CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

The effect of weather on high-stakes exam performance: The case of Czech Republic and Slovakia

Master's thesis

Author: Nikola Borýsková

Study program: Economics and Finance

Supervisor: Mgr. Barbara Pertold-Gebická, M.A., Ph.D.

Year of defense: 2025

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, April 28, 2025

Nikola Borýsková

Abstract

This thesis examines the impact of weather on high-stakes exam performance using data from a standardized university admission test conducted across thirty-two locations in the Czech Republic and Slovakia. By linking individual test scores with localized weather conditions using the Ordinary Least Squares estimator, the study identified temperature and rain as the main factors influencing cognitive performance. On the other hand, the heterogeneous effect was not confirmed across types of cognitive ability and genders. Moreover, mediation analysis accompanied by bootstrapping method suggested that the impact of weather is primarily direct rather than operating through sleep patterns or commuting time. The findings highlight the relevance of weather even under mild conditions in exam performance and underscore the need to consider such external factors in the future research and educational policy.

Keywords High-stakes exams, exam performance, weather effect, OLS regression, mediation analysis, bootstrapping method

Title The effect of weather on high-stakes exam performance: The case of Czech Republic and Slovakia

Abstrakt

Tato práce zkoumá vliv počasí na výsledky u důležitých zkoušek pomocí dat ze standardizovaných přijímacích zkoušek na vysoké školy konajících se na třiceti dvou místech v České republice a na Slovensku. Propojením individuálních výsledků testů s meteorologickými podmínkami v místě zkoušky studie ukázala, že teplota a srážky jsou klíčovými faktory ovlivňující kognitivní výkon. Naopak výsledky nepotvrdily heterogenní efekt počasí napříč typem kognitivních schopností a pohlavím. Mediační analýza ukázala, že vliv počasí je přímý, nikoliv zprostředkovaný spánkovými návyky nebo časem stráveným cestou na zkoušku. Výsledky studie poukazují na významnost počasí ve výsledcích u důležitých zkoušek i za mírných podmínek a zdůrazňují potřebu zohlednit tyto externí faktory v budoucím výzkumu a vzdělávací politice.

Klíčová slova Rozhodové testy, úspěšnost u zkoušek, vliv počasí, metoda nejmenších čtverců, mediační analýza, bootstrapping metoda

Název práce Vliv počasí na úspěšnost u důležtých zkoušek: Případ České republiky a Slovenska

Acknowledgments

I am grateful especially to Mgr. Barbara Pertold-Gebická, M.A., Ph.D. for her suggestions for improvements, overall guidance, and time spent on feedback on the thesis. I would also like to express my gratitude to my family for their support during my studies.

Typeset in LaTeXusing the IES Thesis Template.

Bibliographic Record

Borýsková, Nikola: *The effect of weather on high-stakes exam performance: The case of Czech Republic and Slovakia.* Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2025, pages 85. Advisor: Mgr. Barbara Pertold-Gebická, M.A., Ph.D.

Contents

List of Tables

List of Figures

Acronyms

ČHMÚ Czech Hydrometeorological Institute

SHMÚ Slovak Hydrometeorological Institute

OSP General study requirements test

OLS Ordinary Least Squares

SEM Structural Equation model

ACME Average Causal Mediated effect

ADE Average Direct effect

Chapter 1

Introduction

Academic performance plays a critical role in shaping individual opportunities and future success. This is particularly true for high-stakes exams, where outcomes often determine access to prestigious universities, scholarships, or competitive programs. In this context, even minor differences in exam performance can have significant implications for students' educational and professional trajectories. Therefore, understanding the factors that influence exam outcomes is crucial, not only for students and educators but also for policymakers striving to ensure equality in educational systems.

While individual factors such as preparation, adequate level of stress, and intelligence are acknowledged as significant exam determinants of success, external factors have gathered comparatively less attention. Among these external influences, weather is a critical yet often underestimated one. Weather affects nearly every aspect of daily life, impacting productivity, emotional well-being, and physical health. In the context of high-stakes testing, unfavourable weather conditions may create unequal opportunities for test-takers, introducing variability beyond their control. This issue is particularly relevant in countries like the Czech Republic and Slovakia, where standardised exams are a key gateway to higher education, yet the potential influence of weather has not been explored.

In recent years, the growing variability of weather patterns has elevated the importance of understanding its impact. According to the World Meteorological Organization (2024), the past decade has been the warmest in the recorded history, accompanied by the highest levels of rainfall and flooding. In Europe, these trends are no exception (Copernicus 2024). Based on past years' events, it is expected that shifts in seasonal weather, more frequent heatwaves and

unpredictable rain patterns will continue to occur even more frequently. Such changes disrupt daily life and pose challenges for tasks that require sustained cognitive effort and focus, such as exams.

The relationship between weather and performance is complex, involving both direct and indirect effects. On the one hand, weather conditions can directly impair or enhance cognitive functions. On the other hand, weather may influence performance indirectly by affecting broader individual patterns such as sleep or commuting.

Despite the growing evidence of weather's significance, its impact on exam performance remains insufficiently explored, particularly in Europe. Majority of existing literature focuses on other regions, such as the United States or countries in Asia, and often analyses weather variables in isolation, most frequently temperature. Few studies have considered the broader inclusion of environmental factors, such as wind, sunshine and rain, or the indirect pathways through which these variables might shape academic outcomes.

