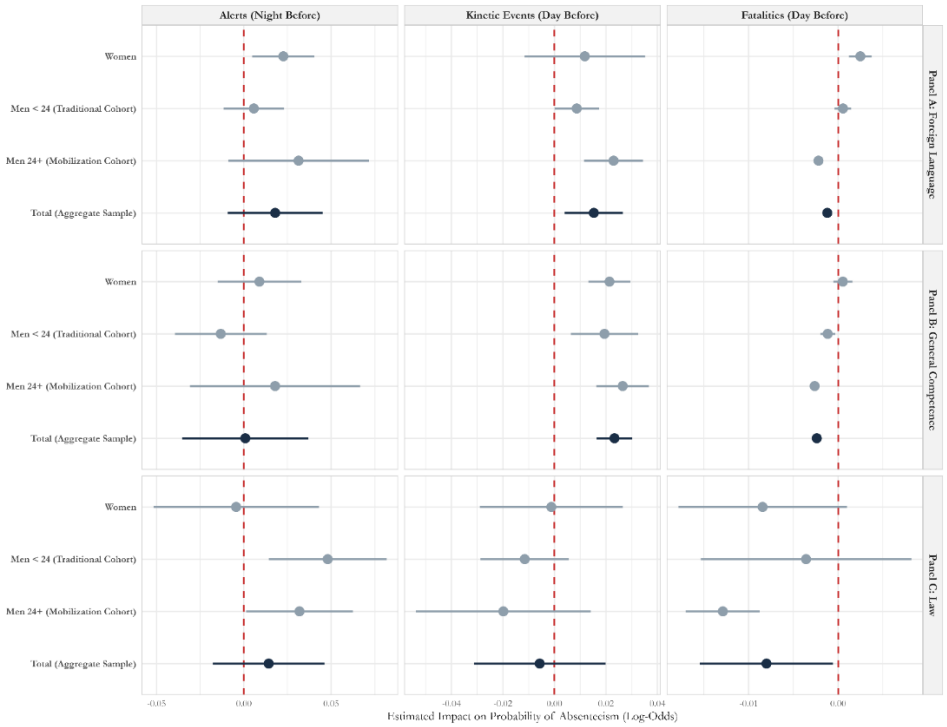


Figure 12. Forest Plots for TWFE Logistic regression (Absenteeism) Estimation Results for 3 Main Exams

Note: TWFE Logistic regression (Absenteeism) estimates and 95% CI across all main subjects and all war variables.

Table 38. Average Marginal Effects for Exam Absenteeism

Exposure	Aggregate	Women	Men < 24	Men 24+
Panel A: Foreign Language				
Alerts (Night Before)	0.0027 (0.0021)	0.0030** (0.0012)	0.0005 (0.0008)	0.0059 (0.0040)
Kinetic Events (Day Before)	0.0023*** (0.0009)	0.0015 (0.0016)	0.0008 (0.0005)	0.0043*** (0.0011)
Fatalities (Day Before)	-0.0002*** (0.0000)	0.0003*** (0.0001)	0.0000 (0.0000)	-0.0004*** (0.0000)
Panel B: General Competence				
Alerts (Night Before)	0.0001 (0.0028)	0.0012 (0.0017)	-0.0013 (0.0015)	0.0034 (0.0048)
Kinetic Events (Day Before)	0.0035*** (0.0005)	0.0030*** (0.0006)	0.0019* (0.0011)	0.0050*** (0.0010)
Fatalities (Day Before)	-0.0004*** (0.0000)	0.0001 (0.0001)	-0.0001 (0.0001)	-0.0005*** (0.0000)
Panel C: Law				
Alerts (Night Before)	0.0017 (0.0020)	-0.0005 (0.0026)	0.0042 (0.0049)	0.0069** (0.0033)
Kinetic Events (Day Before)	-0.0007 (0.0015)	-0.0001 (0.0015)	-0.0010 (0.0017)	-0.0043 (0.0036)
Fatalities (Day Before)	-0.0009** (0.0005)	-0.0009 (0.0005)	-0.0003 (0.0008)	-0.0028*** (0.0004)

Note: Coefficients represent Average Marginal Effects (AME), expressing the absolute change in probability (ranging from 0 to 1). To interpret as percentage points, multiply the estimate by 100. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 39. Full TWFE Logistic regression (Conditional Failure) Estimation - Foreign Language

Panel A: Alerts Night Before (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.03561**	-0.01768	-0.02522	-0.06202***
	(0.01466)	(0.02203)	(0.01692)	(0.02060)
Age	0.02875***	0.03161***	0.02947	0.02838***
	(0.00162)	(0.00230)	(0.03512)	(0.00264)
Female	0.14268***			
	(0.03445)			
Num.Obs.	289589	79009	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.01171**	-0.01765***	-0.00421	-0.01231**
	(0.00465)	(0.00636)	(0.01177)	(0.00535)
Age	0.02864***	0.03144***	0.02933	0.02831***
	(0.00161)	(0.00230)	(0.03511)	(0.00262)
Female	0.14202***			
	(0.03414)			
Num.Obs.	289589	79009	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00302***	0.00123***	-0.00971***	-0.00346***
	(0.00041)	(0.00039)	(0.00137)	(0.00033)
Age	0.02885***	0.03146***	0.03039	0.02842***
	(0.00165)	(0.00229)	(0.03525)	(0.00265)
Female	0.14122***			
	(0.03399)			
Num.Obs.	289589	79009	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE Logistic regression estimates for 2022-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01

Table 40. Full TWFE Logistic regression (Conditional Failure) Estimation - General Competence

Panel A: Alerts Night Before (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.01229	-0.00403	0.00286	-0.00494
	(0.03713)	(0.05001)	(0.03044)	(0.04274)
Age	0.03345***	0.05980***	0.07731	0.02610***
	(0.00600)	(0.01126)	(0.11673)	(0.00817)
Female	-0.32446**			
	(0.14141)			
Num.Obs.	289589	77865	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.02308	-0.03864	-0.00351	-0.03553
	(0.03711)	(0.02614)	(0.02710)	(0.05032)
Age	0.03323***	0.05929***	0.07724	0.02600***
	(0.00600)	(0.01143)	(0.11686)	(0.00816)
Female	-0.32555**			
	(0.14110)			
Num.Obs.	289589	77865	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00501***	0.00016	-0.02781***	0.00023
	(0.00092)	(0.00121)	(0.00339)	(0.00210)
Age	0.03368***	0.05975***	0.07789	0.02609***
	(0.00600)	(0.01143)	(0.11690)	(0.00816)
Female	-0.32593**			
	(0.14070)			
Num.Obs.	289589	77865	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE Logistic regression estimates for 2022-2024. Standard errors clustered by physical testing region. The Num.Obs. discrepancy for 'Women' between Foreign Language and General Competence is due to perfect separation in the latter – Kirovohrad Oblast was automatically dropped because zero women failed the exam in that region. For other subsamples the only perfect separation is in year 2022 (when no fails happened).

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01

Table 41. Full TWFE Logistic regression (Conditional Failure) Estimation - Law

Panel A: Alerts Night Before (Law)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.00937	-0.01733	-0.01591	0.07311
	(0.03480)	(0.04127)	(0.04992)	(0.06641)
Age	0.01669**	0.00573	0.11329	-0.01430
	(0.00790)	(0.01399)	(0.12968)	(0.02236)
Female	-0.41508***			
	(0.12166)			
Num.Obs.	21779	11332	7783	2536
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Law)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.01524	0.02065	-0.03907	-0.04198
	(0.01509)	(0.03624)	(0.03244)	(0.02716)
Age	0.01670**	0.00599	0.11435	-0.01453
	(0.00791)	(0.01411)	(0.12921)	(0.02230)
Female	-0.41524***			
	(0.12149)			
Num.Obs.	21779	11332	7783	2536
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Law)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00510	-0.00167	-0.00673	-0.01010
	(0.00386)	(0.00689)	(0.00594)	(0.00813)
Age	0.01664**	0.00563	0.11365	-0.01470
	(0.00792)	(0.01416)	(0.12959)	(0.02243)
Female	-0.41552***			
	(0.12156)			
Num.Obs.	21779	11332	7783	2536
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: Pooled TWFE Logistic regression estimates for 2022-2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of additional perfect separation in 'Men 24+' subsample apart from year 2022 in regions Cherkasy Oblast, Kirovohrad Oblast, and Mykolaiv Oblast.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01

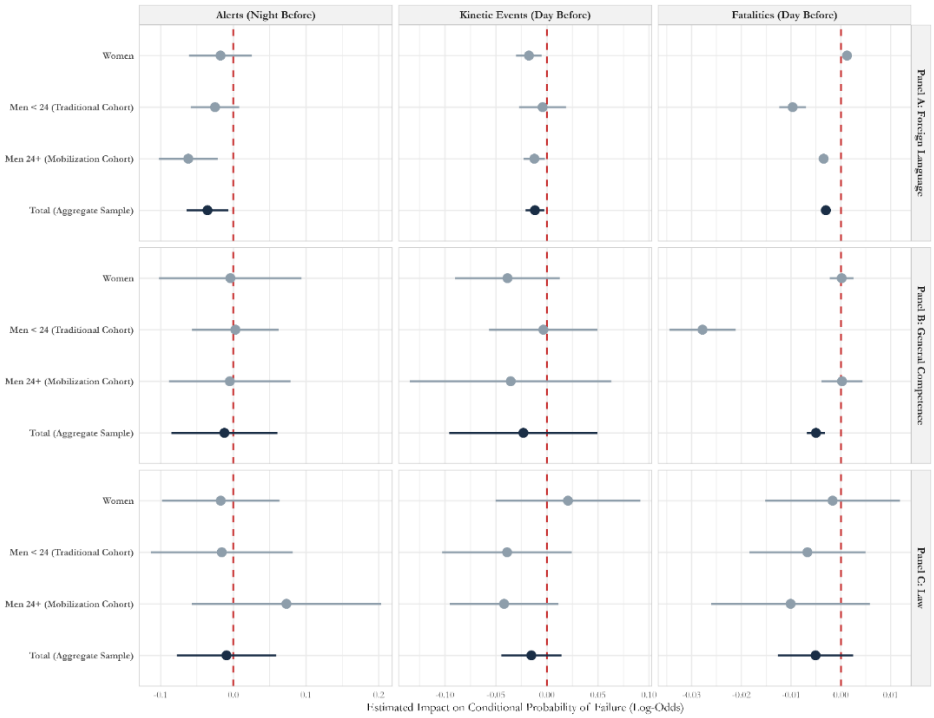


Figure 13. Forest Plots for TWFE Logistic regression (Conditional Failure) Estimation Results for 3 Main Exams

Note: TWFE Logistic regression (Conditional Failure) estimates and 95% CI across all main subjects and all war variables.

Table 42. Average Marginal Effects for Conditional Exam Failure

Exposure	Aggregate	Women	Men < 24	Men 24+
Panel A: Foreign Language				
Alerts (Night Before)	-0.0011***	-0.0005	-0.0005	-0.0025***
	(0.0004)	(0.0006)	(0.0005)	(0.0006)
Kinetic Events (Day Before)	-0.0004***	-0.0005***	-0.0001	-0.0005**
	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Fatalities (Day Before)	-0.0001***	0.0000***	-0.0002	-0.0001***
	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Panel B: General Competence				
Alerts (Night Before)	-0.0000	-0.0000	0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Kinetic Events (Day Before)	-0.0000	-0.0001	-0.0000	-0.0001
	(0.0001)	(0.0000)	(0.0000)	(0.0001)
Fatalities (Day Before)	-0.0000***	0.0000	-0.0000	0.0000
	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Panel C: Law				
Alerts (Night Before)	-0.0002	-0.0003	-0.0003	0.0025
	(0.0007)	(0.0006)	(0.0013)	(0.0035)
Kinetic Events (Day Before)	-0.0003	0.0003	-0.0008	-0.0014
	(0.0003)	(0.0006)	(0.0020)	(0.0010)
Fatalities (Day Before)	-0.0001	-0.0000	-0.0001	-0.0003
	(0.0001)	(0.0001)	(0.0004)	(0.0002)

Note: Coefficients represent Average Marginal Effects (AME), expressing the absolute change in probability (ranging from 0 to 1). To interpret as percentage points, multiply the estimate by 100. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

APPENDIX I. FULL TWFE LOGISTIC REGRESSION ESTIMATIONS (OTHER UPEs)

Table 43. TWFE Logistic regression (Absenteeism) Estimates for Other UPEs

Exposure	Aggregate	Women	Men < 24	Men 24+
Psychology				
Alerts (Night Before)	-0.0035 (0.0080)	-0.0043 (0.0121)	-0.0065 (0.0088)	-0.0120 (0.0397)
Kinetic Events (Day Before)	0.0239*** (0.0080)	0.0298*** (0.0105)	0.0246** (0.0115)	-0.0053 (0.0115)
Fatalities (Day Before)	-0.0135 (0.0258)	-0.0162 (0.0314)	-0.0102 (0.0264)	0.0023 (0.0513)
Finance				
Alerts (Night Before)	-0.0268 (0.0265)	-0.0281 (0.0326)	-0.0567* (0.0308)	-0.0208 (0.0458)
Kinetic Events (Day Before)	0.0077 (0.0156)	0.0185 (0.0165)	0.0079 (0.0143)	-0.0081 (0.0198)
Fatalities (Day Before)	0.0018 (0.0287)	0.0095 (0.0244)	-0.0265* (0.0152)	0.0371 (0.0545)
Management				
Alerts (Night Before)	0.0043 (0.0093)	0.0131 (0.0147)	-0.0068 (0.0114)	-0.0086 (0.0174)
Kinetic Events (Day Before)	-0.0082 (0.0080)	-0.0012 (0.0066)	-0.0191 (0.0126)	-0.0132** (0.0066)
Fatalities (Day Before)	-0.0016 (0.0108)	-0.0212 (0.0276)	-0.0134 (0.0104)	0.0128 (0.0083)
Economics				
Alerts (Night Before)	0.0241** (0.0109)	0.0373*** (0.0112)	0.0409** (0.0208)	-0.0092 (0.0156)
Kinetic Events (Day Before)	0.0098 (0.0151)	0.0431 (0.0313)	0.0285 (0.0259)	-0.0313** (0.0155)
Fatalities (Day Before)	0.0002 (0.0001)	0.0021*** (0.0002)	-0.0009*** (0.0003)	-0.0009*** (0.0003)
Political Science				
Alerts (Night Before)	0.1521* (0.0804)	0.1684*** (0.0590)	0.1054 (0.1237)	0.3389** (0.1531)
Kinetic Events (Day Before)	-0.0132 (0.0251)	-0.0188 (0.0516)	-0.0105 (0.0432)	0.0068 (0.0443)
Fatalities (Day Before)	0.0044*** (0.0004)	0.0062*** (0.0008)	-0.0015 (0.0026)	0.0359*** (0.0130)

Table 43 – Continued

Exposure	Aggregate	Women	Men < 24	Men 24+
Pedagogy				
Alerts (Night Before)	-0.2083***	-0.2461***	-0.2310**	-0.1160**
	(0.0401)	(0.0634)	(0.1003)	(0.0543)
Kinetic Events (Day Before)	0.0775*	0.1668***	0.1146*	-0.0469
	(0.0443)	(0.0518)	(0.0652)	(0.0410)
Fatalities (Day Before)	-0.0202	-0.0334	-0.0340	0.0164
	(0.0355)	(0.0450)	(0.1174)	(0.0235)
Information Technology				
Alerts (Night Before)	-0.0344	0.1828**	-0.2488	-0.0024
	(0.0695)	(0.0861)	(0.2024)	(0.0731)
Kinetic Events (Day Before)	0.0317	0.1232***	0.0034	0.0335
	(0.0308)	(0.0304)	(0.0222)	(0.0481)
Fatalities (Day Before)	-0.0239	0.0851	0.0426	-0.0874***
	(0.0308)	(0.0666)	(0.0334)	(0.0166)

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 44. Full TWFE Logistic regression (Absenteeism) Estimation - Psychology