This thesis aims to address this gap by investigating the relationship between weather conditions and exam performance in the Czech Republic and Slovakia, countries where such research has not yet been conducted. Moreover, the study utilizes complex dataset, combining detailed information on exam performance provided by SCIO, a company specialising in standardised university admission tests, and personal information, such as gender, age, sleep, commuting time, with two distinct weather datasets.

The first weather dataset contains averages of weather conditions throughout the exam day, while the second provides hourly information from before and during the exam, all corresponding to each exam's specific date and location. This comprehensive approach enables more robust findings on the effect of weather on exam performance, offering valuable insights into how environmental factors shape academic outcomes. In addition to examining the effects of multiple weather variables, such as temperature, precipitation, wind, sunshine duration, and sky conditions, we also investigate the heterogeneous effects of weather. The structure of the exam, being split into verbal and analytical parts, allows us to examine how weather influences distinct cognitive domains. Moreover, since gender information is also available, the study explores potential differences in weather sensitivity and exam performance between men and women.

Furthermore, the richness of the dataset enables the investigation of indirect pathways through mediation analysis, examining if weather effects are mediated

by sleep duration, sleep quality, or commuting time. This approach helps determine whether weather impacts academic performance directly as well as indirectly through these mediating factors.

The findings of this study indicate that weather conditions play an important role in academic performance. Temperature and rain appeared as the most important factors, showing a consistent positive association with exam performance. Moreover, few clouds were associated with lower exam performance.

On the other hand, no strict evidence was found supporting the heterogeneous effect of weather across different types of cognitive performance or genders. Even though the estimation outcomes suggested that rain and wind have a stronger influence on verbal reasoning, few clouds on analytical reasoning, and temperature shows a positive relationship with both, these differences were found to be statistically insignificant. Similarly, temperature and few clouds showed a stronger effect on women, while rain on men, however, these differences were again not found to be statistically significant.

The mediation analysis further suggested that the influence of weather is primarily direct and not mediated through sleep duration, sleep quality or commuting time. Overall, these findings demonstrate that weather can significantly impact academic performance and highlight the importance of further research to better understand the complex relationship between weather and exam performance.

The thesis is structured as follows. First, the literature review presents previous research to establish the importance and motivation for our study. In this part, the relationship between various weather variables, cognitive performance, and mood is discussed, highlighting how specific weather conditions can influence individuals' emotional states. Additionally, past research focusing on the effect of weather in real-life scenarios, particularly in the context of high-stakes exams, is summarised. Moreover, the hypotheses of our study are formulated considering previous research and our contribution is stated.

The second chapter focuses on describing the employed data. It explains how variables - exam performance, different weather conditions, and additional factors - used in this study were collected and what they represent, along with their summary statistics. The third chapter details the methodology, including the use of Ordinary Least Squares (OLS) estimation, dealing with endogeneity, and application of the mediation analysis with bootstrapping method, which are employed to investigate the direct and indirect effect of weather on exam performance.

In the fourth chapter, the analysis results are presented and discussed. Moreover, robustness checks are performed to ensure the reliability of the findings and the hypothesis fulfillment is commented. Finally, the concluding chapter summarises the key findings, discusses the study's limitations, and offers directions for future research.

Chapter 2

Literature review

Understanding the factors influencing academic performance is crucial, particularly in the context of high-stakes exams, which can significantly shape a person's academic future and career path. Although a person's level of preparedness, along with innate or developed cognitive ability, remains the primary determinant of their success, it is increasingly recognised that external factors such as environmental conditions and weather patterns can also play a significant role.

Weather has been proven to influence numerous aspects of our life, including our emotional state (Denissen *et al.* 2008), level of physical activity (Bélanger *et al.* 2009) or health (Weilnhammer *et al.* 2021). Still, the effect on exam performance only started to gather more attention in recent years.

This literature review aims to provide a comprehensive overview of existing knowledge on the effect of weather on high-stakes exam performance. The chapter is structured into four parts. Firstly, the effect of different weather variables on mood and cognitive performance, the brain's ability to process, store, and retrieve information, is described. This part includes evidence from laboratory and field studies to showcase weather as an influencing factor on people's mental state. The second part focuses on the direct relationship between weather variables and exam performance to highlight the importance of such studies in an educational environment. The third part introduces other variables which can also through weather impact exam scores to bring a complex understanding of different influential factors to an individual's exam performance. Finally, the last part describes the hypothesis of our study formulated based on the previous research.

2.1 Impact of weather on cognitive performance and mood

Weather has been recognised as an essential factor in shaping human behaviour and overall well-being. However, physical comfort is not the only aspect of the human body impacted. Different weather conditions have been shown to affect the mental aspect, such as cognitive abilities or mood, and alter how individuals think, feel, and perform in various situations.

Reasoning, memory, attention, and problem-solving are all crucial cognitive functions in people's daily tasks. Though largely shaped by genetic inheritance and lifelong learning, external conditions, including weather, can also significantly influence these abilities. In recent years, there has been growing evidence of the vital role of external factors on emotions, which are a critical aspect in a person's life, affecting concentration, motivation or readiness to perform specific tasks. This section examines the diverse findings on the relationship between weather and cognitive performance and mood, showing how different weather conditions can either enhance or hinder tasks requiring focus and precision, such as exams.