Panel A: Alerts Night Before (Psychology)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.00350	-0.00427	-0.00649	-0.01195
	(0.00802)	(0.01213)	(0.00875)	(0.03973)
Age	0.03276***	0.02956***	0.04848	0.03251***
	(0.00273)	(0.00299)	(0.03916)	(0.00762)
Female	-0.25554***			
	(0.03403)			
Num.Obs.	16293	11233	2285	2775
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Psychology)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.02389***	0.02975***	0.02455**	-0.00532
	(0.00796)	(0.01051)	(0.01154)	(0.01151)
Age	0.03261***	0.02938***	0.04788	0.03254***
	(0.00270)	(0.00296)	(0.03900)	(0.00763)
Female	-0.25599***			
	(0.03390)			
Num.Obs.	16293	11233	2285	2775
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Psychology)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.01352	-0.01619	-0.01024	0.00228
	(0.02580)	(0.03138)	(0.02641)	(0.05128)
Age	0.03274***	0.02953***	0.04834	0.03251***
	(0.00274)	(0.00302)	(0.03899)	(0.00762)
Female	-0.25528***			
	(0.03401)			
Num.Obs.	16293	11233	2285	2775
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 45. Full TWFE Logistic regression (Absenteeism) Estimation - Finance

Panel A: Alerts Night Before (Finance)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.02683	-0.02811	-0.05674*	-0.02076
	(0.02651)	(0.03262)	(0.03078)	(0.04576)
Age	0.05223***	0.03701***	0.26131***	0.04277***
	(0.00308)	(0.00549)	(0.06940)	(0.00404)
Female	-0.08079			
	(0.05470)			
Num.Obs.	14834	6652	4197	3985
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Finance)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.00770	0.01852	0.00786	-0.00811
	(0.01565)	(0.01648)	(0.01430)	(0.01985)
Age	0.05232***	0.03720***	0.26049***	0.04272***
	(0.00311)	(0.00552)	(0.06989)	(0.00405)
Female	-0.08111			
	(0.05432)			
Num.Obs.	14834	6652	4197	3985
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Finance)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.00179	0.00950	-0.02652*	0.03715
	(0.02867)	(0.02440)	(0.01522)	(0.05451)
Age	0.05233***	0.03709***	0.26096***	0.04268***
	(0.00310)	(0.00554)	(0.06980)	(0.00406)
Female	-0.08132			
	(0.05443)			
Num.Obs.	14834	6652	4197	3985
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 46. Full TWFE Logistic regression (Absenteeism) Estimation - Management

Panel A: Alerts Night Before (Management)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.00426	0.01312	-0.00680	-0.00856
	(0.00927)	(0.01466)	(0.01141)	(0.01741)
Age	0.05688***	0.04202***	0.26170***	0.05026***
	(0.00318)	(0.00369)	(0.06368)	(0.00300)
Female	0.09033**			
	(0.04493)			
Num.Obs.	42747	15136	13003	14608
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Management)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.00817	-0.00116	-0.01910	-0.01320**
	(0.00801)	(0.00659)	(0.01256)	(0.00657)
Age	0.05686***	0.04210***	0.26127***	0.05027***
	(0.00318)	(0.00365)	(0.06376)	(0.00300)
Female	0.09085**			
	(0.04479)			
Num.Obs.	42747	15136	13003	14608
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Management)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00162	-0.02124	-0.01345	0.01282
	(0.01084)	(0.02756)	(0.01036)	(0.00830)
Age	0.05689***	0.04212***	0.26130***	0.05022***
	(0.00319)	(0.00363)	(0.06394)	(0.00301)
Female	0.09044**			
	(0.04475)			
Num.Obs.	42747	15136	13003	14608
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 47. Full TWFE Logistic regression (Absenteeism) Estimation - Economics

Panel A: Alerts Night Before (Economics)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.02412**	0.03731***	0.04086**	-0.00924
	(0.01091)	(0.01123)	(0.02082)	(0.01558)
Age	0.06391***	0.04399***	0.38536***	0.04576***
	(0.00432)	(0.00825)	(0.06143)	(0.00420)
Female	0.11291**			
	(0.05681)			
Num.Obs.	15507	5241	5275	4991
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Economics)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.00978	0.04314	0.02855	-0.03131**
	(0.01513)	(0.03128)	(0.02591)	(0.01553)
Age	0.06405***	0.04444***	0.38519***	0.04586***
	(0.00431)	(0.00835)	(0.06227)	(0.00418)
Female	0.11259**			
	(0.05685)			
Num.Obs.	15507	5241	5275	4991
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Economics)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.00016	0.00215***	-0.00087***	-0.00089***
	(0.00015)	(0.00022)	(0.00029)	(0.00025)
Age	0.06406***	0.04451***	0.38578***	0.04575***
	(0.00430)	(0.00832)	(0.06186)	(0.00422)
Female	0.11280**			
	(0.05671)			
Num.Obs.	15507	5241	5275	4991
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 48. Full TWFE Logistic regression (Absenteeism) Estimation - Political Science

Panel A: Alerts Night Before (Political Science)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.15208*	0.16843***	0.10537	0.33893**
	(0.08036)	(0.05899)	(0.12369)	(0.15310)
Age	0.08744***	0.10023***	0.41251***	0.02812**
	(0.00515)	(0.01077)	(0.07826)	(0.01117)
Female	0.07195			
	(0.06390)			
Num.Obs.	5329	2467	1840	1007
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Political Science)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.01318	-0.01883	-0.01054	0.00679
	(0.02506)	(0.05164)	(0.04316)	(0.04430)
Age	0.08807***	0.10315***	0.41230***	0.02717**
	(0.00512)	(0.01043)	(0.07707)	(0.01093)
Female	0.07115			
	(0.06363)			
Num.Obs.	5329	2467	1840	1007
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Political Science)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.00442***	0.00620***	-0.00147	0.03587***
	(0.00042)	(0.00083)	(0.00259)	(0.01301)
Age	0.08761***	0.10188***	0.41216***	0.02647**
	(0.00515)	(0.01087)	(0.07739)	(0.01060)
Female	0.07100			
	(0.06361)			
Num.Obs.	5329	2467	1840	1007
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of a perfect separation in 'Men < 24' and 'Men 24+' subsample in region Kirovohrad Oblast.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 49. Full TWFE Logistic regression (Absenteeism) Estimation - Pedagogy

Panel A: Alerts Night Before (Pedagogy)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.20825***	-0.24615***	-0.23102**	-0.11601**
	(0.04007)	(0.06338)	(0.10033)	(0.05432)
Age	0.05061***	0.04807***	0.18691***	0.03538***
	(0.00394)	(0.00353)	(0.05504)	(0.00592)
Female	-0.39600***			
	(0.06242)			
Num.Obs.	20895	12663	3191	5041
FE: PTReg	X	X	X	X
Panel B: Events Day Before (Pedagogy)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.07752*	0.16677***	0.11461*	-0.04685
	(0.04433)	(0.05179)	(0.06515)	(0.04101)
Age	0.05105***	0.04848***	0.18911***	0.03562***
	(0.00401)	(0.00349)	(0.05445)	(0.00591)
Female	-0.39798***			
	(0.06220)			
Num.Obs.	20895	12663	3191	5041
FE: PTReg	X	X	X	X
Panel C: Fatalities Day Before (Pedagogy)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.02018	-0.03341	-0.03397	0.01641
	(0.03554)	(0.04504)	(0.11740)	(0.02352)
Age	0.05116***	0.04874***	0.18874***	0.03560***
	(0.00402)	(0.00359)	(0.05446)	(0.00594)
Female	-0.39788***			
	(0.06214)			
Num.Obs.	20895	12663	3191	5041
FE: PTReg	X	X	X	X

Note: TWFE Logistic regression estimates for 2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 50. Full TWFE Logistic regression (Absenteeism) Estimation - Information Technology