Recent studies have shown that different temperature levels have both negative and positive effects on cognitive performance. A study conducted by Yin *et al.* (2024) focused on analysing the effect of extremely warm temperatures on the cognitive performance of more than 53.000 participants in China using the difference-in-differences approach. The research found that both short-term (hourly, daily) and long-term (annually) exposure to extremely high temperatures (above 32° Celsius) results in cognitive decline. Additionally, the study revealed that the longer individuals were exposed to such high temperatures, the stronger the negative effect on their cognitive performance was, suggesting a potential cumulative impact of heat exposure on mental functioning, which can have implications, especially for regions facing prolonged heat waves due to climate change. Similarly, cold temperatures have also been associated with a decline in cognitive abilities. Khan *et al.* (2021) analysed the effect of cold temperatures (around 0 and 4° Celsius) on episodic memory of ageing adults in the United States, showcasing a significant negative relationship.

These findings are also supported by meta-analyses focusing on the broader impact of temperature on cognitive performance tasks, such as memory tests, attention or decision-making processes. Falla *et al.* (2021) confirmed the sig-

nificant effect of cold temperature on cognition, reviewing studies on healthy subjects exposed to cold environments in both field and laboratory settings. Additionally, Yeganeh *et al.* (2018) studied research performed only in laboratory settings and suggested that temperature has a bell-shaped effect on cognitive performance. While heat (above 25° Celsius) and cold (below 20° Celsius) showed a decline in studied abilities, an increase was measured in an optimal range between 20 and 25° Celsius. Additionally, both systematic reviews pointed out gender differences, noting that males and females may vary in their sensitivity to temperature changes, potentially due to physiological and hormonal differences.

The findings on temperature's effect on mood show more variability than the ones on cognition, reflecting a complex relationship between temperature, emotional responses, and seasonal variations. Denissen *et al.* (2008) found a positive effect of temperature on mood by analysing daily data from online diaries of individuals in Germany. This positive effect has been confirmed by Bundo *et al.* (2023), which studied participants in Switzerland and found that some of the effect was moderated by sunshine, suggesting that sunlight exposure may play a critical role in elevating mood. On the other hand, Keller *et al.* (2005) conducted a study in the United States and found that the effect of temperature on mood differed across seasons. The findings revealed that temperature positively impacted mood during the spring when people were deprived of warmer temperatures after winter months and were more appreciative of warmer conditions. Different individual characteristics might explain these inconsistencies in findings, as emotionally stable individuals are less prone to be strongly affected by weather changes. (Spasova 2011) Moreover, the temperature has also been found to affect attention negatively. Based on participants of a study conducted by Yoorim Choi and Minjung Kim and Chungyoon Chun (2019), their attention was the worst during warm temperatures, and their lowest measured brain activity was in lower temperature conditions.

While temperature has been widely studied in terms of cognitive abilities and mood, other weather variables, such as rainfall, humidity, sunshine, wind or air pressure, have yet to gather as much attention. These factors, however, can also significantly shape cognitive performance and mental well-being, offering a broader perspective on how environmental conditions impact human functioning.

Pazos *et al.* (2024) looked into the effect of rainfall shocks in early childhood in Peru and discovered evidence of negative influence, especially on memory.

Furthermore, Finlay *et al.* (2020) found out that long-term exposure to rain for older adults in the United States can also significantly impact cognition. Studies showed that cognitive abilities were higher for adults experiencing moderately frequent rainfall (20 - 40% days in a year) than for the ones with infrequent (more than 40% days in a year) or highly frequent (less than 20% days in a year) rain, indicating that a balanced exposure to rainfall may support cognitive maintenance in older age, potentially through mechanisms like reduced stress or more varied daily routines. While these researches focused primarily on the long-term effect of precipitation, Proverbio *et al.* (2018) have suggested that rainfall can also have a short-term effect on cognitive abilities. The study focused on the impact of different sounds on cognition and detected increased performance while listening to intense rain or occasional thunder. Additionally, Lee *et al.* (2014) found out that bad weather, notably rain, positively impacts productivity, probably due to missing cognitive distractions.

Wind has also been identified as a potential factor impacting cognitive performance, particularly in high-exposure environments. A field study by S. Lamb (2014) explored the effect of wind on the performance and comfort of office workers working on high floors of a building in New Zealand, finding a negative impact. Moreover, Heshmati *et al.* (2024) performed a laboratory experiment in England and, alongside well-being and comfort, also focused on the effect of wind on cognition. The results have shown that wind can make higher buildings vibrate, negatively affecting the office worker's cognitive performance, comfort and well-being.

Studies about the effect of air pressure on cognitive abilities were primarily conducted in high altitudes where hypoxia, a low level of oxygen in a person's body, was observed to cause changes in people's cognitive performance. (Pun *et al.* 2018) Nonetheless, Keller *et al.* (2005) found that barometric pressure also influences cognition in normal altitudes. The research focused on studying the effect on young adults in the United States and discovered that during spring when individuals are deprived of good weather in the winter and now spend more time outside, air pressure enhances cognition and mood.

Sunshine has also been shown to affect cognitive performance. Kent *et al.* (2009) focused on the long-term and short-term impact of sun deficiency, where two weeks of isolation caused a decline in cognitive abilities, especially in the case of depressed individuals. On the contrary, same-day isolation has no significant relationship, indicating that the effect of sun exposure might accumulate over time. On the other hand, Gao *et al.* (2018) investigated the effect of sun

exposure by measuring time spent outside of elderly individuals, confirming the positive effect of high sun exposure to low one. This positive relationship between sunlight and cognition may be partly due to the production of vitamin D, which supports various aspects of brain health (Raymond-Lezman & Riskin 2023). On the contrary, its deficiency is associated with cognitive impairments. (Lee *et al.* 2009; Llewellyn *et al.* 2010)

Regarding humidity, its impact on cognitive performance has been mainly studied alongside high temperatures. Humidity can influence how people perceive and tolerate heat, potentially altering cognitive outcomes in warm conditions. Sharma VM (1983) discovered that the effect on cognitive performance has decreased with high temperatures, and a more substantial effect was captured in hot, humid conditions compared to hot, dry ones. These effects have also been confirmed by Tian *et al.* (2020), which exposed individuals in China to different temperature and humidity levels. Results indicated that while higher temperature decreased cognition abilities such as perception, concentration or memory, lowering humidity levels in such hot environments helped to improve overall performance, likely by reducing thermal stress. In addition to these findings, other studies discovered a decrease in concentration (Howarth E 1984) and an increase in sleepiness (Tsutsumi *et al.* 2007) in highly humid environments.

While temperature has been the primary focus of most studies on weather's influence on cognition, the research mentioned above highlights that other variables, such as rainfall, wind, air pressure, humidity, and sunshine, may also significantly impact cognitive performance.

2.2 Impact of weather on exam performance

The previous section showcased the effect of different weather variables on cognitive abilities, highlighting how extreme heat and cold, high humidity, wind and sun deficiency can impair mental processes such as memory, attention, and problem-solving. In contrast, moderate temperatures, low humidity, rainfall, air pressure and sun exposure enhance these cognitive abilities. While these findings provide valuable insights into the general relationship between weather and cognitive performance, the direct impact of weather on academic performance remains an important point of research.

Exams present a unique setting in which students' cognitive abilities are tested under real-world conditions. Therefore, weather affects performance dif-

ferently than what might be apparent in controlled laboratory experiments. This section will provide a comprehensive overview of studies conducted across various regions to understand better how weather influences academic outcomes.

Several studies have examined the effect of temperature on overall test scores and found that temperature plays a significant role in determining how well students perform during exams. A study conducted by (Zivin *et al.* 2015) in the United States on a sample of men and women aged 14-22 years old using a fixed effects model showed that short-term increases in temperature (above 26° Celsius) led to a decrease in math scores, while reading recognition and comprehension did not show any significant relationship, suggesting that specific cognitive abilities, such as those required for quantitative reasoning, may be more sensitive to temperature changes, possibly due to the higher cognitive load associated with these tasks. Similarly, Cho (2017) showed that summer heat (above 34° Celsius) negatively influenced test scores in math and English and confirmed no particular impact on reading exams. This study also highlighted that the negative effect of heat was more substantial in colder cities, where students were less used to high temperatures, while no difference based on gender was found. These findings were further confirmed by Roach & Whitney (2021) in their analysis of standardised math and English/Language Arts tests, corroborating the stronger effect in colder areas.

The impact of weather on high-stakes exams has also been a focus of research, given the significant consequences these exams have for students' academic and professional futures. In China, a study of the national college entrance exam, which includes Chinese, mathematics and foreign language (typically English), found that rising temperatures resulted in lower test scores, with top-performing students being the most affected. (Graff Zivin *et al.* 2020) This is an interesting finding as it suggests that even the students who typically excel in academic settings are not immune to the adverse effects of heat. Furthermore, the study also explored the effect of other weather variables, showcasing a significant positive effect of wind speed and sunshine duration, as opposed to a statistically insignificant effect of rainfall, humidity, and air pressure. Park (2022) confirmed the negative effect of heat on high-stakes exams and suggested that teachers sometimes compensate for these effects by adjusting scores for borderline cases during hot weather. Improved ventilation was shown to mitigate some of the negative impacts of heat (Haverinen-Shaughnessy & Shaughnessy 2015), suggesting that different measures can be

taken and make a meaningful difference in maintaining optimal conditions for testing.

Although many studies related to exam performance focused on warm temperatures, cold weather also plays a significant role in affecting exam performance. A study conducted by Cook & Heyes (2020) in Canada found using a fixed effect model that cold temperatures outside were associated with lower student performance. This finding indicates that the negative effects of cold, perhaps during the commute to school or in inadequately heated buildings, can persist, even if students are not physically exposed to the cold for long periods before the exam. Adaptation again played a notable role, as the study showed that students who lived in colder climates were less affected. This effect has been confirmed in the case of high-stakes exams as well. Suzuki (2024) studied the test scores of college entrance exams in Japan using a fixed effect model and discovered a negative influence of low temperatures and snow. Moreover, a study from Australia supported the idea that adaptability plays a crucial role in moderating the effects of extreme weather, both hot and cold. Students living in regions where extreme weather is more common tend to cope better with it as the negative impact on their exam performance is weaker. (Srivastava *et al.* 2022) This adaptability could include behavioural changes in response to weather, which has been observed as well. For instance, some students adjust their study habits based on weather conditions, allocating less time to study and more time to leisure activities suitable to current weather conditions during periods of cold temperature levels. (Alberto *et al.* 2021)

Nevertheless, not all studies have found a strong connection between weather and exam performance. For example, a study conducted by Li & Patel (2021) in Brazil focusing on high-stakes exam performance reported only a small effect of temperature on test scores. Moreover, there were no significant differences in impact based on the type of test, gender, or race. This finding suggests that the relationship between weather and exam outcomes is not always straightforward. It is necessary to examine the impact of weather on academic outcomes across different countries to understand better when and where these influential effects play an essential role.