Panel A: Alerts Night Before (Information Technology)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.03438	0.18278**	-0.24877	-0.00237
	(0.06954)	(0.08610)	(0.20242)	(0.07313)
Age	0.10483***	0.07270***	0.47103***	0.06602***
	(0.00984)	(0.01450)	(0.10196)	(0.00496)
Female	0.01883			
	(0.07866)			
Num.Obs.	17649	1524	9016	7109
FE: PTReg	X	X	X	X
Panel B: Events Day Before (Information Technology)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.03169	0.12321***	0.00339	0.03354
	(0.03081)	(0.03039)	(0.02220)	(0.04809)
Age	0.10459***	0.07227***	0.47017***	0.06587***
	(0.00998)	(0.01502)	(0.10169)	(0.00511)
Female	0.01765			
	(0.07825)			
Num.Obs.	17649	1524	9016	7109
FE: PTReg	X	X	X	X
Panel C: Fatalities Day Before (Information Technology)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.02390	0.08512	0.04258	-0.08736***
	(0.03084)	(0.06661)	(0.03343)	(0.01660)
Age	0.10489***	0.07238***	0.46919***	0.06625***
	(0.00988)	(0.01474)	(0.10123)	(0.00503)
Female	0.01869			
	(0.07814)			
Num.Obs.	17649	1524	9016	7109
FE: PTReg	X	X	X	X

Note: TWFE Logistic regression estimates for 2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 51. TWFE Logistic regression (Conditional Failure) Estimates for Other UPEs

Exposure	Aggregate	Women	Men < 24	Men 24+
Psychology				
Alerts (Night Before)	0.0417 (0.0824)	-0.0445 (0.0728)	0.0472 (0.4508)	0.5808 (0.6913)
Kinetic Events (Day Before)	0.0502 (0.0397)	0.0516 (0.0589)	0.1954** (0.0828)	-0.0952 (0.0747)
Fatalities (Day Before)	-0.0494 (0.1630)	-0.3623*** (0.1166)	0.3619 (0.4349)	15.6739*** (0.5723)
Finance				
Alerts (Night Before)	0.0071 (0.0209)	-0.0191 (0.0272)	0.0362 (0.0363)	0.0557 (0.0564)
Kinetic Events (Day Before)	0.0131 (0.0181)	0.0202 (0.0131)	0.0189 (0.0216)	-0.0021 (0.0609)
Fatalities (Day Before)	0.0237 (0.0196)	0.0264** (0.0118)	0.0200 (0.0157)	0.0262** (0.0114)
Management				
Alerts (Night Before)	0.0030 (0.0292)	-0.1728 (0.1122)	0.0651 (0.0400)	-0.0081 (0.0550)
Kinetic Events (Day Before)	-0.0074 (0.0328)	-0.0499 (0.0546)	0.0536** (0.0255)	-0.0447 (0.0524)
Fatalities (Day Before)	0.0805* (0.0454)	0.1145 (0.1049)	-0.0204 (0.0237)	0.1461** (0.0720)
Economics				
Alerts (Night Before)	0.0409 (0.0428)	0.0105 (0.0602)	0.0986* (0.0539)	-0.0743 (0.0634)
Kinetic Events (Day Before)	0.0422 (0.0611)	0.0172 (0.0617)	0.0662 (0.0922)	0.0326 (0.1085)
Fatalities (Day Before)	-0.0014*** (0.0004)	-0.0008 (0.0007)	-0.0076* (0.0044)	-0.0097 (0.0078)
Political Science				
Alerts (Night Before)	0.0541 (0.0620)	0.0575 (0.0643)	0.8242* (0.4306)	-116.1044*** (40.2868)
Kinetic Events (Day Before)	0.3649*** (0.0720)	0.3388*** (0.1009)	17.6674*** (1.3061)	3.2277*** (0.8570)
Fatalities (Day Before)	0.0052 (0.0048)	0.0045 (0.0047)	1.7440*** (0.0545)	3.9533*** (0.3804)
Pedagogy				
Alerts (Night Before)	0.2560 (0.1875)	0.1829 (0.1906)	1.1890 (2.9483)	0.2302 (0.2764)
Kinetic Events (Day Before)	-0.1820 (0.1790)	-0.1886 (0.2441)	-0.1796 (0.2042)	-0.1833 (0.1484)
Fatalities (Day Before)	0.0635 (0.0755)	0.0996 (0.0724)	-0.0425 (0.1574)	0.0323 (0.0924)

Table 51 – Continued

Exposure **Information Technology**	**Aggregate**	**Women**	**Men < 24**	**Men 24+**
Alerts (Night Before)	0.0862 (0.1745)	0.3411 (0.3176)	0.1852 (0.1799)	0.0054 (0.2262)
Kinetic Events (Day Before)	-0.2559 (0.2026)	-0.3198 (0.2581)	-0.1172 (0.1906)	-0.3351 (0.2197)
Fatalities (Day Before)	-1.9053*** (0.0055)	-1.9719*** (0.0032)	-1.9005*** (0.0050)	-2.0694*** (0.0047)

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 52. Full TWFE Logistic regression (Conditional Failure) Estimation - Psychology

Panel A: Alerts Night Before (Psychology)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.04167	-0.04447	0.04717	0.58080
	(0.08238)	(0.07279)	(0.45076)	(0.69127)
Age	-0.06160*	-0.09681**	-0.42526*	-0.05157
	(0.03361)	(0.03863)	(0.24329)	(0.04989)
Female	-0.83297***			
	(0.20520)			
Num.Obs.	9955	6529	1057	1116
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Psychology)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.05022	0.05163	0.19537**	-0.09521
	(0.03968)	(0.05889)	(0.08279)	(0.07468)
Age	-0.06212*	-0.09746**	-0.42867*	-0.05337
	(0.03363)	(0.03877)	(0.23224)	(0.05133)
Female	-0.83346***			
	(0.20538)			
Num.Obs.	9955	6529	1057	1116
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Psychology)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.04942	-0.36229***	0.36191	15.67391***
	(0.16298)	(0.11655)	(0.43494)	(0.57233)
Age	-0.06169*	-0.09808**	-0.42180*	-0.05154
	(0.03343)	(0.03865)	(0.24146)	(0.04831)
Female	-0.83305***			
	(0.20467)			
Num.Obs.	9955	6529	1057	1116
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of a perfect separation in each subsample in multiple regions.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 53. Full TWFE Logistic regression (Conditional Failure) Estimation - Finance

Panel A: Alerts Night Before (Finance)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.00712	-0.01908	0.03620	0.05572
	(0.02091)	(0.02721)	(0.03634)	(0.05639)
Age	-0.00984	0.00221	0.04402	-0.04464***
	(0.00606)	(0.00804)	(0.07968)	(0.01486)
Female	-0.08548			
	(0.07640)			
Num.Obs.	10658	4974	3274	2379
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Finance)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.01313	0.02020	0.01891	-0.00206
	(0.01806)	(0.01314)	(0.02158)	(0.06089)
Age	-0.00994*	0.00229	0.04400	-0.04443***
	(0.00600)	(0.00806)	(0.07938)	(0.01501)
Female	-0.08490			
	(0.07656)			
Num.Obs.	10658	4974	3274	2379
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Finance)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.02366	0.02639**	0.02003	0.02620**
	(0.01955)	(0.01175)	(0.01569)	(0.01137)
Age	-0.00992*	0.00224	0.04414	-0.04444***
	(0.00600)	(0.00804)	(0.07930)	(0.01501)
Female	-0.08554			
	(0.07654)			
Num.Obs.	10658	4974	3274	2379
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of a perfect separation in 'Men 24+' subsample in region Sumy Oblast.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 54. Full TWFE Logistic regression (Conditional Failure) Estimation - Management

Panel A: Alerts Night Before (Management)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.00296	-0.17282	0.06506	-0.00806
	(0.02919)	(0.11219)	(0.04003)	(0.05499)
Age	0.01442	0.01276	0.09197	0.01085
	(0.00929)	(0.01470)	(0.15536)	(0.01715)
Female	-0.23834***			
	(0.08714)			
Num.Obs.	31263	11019	10113	9250
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Management)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.00736	-0.04989	0.05357**	-0.04470
	(0.03284)	(0.05457)	(0.02554)	(0.05236)
Age	0.01440	0.01273	0.08783	0.01122
	(0.00933)	(0.01484)	(0.15446)	(0.01707)
Female	-0.23791***			
	(0.08741)			
Num.Obs.	31263	11019	10113	9250
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Management)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.08052*	0.11448	-0.02037	0.14609**
	(0.04537)	(0.10487)	(0.02370)	(0.07201)
Age	0.01447	0.01272	0.09058	0.01047
	(0.00933)	(0.01464)	(0.15530)	(0.01721)
Female	-0.23815***			
	(0.08716)			
Num.Obs.	31263	11019	10113	9250
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of a perfect separation in each subsample in multiple regions.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 55. Full TWFE Logistic regression (Conditional Failure) Estimation - Economics

Panel A: Alerts Night Before (Economics)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.04088	0.01047	0.09863*	-0.07427
	(0.04283)	(0.06024)	(0.05388)	(0.06336)
Age	-0.00536	0.03699**	0.08010	-0.02828
	(0.01367)	(0.01854)	(0.11178)	(0.02930)
Female	-0.16332			
	(0.11340)			
Num.Obs.	10679	3630	4093	2754
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Economics)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.04221	0.01715	0.06618	0.03259
	(0.06110)	(0.06174)	(0.09218)	(0.10853)
Age	-0.00531	0.03691**	0.07822	-0.02848
	(0.01370)	(0.01852)	(0.11282)	(0.02920)
Female	-0.16309			
	(0.11335)			
Num.Obs.	10679	3630	4093	2754
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Economics)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00142***	-0.00083	-0.00760*	-0.00975
	(0.00042)	(0.00067)	(0.00443)	(0.00784)
Age	-0.00512	0.03707**	0.08154	-0.02819
	(0.01366)	(0.01870)	(0.11203)	(0.02926)
Female	-0.16387			
	(0.11282)			
Num.Obs.	10679	3630	4093	2754
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of a perfect separation in each subsample in multiple regions.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 56. Full TWFE Logistic regression (Conditional Failure) Estimation - Political Science

Panel A: Alerts Night Before (Political Science)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.05413	0.05753	0.82415*	-116.10438***
	(0.06199)	(0.06431)	(0.43059)	(40.28677)
Age	0.06058	-0.78755	0.64109	0.09819
	(0.07289)	(1.16065)	(1.07423)	(0.16662)
Female	0.23996			
	(0.62629)			
Num.Obs.	2592	1118	800	299
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Events Day Before (Political Science)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.36493***	0.33877***	17.66741***	3.22769***
	(0.07201)	(0.10086)	(1.30611)	(0.85696)
Age	0.06084	-0.85349	0.62839	0.09494
	(0.07688)	(1.20314)	(1.01084)	(0.16824)
Female	0.26268			
	(0.64705)			
Num.Obs.	2592	1118	800	299
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities Day Before (Political Science)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.00520	0.00448	1.74396***	3.95333***
	(0.00480)	(0.00465)	(0.05453)	(0.38041)
Age	0.06015	-0.77969	0.60953	0.09494
	(0.07214)	(1.15758)	(0.94086)	(0.16824)
Female	0.23738			
	(0.62547)			
Num.Obs.	2592	1118	800	299
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates for 2023-2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of a perfect separation in each subsample in multiple regions.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 57. Full TWFE Logistic regression (Conditional Failure) Estimation - Pedagogy

Panel A: Alerts Night Before (Pedagogy)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.25603	0.18291	1.18901	0.23019
	(0.18750)	(0.19065)	(2.94828)	(0.27643)
Age	0.03073***	0.03684***	0.21683	0.02918
	(0.01063)	(0.00964)	(0.25309)	(0.02804)
Female	-0.65475***			
	(0.18328)			
Num.Obs.	15402	9333	2043	2787
FE: PTReg	X	X	X	X
Panel B: Events Day Before (Pedagogy)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.18198	-0.18855	-0.17962	-0.18329
	(0.17899)	(0.24415)	(0.20423)	(0.14838)
Age	0.03032***	0.03669***	0.21103	0.02891
	(0.01054)	(0.00943)	(0.25460)	(0.02815)
Female	-0.65242***			
	(0.18260)			
Num.Obs.	15402	9333	2043	2787
FE: PTReg	X	X	X	X
Panel C: Fatalities Day Before (Pedagogy)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	0.06353	0.09964	-0.04248	0.03234
	(0.07552)	(0.07240)	(0.15739)	(0.09239)
Age	0.03016***	0.03677***	0.20931	0.02865
	(0.01062)	(0.00949)	(0.25511)	(0.02821)
Female	-0.65087***			
	(0.18259)			
Num.Obs.	15402	9333	2043	2787
FE: PTReg	X	X	X	X

Note: TWFE Logistic regression estimates for 2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of a perfect separation in each subsample in multiple regions.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 58. Full TWFE Logistic regression (Conditional Failure) Estimation - Information Technology

Panel A: Alerts Night Before (Information Technology)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.08621	0.34106	0.18516	0.00537
	(0.17454)	(0.31763)	(0.17988)	(0.22618)
Age	0.02207**	-0.04731*	0.33239**	0.03359*
	(0.00958)	(0.02653)	(0.16820)	(0.01939)
Female	0.05722			
	(0.23891)			
Num.Obs.	13527	987	7485	4597
FE: PTReg	X	X	X	X
Panel B: Events Day Before (Information Technology)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.25587	-0.31983	-0.11717	-0.33509
	(0.20261)	(0.25806)	(0.19059)	(0.21969)
Age	0.02250**	-0.04445*	0.33238**	0.03404*
	(0.00961)	(0.02697)	(0.16819)	(0.01938)
Female	0.05874			
	(0.23918)			
Num.Obs.	13527	987	7485	4597
FE: PTReg	X	X	X	X
Panel C: Fatalities Day Before (Information Technology)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-1.90533***	-1.97189***	-1.90055***	-2.06943***
	(0.00548)	(0.00322)	(0.00495)	(0.00473)
Age	0.02253**	-0.04429*	0.33279**	0.03409*
	(0.00955)	(0.02682)	(0.16856)	(0.01920)
Female	0.05778			
	(0.23890)			
Num.Obs.	13527	987	7485	4597
FE: PTReg	X	X	X	X

Note: TWFE Logistic regression estimates for 2024. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of a perfect separation in each subsample in multiple regions.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

APPENDIX J. TWFE LOGISTIC REGRESSION ROBUSTNESS CHECKS FULL ESTIMATION RESULTS

Table 59. Full LPM (Absenteeism) Estimation - Foreign Language

Panel A: Alerts (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.00230	0.00291**	0.00060	0.00593
	(0.00198)	(0.00121)	(0.00086)	(0.00397)
Age	0.01086***	0.00801***	0.02379***	0.00968***
	(0.00024)	(0.00036)	(0.00273)	(0.00023)
Female	0.00122			
	(0.00412)			
Num.Obs.	373427	101665	98469	173293
R2	0.051	0.029	0.009	0.030
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.00210**	0.00175	0.00082	0.00443***
	(0.00078)	(0.00184)	(0.00049)	(0.00118)
Age	0.01088***	0.00805***	0.02379***	0.00969***
	(0.00024)	(0.00037)	(0.00273)	(0.00023)
Female	0.00129			
	(0.00414)			
Num.Obs.	373427	101665	98469	173293
R2	0.051	0.028	0.009	0.030
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00023***	0.00038***	0.00006	-0.00042***
	(0.00004)	(0.00010)	(0.00004)	(0.00001)
Age	0.01089***	0.00802***	0.02379***	0.00970***
	(0.00024)	(0.00037)	(0.00273)	(0.00022)
Female	0.00124			
	(0.00415)			
Num.Obs.	373427	101665	98469	173293
R2	0.051	0.028	0.009	0.030
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: LPM estimates for 2022-2024. Standard errors clustered by physical testing region.
Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 60. Full LPM (Absenteeism) Estimation - General Competence

Panel A: Alerts (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.00018	0.00110	-0.00150	0.00340
	(0.00270)	(0.00179)	(0.00151)	(0.00476)
Age	0.01052***	0.00728***	0.02351***	0.00957***
	(0.00024)	(0.00041)	(0.00274)	(0.00023)
Female	0.00129			
	(0.00410)			
Num.Obs.	395646	113844	105092	176710
R2	0.047	0.026	0.012	0.029
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	0.00356***	0.00354***	0.00240**	0.00524***
	(0.00058)	(0.00066)	(0.00093)	(0.00106)
Age	0.01054***	0.00732***	0.02351***	0.00957***
	(0.00024)	(0.00042)	(0.00275)	(0.00023)
Female	0.00130			
	(0.00411)			
Num.Obs.	395646	113844	105092	176710
R2	0.047	0.027	0.013	0.030
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00041***	0.00008	-0.00014***	-0.00050***
	(0.00002)	(0.00009)	(0.00003)	(0.00002)
Age	0.01055***	0.00729***	0.02353***	0.00959***
	(0.00023)	(0.00042)	(0.00274)	(0.00023)
Female	0.00124			
	(0.00412)			
Num.Obs.	395646	113844	105092	176710
R2	0.047	0.026	0.012	0.029
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: LPM estimates for 2022-2024. Standard errors clustered by physical testing region.
Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 61. Full LPM (Absenteeism) Estimation - Law