While most studies have focused on temperature, other weather variables can have a meaningful impact on exam performance as well. A study examining the effect of extreme rainfall on high-stakes exams in Brazil revealed increased absence from the exam. This effect was stronger for students with lower socioeconomic status, especially the ones with longer commutes. (Prenhaca 2023)

These findings suggest the importance of investigating how different weather variables, such as humidity, wind, and sun exposure, affect exam performance to fill crucial gaps in understanding how a broader range of weather conditions changes students' academic outcomes, offering new insights for policymakers and educators to create an optimal learning environment for everyone.

In conclusion, recent studies suggest that extreme weather, particularly high temperatures, negatively impact exam performance, especially in subjects like math requiring higher cognitive effort. Cold weather also appears to impair performance, though students better adapted to their local climate tend to be less affected. While most research has focused on temperature, more research must be done on the impact of other weather variables, such as humidity, wind, rainfall, air pressure, and sun exposure on students' test performance. As climate change affects weather patterns and makes extreme weather events more frequent, a concern should be raised about the impact on students' performance. These findings show the need for adaptive strategies in educational institutions, including improving school conditions and optimising exam schedules to mitigate the effects of weather on testing days.

2.3 Impact of other variables on exam performance

This part describes other variables that can significantly affect exam performance. Analysing these factors is essential to understanding the complexity of academic performance and highlighting how weather can influence them, indirectly affecting test outcomes. By examining these interrelated elements, we can explore how external conditions shape the outcomes of even the most well-prepared students.

Student preparation and inherent cognitive abilities are critical to exam performance. While intelligence and preparation are often considered stable predictors, stress levels can significantly moderate their impact. Studies have shown that elevated stress negatively affects exam scores, particularly in high-stakes exams where stress is naturally heightened (Heissel *et al.* 2021). The higher levels of stress can be also caused by different reasons than by the exam itself, for example by longer commuting time disrupted by weather or accident. The interaction between these factors reveals that academic success cannot

solely be attributed to intellectual ability but must also account for students' psychological pressures.

Gender has also been recognised as an influential characteristic in academic performance, particularly in relation to test structure and grading. For example, Arthur & Everaert (2012) studied the performance of undergraduate and postgraduate students in Belgium taking accounting exams. While females generally performed better, this advantage diminished when the format shifted to multiple-choice questions, where their performance slightly declined compared to constructed-response questions. Moreover, Funk & Perrone (2023) analysed the effect of gender on point penalties for wrong answers in Microeconomics courses at a Spanish university. Women consistently outperformed men, even when point subtraction for incorrect answers was applied. Interestingly, the stress-inducing nature of point penalties did not negatively impact female students' performance, suggesting a nuanced relationship between gender and stress.

In contrast, studies on high-stakes exams indicate that women often perform worse than men (Cai *et al.* 2019). This disparity may stem from differing perceptions of the test. Men often approach such exams more casually, viewing them as social opportunities, while women tend to view them as more significant, resulting in higher anxiety levels (Leiner *et al.* 2018). These gendered differences in perception underscore the role of psychological factors in shaping exam outcomes. External conditions, such as weather, may further amplify these differences. Research suggests that women are more sensitive to weather variations (Karjalainen 2007), which could significantly impact their performance. However, not all studies agree, with some finding no significant gender-based differences in performance (Cho 2017; Li & Patel 2021). These mixed findings highlight the need for further research to understand better how gender, stress, and weather interact, particularly in the context of high-stakes exams where even small external changes can have a significant impact.

Another critical factor influencing exam performance is sleep quality and duration, which are fundamental to cognitive functioning. Sleep disruptions can lead to poorer academic outcomes. Fakhari *et al.* (2016) examined the effect of sleep duration on the night before an exam among university students in Iran. The study revealed that students who slept six to eight hours (including naps and nighttime) scored higher than those with shorter durations. Conversely, sleep loss was associated with reduced effort, prompting students to choose easier tasks to maintain accuracy (Engle-Friedman *et al.* 2003). Beyond

duration, sleep quality is equally important. Ahrberg *et al.* (2012) studied medical students in the United States and found that poor sleep quality was linked to elevated stress levels, which collectively resulted in poorer test performance. Even though Okano *et al.* (2019) did not find a direct relationship between test scores and the quality or duration of sleep the night before an exam, the study emphasised the importance of long-term sleep patterns. Students with consistent, high-quality sleep routines tended to perform better academically than their peers with erratic sleep habits. This long-term impact of sleep disturbances has been confirmed by Beebe *et al.* (2010) in laboratory settings and a meta-analysis conducted by Wolfson & Carskadon (2003). Poor sleep quality is often tied to elevated stress levels (Herawati & Gayatri 2019), further compounding its negative effects on test scores.