Panel A: Alerts (Law)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.00155	-0.00055	0.00486**	0.00584*
	(0.00216)	(0.00246)	(0.00180)	(0.00332)
Age	0.01478***	0.01135***	0.03478***	0.00771***
	(0.00075)	(0.00116)	(0.00608)	(0.00095)
Female	-0.02079***			
	(0.00275)			
Num.Obs.	39128	20137	13354	5637
R2	0.061	0.031	0.014	0.031
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Law)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.00046	-0.00009	-0.00104	-0.00374
	(0.00148)	(0.00151)	(0.00074)	(0.00362)
Age	0.01479***	0.01134***	0.03495***	0.00772***
	(0.00074)	(0.00115)	(0.00608)	(0.00096)
Female	-0.02084***			
	(0.00277)			
Num.Obs.	39128	20137	13354	5637
R2	0.061	0.031	0.014	0.031
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Law)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00095**	-0.00097*	-0.00032	-0.00261***
	(0.00046)	(0.00049)	(0.00055)	(0.00058)
Age	0.01480***	0.01133***	0.03497***	0.00772***
	(0.00073)	(0.00115)	(0.00611)	(0.00097)
Female	-0.02101***			
	(0.00278)			
Num.Obs.	39128	20137	13354	5637
R2	0.061	0.032	0.014	0.033
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: LPM estimates for 2022-2024. Standard errors clustered by physical testing region.
Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 62. Full LPM (Conditional Failure) Estimation - Foreign Language

Panel A: Alerts (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.00029	0.00023	-2e-04	-0.00171**
	(0.00050)	(0.00058)	(0.00040)	(0.00081)
Age	0.00112***	0.00124***	5e-04	0.00123***
	(0.00007)	(0.00012)	(0.00071)	(0.00008)
Female	0.00377***			
	(0.00099)			
Num.Obs.	301188	85284	88273	127631
R2	0.014	0.019	0.014	0.007
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.00019	-0.00043	4e-05	-0.00042**
	(0.00013)	(0.00028)	(0.00019)	(0.00019)
Age	0.00111***	0.00123***	5e-04	0.00123***
	(0.00007)	(0.00012)	(0.00071)	(0.00008)
Female	0.00377***			
	(0.00099)			
Num.Obs.	301188	85284	88273	127631
R2	0.014	0.019	0.014	0.007
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00011***	0.00008**	-0.00023***	-0.00015***
	(0.00002)	(0.00003)	(0.00003)	(1e-05)
Age	0.00112***	0.00123***	0.00052	0.00124***
	(0.00008)	(0.00012)	(0.00072)	(8e-05)
Female	0.00375***			
	(0.00098)			
Num.Obs.	301188	85284	88273	127631
R2	0.014	0.019	0.014	0.007
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: LPM estimates for 2022-2024. Standard errors clustered by physical testing region. Num.Obs is higher in LPM than in the corresponding Logistic regression specification because OLS estimator does not drop perfectly separated groups unlike maximum likelihood estimation.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 63. Full LPM (Conditional Failure) Estimation - General Competence

Panel A: Alerts (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	-0.00002	0.00002	0.00001	-1e-05
	(0.00006)	(7e-05)	(0.00005)	(8e-05)
Age	0.00006***	0.00011***	0.00012	5e-05**
	(0.00002)	(2e-05)	(0.00019)	(2e-05)
Female	-0.00041*			
	(0.00021)			
Num.Obs.	317837	94341	93465	130031
R2	0.000	0.001	0.001	0.001
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.00004	-0.00007**	-0.00001	-6e-05
	(0.00006)	(3e-05)	(0.00004)	(1e-04)
Age	0.00006***	0.00011***	0.00012	5e-05**
	(0.00002)	(2e-05)	(0.00019)	(2e-05)
Female	-0.00041*			
	(0.00021)			
Num.Obs.	317837	94341	93465	130031
R2	0.000	0.001	0.001	0.001
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00001***	0.00000**	-0.00003***	0e+00
	(0.00000)	(0e+00)	(0.00000)	(0e+00)
Age	0.00006***	0.00011***	0.00013	5e-05**
	(0.00002)	(2e-05)	(0.00019)	(2e-05)
Female	-0.00042*			
	(0.00021)			
Num.Obs.	317837	94341	93465	130031
R2	0.000	0.001	0.001	0.001
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: LPM estimates for 2022-2024. Standard errors clustered by physical testing region. Num.Obs is higher in LPM than in the corresponding Logistic regression specification because OLS estimator does not drop perfectly separated groups unlike maximum likelihood estimation.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 64. Full LPM (Conditional Failure) Estimation - Law

Panel A: Alerts (Law)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts (Night Before)	0.00005	0.00019	-0.00029	0.00071
	(0.00030)	(0.00035)	(0.00041)	(0.00105)
Age	0.00027*	0.00007	0.00160	-0.00027
	(0.00014)	(0.00016)	(0.00187)	(0.00043)
Female	-0.00532***			
	(0.00178)			
Num.Obs.	33378	17607	12028	3743
R2	0.009	0.007	0.012	0.020
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Law)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events (Day Before)	-0.00018	-8e-05	-0.00024	-0.00060
	(0.00018)	(0.00019)	(0.00039)	(0.00043)
Age	0.00027*	7e-05	0.00160	-0.00026
	(0.00014)	(0.00016)	(0.00186)	(0.00043)
Female	-0.00532***			
	(0.00178)			
Num.Obs.	33378	17607	12028	3743
R2	0.009	0.007	0.012	0.021
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Law)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities (Day Before)	-0.00013**	-2e-05	-0.00019*	-0.00029*
	(0.00006)	(0.00007)	(0.00010)	(0.00014)
Age	0.00027*	7e-05	0.00162	-0.00026
	(0.00014)	(0.00016)	(0.00186)	(0.00043)
Female	-0.00534***			
	(0.00177)			
Num.Obs.	33378	17607	12028	3743
R2	0.009	0.007	0.012	0.021
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: LPM estimates for 2022-2024. Standard errors clustered by physical testing region. Num.Obs is higher in LPM than in the corresponding Logistic regression specification because OLS estimator does not drop perfectly separated groups unlike maximum likelihood estimation.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 65. Full TWFE Logistic regression (Absenteeism) Estimation with Binary Treatment - Foreign Language

Panel A: Alerts (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts Dummy (Night Before)	0.07208***	0.02966	0.08417**	0.06180**
	(0.02568)	(0.05288)	(0.04125)	(0.03078)
Age	0.06072***	0.04851***	0.24825***	0.04822***
	(0.00156)	(0.00266)	(0.03200)	(0.00126)
Female	-0.02344			
	(0.03065)			
Num.Obs.	373427	101665	98469	173293
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events Dummy (Day Before)	0.24537**	0.20569***	0.13964	0.31962**
	(0.11732)	(0.06619)	(0.12268)	(0.13134)
Age	0.06122***	0.04880***	0.24999***	0.04860***
	(0.00161)	(0.00265)	(0.03187)	(0.00136)
Female	-0.02755			
	(0.03054)			
Num.Obs.	373427	101665	98469	173293
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities Dummy (Day Before)	0.14827***	0.15925	0.13949**	0.15074
	(0.04448)	(0.09838)	(0.06298)	(0.09798)
Age	0.06092***	0.04862***	0.24930***	0.04832***
	(0.00159)	(0.00260)	(0.03222)	(0.00125)
Female	-0.02526			
	(0.03126)			
Num.Obs.	373427	101665	98469	173293
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates using binary treatment dummies. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 66. Full TWFE Logistic regression (Absenteeism) Estimation with Binary Treatment - General Competence