What makes this connection particularly relevant is the role of weather. Hot weather has been shown to disrupt sleep by making it harder for the body to cool down sufficiently for deep sleep stages (Zhu *et al.* 2016; Zheng *et al.* 2019). Conversely, extreme cold can cause physical discomfort, particularly in inadequately heated environments, further interfering with restful sleep (Lan *et al.* 2017). These findings underscore how weather conditions can indirectly influence exam outcomes through their impact on sleep quality and duration.

By examining these factors - stress levels, gender, sleep - alongside their connections to weather, it becomes clear that a complex set of different variables influences academic outcomes. Weather, even though being an external and uncontrollable factor, intersects with various aspects of students' lives and seems to indirectly affect their preparedness, mental state, and physical health. These interactions demonstrate the need for a comprehensive approach, by studying not only direct effects but also students' characteristics and external conditions to better understand the impact behind exam performance.

2.4 Hypothesis

The studies discussed above have demonstrated that weather can influence cognitive performance through direct and indirect mechanisms, as well as in interaction with individual characteristics. Building on these findings the following hypotheses were formulated to examine the relationship between weather conditions and high-stakes exam performance.

Hypothesis 1: *Different weather variables (temperature, precipitation, sun-*

shine, wind and sky condition) have a direct significant effect on exam performance.

This hypothesis is based on the studies discussed in Section 2.1 and Section 2.2, which suggest that weather conditions can directly influence exam performance by affecting cognitive abilities. Previous research indicates that rain and sun exposure enhance cognitive abilities, whereas wind, and lack of sunlight may have the opposite effect. On the other hand, temperature has been found to exert both positive and negative influences depending on its intensity and duration.

Hypothesis 2: *Different weather variables (temperature, rain, sunshine, wind and sky conditions) have an indirect effect on exam performance through sleep length, sleep quality and commuting time.*

This hypothesis assumes, based on literature discussed in Section 2.3, that weather conditions influence exam performance also indirectly by affecting factors such as sleep and commuting time. Adverse weather, such as extreme temperatures or heavy precipitation, may disrupt sleep patterns or reduce sleep quality, impairing cognitive performance. Similarly, weather conditions during the commute, such as wind, rain, or snow, may increase stress, indirectly affecting students' readiness and focus during the exam.

Hypothesis 3: *Weather has a stronger effect on activities with higher cognitive load, meaning the effects proposed in H1 are stronger for the analytical part of the test than the verbal one.*

This hypothesis is based on the assumption that tasks requiring higher cognitive effort, such as analytical reasoning, are more sensitive to external environmental conditions, as research mentioned in Section 2.2 suggested. Analytical tasks often demand sustained attention, problem-solving, and logical processing, which are more likely to be disrupted by external factors. On the contrary, verbal tasks may involve cognitive processes that are less intensive, such as linguistic skills or long-term memory, and thus, be more resilient to the temporary effects of weather variations.

Hypothesis 4: *Women are more sensitive to weather than men, meaning the*

effects proposed in H1 are stronger for women.

This hypothesis is based on evidence from previously mentioned studies in Section 2.3 suggesting that women may exhibit greater sensitivity to environmental conditions, including weather. Physiological and psychological differences, such as hormonal responses and mood regulation, could make women more responsive to changes in weather conditions. The heightened sensitivity provides a basis for expecting more substantial weather effects on exam performance among female participants compared to their male counterparts.

2.5 Our contribution

This thesis contributes to the literature by examining the effect of weather on high-stakes exam performance in the Czech Republic and Slovakia, two countries where this relationship has not been explored before. Additionally, this study utilizes a unique, previously unexplored dataset on exam performance, offering detailed, standardized test scores and participants' characteristics across multiple locations. Unlike prior studies that often focus on temperature alone, this research considers multiple weather variables simultaneously, including temperature, wind, sunshine, rainfall and sky conditions, providing a more comprehensive perspective.

The study's methodological approach and variable selection are motivated by insights from the literature. For instance, key weather variables such as temperature, precipitation, wind and sunshine have been selected based on their availability and prior research highlighting their influence on cognitive abilities and mood. These insights also inspired the decision to investigate both direct and indirect effects of weather on exam performance, such as those mediated through sleep length and quality or commuting time. This research also builds on prior studies by differentiating between cognitive domains within standardised tests and separately analysing the analytical and verbal components. As the literature suggests that weather may have varying impacts on tasks with different cognitive demands, a hypothesis this thesis aims to investigate. The examination of gender differences in sensitivity to weather effects reflects a broader application of findings from existing studies that emphasized the role of personal characteristics.

By using a unique dataset, not publicly shared, from countries - the Czech Republic and Slovakia, where such effects of weather on exam performance have not yet been explored, this thesis fills a gap in the literature. Existing studies often focus on the United States, Asia or Western Europe leaving Central European contexts underrepresented. The study also uses two weather datasets - average weather conditions throughout the exam day and hourly weather data before and during the exam linked to the location of each exam to provide further precision, offering a more granular analysis of weather's direct and indirect effects.

Overall, this thesis leverages insights from existing studies while addressing gaps in geography, and scope. It combines unique and difficult to access data sources and a comprehensive analytical framework to provide nuanced evidence on the role of weather in shaping academic outcomes. In doing so, it contributes to current research and offers practical implications for educators and policymakers and the design of testing environments.

Chapter 3

Data

This section provides an overview of the data sources and variables used in this study to analyse the relationship between weather conditions and exam performance. The data includes three main components: exam scores, weather data and additional information about participants of the exam.

3.1 Exam performance data

The exam data used in this study as a dependent variable was obtained from SCIO, a company specialising in standardised testing for university admissions in the Czech Republic and Slovakia. The results of these exams are used by twenty-four universities in the Czech Republic and twelve in Slovakia as the main requirement for admission selection, making the test a critical part of students' academic and professional future. SCIO tests are high-stakes exams, meaning they significantly impact student's educational opportunities, making this topic even more important to research.

SCIO exams are conducted multiple times each academic year across various locations in the Czech Republic and Slovakia. While there are typically seven exam dates scheduled between December and May, universities primarily consider results from the ones held in February and March. The final exam results are expressed as percentiles, ranking each student in comparison to others, who took the exam on the same day. SCIO offers several different types of tests, such as math, general study requirements (referred to as OSP), English, Spanish, German, biology, chemistry or fundamentals of social sciences. For the purpose of our study, data from the general study requirements test were selected, as this exam has the largest number of participants, is the most

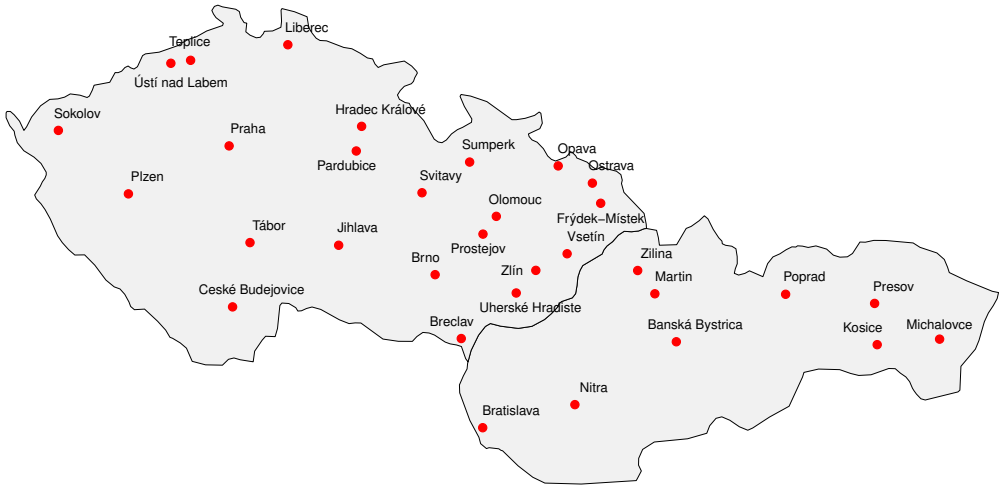
often required type of test for university admissions, and evaluates the overall skills of the student that universities value for admissions. The OSP test provides a well-rounded assessment of students' verbal and analytical abilities, making it highly relevant for studying of external factors that could influence test performance.

In this study, we use data from the first widely accepted term of the exam, which took place on March 7, 2020 - recognised by all participating universities. It is important to note that the exam took place before the onset of the COVID-19 pandemic in the Czech Republic and Slovakia, ensuring that the pandemic did not cause any disruptions. At the time of the exam, only thirty cases had been reported across both countries and the first nation-wide measures were introduced only in the following weeks. The test was held on this specified date across twenty-three cities in the Czech Republic and nine in Slovakia, making it a good representation of both countries. Using data from a single term provides a snapshot of student performance on a single day, which allows us to examine the impact of environmental factors, such as weather, without the confounding effects of students re-taking the test. The exam's standardised timing (9 a.m.) across all locations ensures a consistent baseline, allowing for reliable analysis of weather's potential impact on student performance.

In total, 2864 students participated in the March 7, 2020 exam across thirty-two different cities in Czech Republic and Slovakia which are displayed in Figure 3.1 and the distribution of the number of participants in each location can be seen in Table 3.1. 96% of students (2758 out of 2864) were from the Czech Republic, while the remaining 4% were from Slovakia (106 out of 2864). Although SCIO exams are accepted by 24 universities in the Czech Republic and 12 universities in Slovakia, the proportion of Slovak test-takers remains significantly lower. This disparity is likely due to differences in university admission practices between the two countries. In the Czech Republic, SCIO tests play a major role in the selection process for many universities, making them a widely used part of university applications. In contrast, Slovak universities rely more on their own entrance exams or alternative admission criteria, which may reduce the incentive for Slovak students to take SCIO exams. Due to this, we perform the analysis also separately only for Czech locations and comment on the results in Section 5.3. Prague was the most attended exam place in the Czech Republic, with 701 participants, opposed to Svitavy, with the fewest participants - 21. Bratislava, with 32 participants, was the place with the highest number of attendees in Slovakia, while Michalovce and Poprad, with 1 and 2

participants respectively, were the ones with the lowest. To ensure robust results, places (Michalovce, Poprad) with less than five participants are excluded from the analysis. These statistics reflect the broad geographical reach of the SCIO exam and the diversity of conditions students may have encountered on exam day. This diversity across locations allows for a nuanced examination of regional weather differences and their potential influence on academic outcomes.