Panel A: Alerts (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts Dummy (Night Before)	0.08369***	0.05581*	0.09403**	0.06984**
	(0.02439)	(0.02954)	(0.03845)	(0.03084)
Age	0.05830***	0.04330***	0.23230***	0.04757***
	(0.00163)	(0.00275)	(0.02963)	(0.00126)
Female	-0.01935			
	(0.02879)			
Num.Obs.	395646	113844	105092	176710
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events Dummy (Day Before)	0.24137**	0.18602**	0.15205	0.31971**
	(0.11870)	(0.08352)	(0.11852)	(0.13064)
Age	0.05881***	0.04359***	0.23433***	0.04795***
	(0.00171)	(0.00279)	(0.02964)	(0.00136)
Female	-0.02345			
	(0.02881)			
Num.Obs.	395646	113844	105092	176710
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities Dummy (Day Before)	0.14593***	0.11572	0.13148**	0.14953
	(0.04937)	(0.07853)	(0.06365)	(0.09888)
Age	0.05855***	0.04346***	0.23370***	0.04768***
	(0.00165)	(0.00273)	(0.02995)	(0.00126)
Female	-0.02155			
	(0.02929)			
Num.Obs.	395646	113844	105092	176710
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates using binary treatment dummies. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 67. Full TWFE Logistic regression (Absenteeism) Estimation with Binary Treatment - Law

Panel A: Alerts (Law)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts Dummy (Night Before)	-0.03105	0.05826	-0.10702	-0.10619
	(0.04985)	(0.07207)	(0.07693)	(0.07393)
Age	0.08461***	0.07233***	0.37517***	0.03445***
	(0.00332)	(0.00534)	(0.06910)	(0.00453)
Female	-0.19488***			
	(0.02324)			
Num.Obs.	39128	20137	13354	5637
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Law)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events Dummy (Day Before)	-0.34259***	-0.41356***	-0.28921**	-0.41600***
	(0.07657)	(0.09134)	(0.12013)	(0.07424)
Age	0.08445***	0.07199***	0.37416***	0.03426***
	(0.00328)	(0.00519)	(0.06914)	(0.00455)
Female	-0.19640***			
	(0.02324)			
Num.Obs.	39128	20137	13354	5637
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Law)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities Dummy (Day Before)	-0.16608	-0.16652	0.00634	-0.33556***
	(0.16772)	(0.20189)	(0.24347)	(0.12493)
Age	0.08459***	0.07226***	0.37332***	0.03446***
	(0.00328)	(0.00522)	(0.06928)	(0.00453)
Female	-0.19585***			
	(0.02329)			
Num.Obs.	39128	20137	13354	5637
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates using binary treatment dummies. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 68. Full TWFE Logistic regression (Conditional Failure) Estimation with Binary Treatment - Foreign Language

Panel A: Alerts (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts Dummy (Night Before)	-0.03336	-0.01555	0.00795	-0.06352
	(0.05814)	(0.09066)	(0.06815)	(0.09688)
Age	0.02876***	0.03156***	0.02920	0.02842***
	(0.00153)	(0.00231)	(0.03550)	(0.00253)
Female	0.14146***			
	(0.03460)			
Num.Obs.	289589	79009	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events Dummy (Day Before)	0.00572	-0.09541	-0.10222	0.08948
	(0.10306)	(0.10232)	(0.08598)	(0.13565)
Age	0.02868***	0.03141***	0.02873	0.02839***
	(0.00167)	(0.00230)	(0.03500)	(0.00272)
Female	0.14224***			
	(0.03437)			
Num.Obs.	289589	79009	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Foreign Language)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities Dummy (Day Before)	0.17111	0.25322***	0.22143*	0.10417
	(0.16319)	(0.07907)	(0.12451)	(0.23094)
Age	0.02874***	0.03157***	0.02967	0.02835***
	(0.00161)	(0.00228)	(0.03514)	(0.00263)
Female	0.14168***			
	(0.03419)			
Num.Obs.	289589	79009	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates using binary treatment dummies. Standard errors clustered by physical testing region.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 69. Full TWFE Logistic regression (Conditional Failure) Estimation with Binary Treatment - General Competence

Panel A: Alerts (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts Dummy (Night Before)	0.10566	0.05843	0.39666**	0.03503
	(0.13385)	(0.31115)	(0.18517)	(0.15524)
Age	0.03318***	0.05969***	0.07274	0.02605***
	(0.00607)	(0.01138)	(0.11698)	(0.00820)
Female	-0.32286**			
	(0.14068)			
Num.Obs.	289589	77865	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events Dummy (Day Before)	-0.45945**	-0.60259	-0.77581**	-0.30494
	(0.22554)	(0.60051)	(0.36174)	(0.45828)
Age	0.03292***	0.05944***	0.07345	0.02597***
	(0.00604)	(0.01132)	(0.11583)	(0.00822)
Female	-0.32340**			
	(0.14212)			
Num.Obs.	289589	77865	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (General Competence)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities Dummy (Day Before)	0.83850*	0.27418	0.10689	1.64293***
	(0.43000)	(0.27879)	(0.10675)	(0.58774)
Age	0.03386***	0.05986***	0.07731	0.02668***
	(0.00604)	(0.01141)	(0.11679)	(0.00830)
Female	-0.32658**			
	(0.14147)			
Num.Obs.	289589	77865	84028	126552
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates using binary treatment dummies. Standard errors clustered by physical testing region. The Num.Obs. discrepancy for 'Women' between Foreign Language and General Competence is due to perfect separation in the latter – Kirovohrad Oblast was automatically dropped because zero women failed the exam in that region. For other subsamples the only perfect separation is in year 2022 (when no fails happened).

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 70. Full TWFE Logistic regression (Conditional Failure) Estimation with Binary Treatment - Law

Panel A: Alerts (Law)				
	Aggregate	Women	Men < 24	Men 24+
Air Raid Alerts Dummy (Night Before)	-0.68937***	-0.61950***	-0.89638***	-0.23634
	(0.15380)	(0.14756)	(0.20211)	(0.32524)
Age	0.01576**	0.00489	0.12540	-0.01450
	(0.00785)	(0.01409)	(0.13108)	(0.02236)
Female	-0.41675***			
	(0.12296)			
Num.Obs.	21779	11332	7783	2536
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel B: Kinetic Events (Law)				
	Aggregate	Women	Men < 24	Men 24+
Kinetic Events Dummy (Day Before)	-0.18062*	-0.08276	-0.25481	-0.33948
	(0.10830)	(0.18954)	(0.28401)	(0.22092)
Age	0.01662**	0.00560	0.11368	-0.01459
	(0.00795)	(0.01416)	(0.12966)	(0.02235)
Female	-0.41540***			
	(0.12156)			
Num.Obs.	21779	11332	7783	2536
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X
Panel C: Fatalities (Law)				
	Aggregate	Women	Men < 24	Men 24+
Fatalities Dummy (Day Before)	-0.08720	-0.06970	-0.01941	-0.27678
	(0.13649)	(0.22344)	(0.25149)	(0.21782)
Age	0.01665**	0.00557	0.11350	-0.01473
	(0.00794)	(0.01423)	(0.12981)	(0.02246)
Female	-0.41538***			
	(0.12159)			
Num.Obs.	21779	11332	7783	2536
FE: PTReg	X	X	X	X
FE: Year	X	X	X	X

Note: TWFE Logistic regression estimates using binary treatment dummies. Standard errors clustered by physical testing region. Subsample Num.Obs do not sum to the aggregate Num.Obs because of additional perfect separation in 'Men 24+' subsample apart from year 2022 in regions Cherkasy Oblast, Kirovohrad Oblast, and Mykolaiv Oblast.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

APPENDIX K. MULTIPLE HYPOTHESIS TEST CORRECTIONS

Table 71. MHT Correction (TWFE Continuous DID)

Demographic	Original p-value	FDR Adjusted p-value
Foreign Language		
Alerts		
Aggregate	0.1341	0.2195
Women	0.2597	0.3895
Men < 24	0.3001	0.4321
Men 24+	0.0002 ***	0.0014 ***
Kinetic Events		
Aggregate	0.9932	0.9932
Women	0.8002	0.8528
Men < 24	0.8942	0.9197
Men 24+	0.0016 ***	0.0074 ***
Fatalities		
Aggregate	0.4398	0.5864
Women	0.5589	0.7186
Men < 24	0.6902	0.8436
Men 24+	0.0424 **	0.0954 *
General Competence		
Alerts		
Aggregate	0.8017	0.8528
Women	0.7523	0.8528
Men < 24	0.8054	0.8528
Men 24+	0.0740 *	0.1332
Kinetic Events		
Aggregate	0.0508 *	0.1053
Women	0.0168 **	0.0433 **
Men < 24	0.0107 **	0.0296 **
Men 24+	0.4003	0.5542
Fatalities		
Aggregate	0.0001 ***	0.0011 ***
Women	0.0014 ***	0.0072 ***
Men < 24	0.0008 ***	0.0049 ***
Men 24+	0.7030	0.8436

Table 71 – Continued

Demographic	Original p-value	FDR Adjusted p-value
Law		
Alerts		
Aggregate	0.0556 *	0.1053
Women	0.1458	0.2283
Men < 24	0.0532 *	0.1053
Men 24+	0.0000 ***	0.0002 ***
Kinetic Events		
Aggregate	0.0003 ***	0.0018 ***
Women	0.0001 ***	0.0013 ***
Men < 24	0.0096 ***	0.0289 **
Men 24+	0.0197 **	0.0473 **
Fatalities		
Aggregate	0.0090 ***	0.0289 **
Women	0.0091 ***	0.0289 **
Men < 24	0.0951 *	0.1631
Men 24+	0.0031 ***	0.0123 **

Note: p-values adjusted using the Benjamini-Hochberg (1995) False Discovery Rate (FDR) procedure. The correction is applied across all tests within this specific hypothesis family.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 72. MHT Correction (TWFE OLS)

Demographic	Original p-value	FDR Adjusted p-value
Foreign Language		
Alerts		
Aggregate	0.0053 ***	0.0238 **
Women	0.0014 ***	0.0072 ***
Men < 24	0.3935	0.5449
Men 24+	0.0528 *	0.1559
Kinetic Events		
Aggregate	0.0000 ***	0.0001 ***
Women	0.1554	0.3290
Men < 24	0.0863 *	0.2064
Men 24+	0.0063 ***	0.0253 **
Fatalities		
Aggregate	0.2343	0.4016
Women	0.5128	0.6153
Men < 24	0.0484 **	0.1559
Men 24+	0.0000 ***	0.0000 ***
General Competence		
Alerts		
Aggregate	0.6942	0.7810
Women	0.8930	0.9306
Men < 24	0.3703	0.5332
Men 24+	0.1934	0.3481
Kinetic Events		
Aggregate	0.7655	0.8351
Women	0.4339	0.5785
Men < 24	0.1862	0.3481
Men 24+	0.9687	0.9687
Fatalities		
Aggregate	0.0002 ***	0.0010 ***
Women	0.0002 ***	0.0010 ***
Men < 24	0.0000 ***	0.0001 ***
Men 24+	0.0000 ***	0.0001 ***

Table 72 – Continued

Demographic	Original p-value	FDR Adjusted p-value
Law		
Alerts		
Aggregate	0.0563 *	0.1559
Women	0.0735 *	0.1889
Men < 24	0.1830	0.3481
Men 24+	0.0125 **	0.0448 **
Kinetic Events		
Aggregate	0.2923	0.4783
Women	0.5075	0.6153
Men < 24	0.3335	0.5221
Men 24+	0.0917 *	0.2064
Fatalities		
Aggregate	0.9048	0.9306
Women	0.4596	0.5910
Men < 24	0.5902	0.6854
Men 24+	0.3566	0.5332

Note: p-values adjusted using the Benjamini-Hochberg (1995) False Discovery Rate (FDR) procedure. The correction is applied across all tests within this specific hypothesis family.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 73. MHT Correction (TWFE Logistic regression (Absenteeism))

Demographic	Original p-value	FDR Adjusted p-value
Foreign Language		
Alerts		
Aggregate	0.1949	0.3190
Women	0.0126 **	0.0301 **
Men < 24	0.5168	0.6002
Men 24+	0.1269	0.2285
Kinetic Events		
Aggregate	0.0080 ***	0.0205 **
Women	0.3230	0.4497
Men < 24	0.0496 **	0.0992 *
Men 24+	0.0001 ***	0.0004 ***
Fatalities		
Aggregate	0.0000 ***	0.0000 ***
Women	0.0001 ***	0.0004 ***
Men < 24	0.2516	0.3780
Men 24+	0.0000 ***	0.0000 ***
General Competence		
Alerts		
Aggregate	0.9628	0.9628
Women	0.4655	0.5651
Men < 24	0.3248	0.4497
Men 24+	0.4709	0.5651
Kinetic Events		
Aggregate	0.0000 ***	0.0000 ***
Women	0.0000 ***	0.0000 ***
Men < 24	0.0035 ***	0.0114 **
Men 24+	0.0000 ***	0.0000 ***
Fatalities		
Aggregate	0.0000 ***	0.0000 ***
Women	0.3425	0.4567
Men < 24	0.0047 ***	0.0141 **
Men 24+	0.0000 ***	0.0000 ***

Table 73 – Continued

Demographic	Original p-value	FDR Adjusted p-value
Law		
Alerts		
Aggregate	0.3824	0.4916
Women	0.8560	0.9064
Men < 24	0.0052 ***	0.0143 **
Men 24+	0.0413 **	0.0874 *
Kinetic Events		
Aggregate	0.6618	0.7219
Women	0.9323	0.9589
Men < 24	0.1886	0.3190
Men 24+	0.2520	0.3780
Fatalities		
Aggregate	0.0347 **	0.0780 *
Women	0.0789 *	0.1494
Men < 24	0.5501	0.6189
Men 24+	0.0000 ***	0.0000 ***

Note: p-values adjusted using the Benjamini-Hochberg (1995) False Discovery Rate (FDR) procedure. The correction is applied across all tests within this specific hypothesis family.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.

Table 74. MHT Correction (TWFE Logistic regression (Conditional Failure))

Demographic	Original p-value	FDR Adjusted p-value
Foreign Language		
Alerts		
Aggregate	0.0151 **	0.0544 *
Women	0.4223	0.7239
Men < 24	0.1360	0.3582
Men 24+	0.0026 ***	0.0134 **
Kinetic Events		
Aggregate	0.0118 **	0.0474 **
Women	0.0055 ***	0.0249 **
Men < 24	0.7203	0.9358
Men 24+	0.0214 **	0.0699 *
Fatalities		
Aggregate	0.0000 ***	0.0000 ***
Women	0.0018 ***	0.0110 **
Men < 24	0.0000 ***	0.0000 ***
Men 24+	0.0000 ***	0.0000 ***
General Competence		
Alerts		
Aggregate	0.7407	0.9358
Women	0.9358	0.9358
Men < 24	0.9250	0.9358
Men 24+	0.9080	0.9358
Kinetic Events		
Aggregate	0.5340	0.8358
Women	0.1393	0.3582
Men < 24	0.8971	0.9358
Men 24+	0.4802	0.7858
Fatalities		
Aggregate	0.0000 ***	0.0000 ***
Women	0.8926	0.9358
Men < 24	0.0000 ***	0.0000 ***
Men 24+	0.9135	0.9358

Table 74 – Continued

Demographic	Original p-value	FDR Adjusted p-value
Law		
Alerts		
Aggregate	0.7877	0.9358
Women	0.6745	0.9358
Men < 24	0.7499	0.9358
Men 24+	0.2709	0.5134
Kinetic Events		
Aggregate	0.3124	0.5624
Women	0.5688	0.8532
Men < 24	0.2284	0.4838
Men 24+	0.1222	0.3582
Fatalities		
Aggregate	0.1866	0.4479
Women	0.8088	0.9358
Men < 24	0.2578	0.5134
Men 24+	0.2140	0.4814

Note: p-values adjusted using the Benjamini-Hochberg (1995) False Discovery Rate (FDR) procedure. The correction is applied across all tests within this specific hypothesis family.

Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.