Figure 3.1: Exam locations in the Czech Republic and Slovakia



The exam assesses a broad range of cognitive abilities, including analytical thinking, verbal reasoning, and quantitative skills. Each section evaluates specific skills, comprehensively measuring a student's academic capabilities. The standardised nature of the exam allows for consistent evaluation across various student groups, irrespective of their educational background or regional differences, thereby ensuring fair conditions for everyone.

The test is divided into two main segments: verbal and analytical part. The verbal section examines student's ability to understand, analyse, and interpret complex texts. This part is limited to thirty-five minutes and includes thirty-three tasks. It incorporates tasks like reading comprehension, synonym and antonym identification, and critical reasoning. Students must analyse arguments, evaluate assumptions, and draw logical conclusions in this section. An example task might be reading a passage on a social or scientific topic and answering questions that test the understanding of the main idea, logical connections, or the implications of specific statements within the text. Students must process information rapidly, accurately assess meaning, and demonstrate

a deep comprehension of nuanced content while adhering to strict time constraints.

An example exercise (SCIO 2023) from a previous SCIO test illustrates the nature of these questions.

France is experiencing a drop in production at its nuclear power plants and could cut off electricity exports to Italy for two years. Supplies from France account for around five per cent of Italy's electricity consumption. Which of the following statements is implied by the text?
(A) France exports about 5 % of its nuclear power plant output to Italy.
(B) Interrupting French electricity exports for two years will reduce Italian consumption by about one-twentieth.
(C) France's nuclear power plants produce electricity totalling about 5% of Italy's consumption.
(D) The decline in output from French plants threatens Italy's electricity supply, which is about one-twentieth of its consumption. (correct answer)

On the other hand, the analytical reasoning section assesses students' logical thinking and problem-solving skills, which are essential for academic success in higher education. This section is also time-limited; students have fifty minutes to complete thirty-three tasks. This part includes various exercises, such as puzzles, sequences, and mathematical problems. Each task is crafted to challenge students to think methodically, apply deductive reasoning, and solve complex problems without calculators. This section assesses students' ability to interpret relationships, identify patterns, and make reasoned decisions, all of which require a structured approach to problem-solving. Analytical tasks often require students to engage in abstract thinking, drawing on skills that go beyond simple memorisation, thereby providing a robust measure of critical thinking under time pressure. For example, students might be asked to solve a problem involving geometric reasoning or to determine the outcome of a logical scenario based on a set of rules.

Here is an example exercise (SCIO 2023) from a SCIO test that further demonstrates the cognitive demands of this section.

Aleš and Bára competed to see who could solve the puzzle first. They started

simultaneously, and the one who solved it first was the winner. It took Aleš a minutes and Bára b minutes. The following relationship holds:
$a + 15 = 0{,}5b$
Which of the following statements follows from the above relationship?
(A) It took Aleš exactly 30 minutes longer than Bára to solve the puzzle.
(B) Aleš solved the puzzle before Bára, but at most 15 minutes.
(C) Aleš would have won even if it had taken him 15 minutes longer to solve the puzzle than it actually did. (correct answer)
(D) Even if it had taken Aleš 15 minutes less time to solve the puzzle than it actually did, he would have lost.

The exam consists of 66 tasks across both sections, each question contributing to the final score. The scoring system accounts not only for correct answers but also penalises incorrect responses, as points are deducted for wrong answers. This approach encourages students to critically evaluate their certainty in each response, adding a layer of strategy to their decision-making. The entire test lasts one hour and twenty-five minutes, making time management a key factor for students to complete all exercises. Each section is scored separately and then standardised, with the final score calculated as an average of both sections. The scores range from zero to one hundred, enabling a straightforward comparison of student performance across all participants. Students are ranked based on their results, and final scores are presented as percentile placements, indicating each student's relative standing among all test-takers on the same exam date. Table 3.1 shows the distribution of the average test scores for the whole exam and separately for the verbal and analytical parts for each place of the exam in the Czech Republic and Slovakia alongside the summary statistics from the term on March 7, 2020. In Prague, participants achieved the highest average (54.79) score, while Vsetín the lowest one (35.32) in the Czech Republic. In Slovakia, the city with the highest average test score (57) was Michalovce (which will be due to missing weather data and a low number of participants taken out of the dataset) and Banská Bystrica with the second highest (54.36), while in Žilina was the lowest average test score (35.5) in Slovakia. Moreover, the average result for the whole test and separately for the two parts per each place of the exam from the previous year 2019, March 9, was obtained. The average results can be seen in Appendix in Table A.1. The highest average score (57.01) that year was achieved in Tábor, while the lowest one in Vsetín (33.56) and Břeclav (33.96) in the Czech Republic. In Slovakia, the city with

the highest average test score (54.9) was Martin and Banská Bystrica with the lowest average test score (42.2).

Overall, the SCIO exam data provides a comprehensive view of students' cognitive abilities on exam day. This dataset is valuable for its uniqueness in the Czech Republic and Slovakia, as SCIO is the leading provider of university admission testing, and it allows us to examine the effect of weather on the cognitive performance of thousands of students in a high-stakes scenario.

Table 3.1: Average test scores and participants count per each place of the exam in the Czech Republic and Slovakia and overall summary statistics, March 7, 2020

Czech Republic				
City	**Participants**	**Verbal Part**	**Analytical Part**	**Overall Test**
Břeclav	38	36.84	43.34	40.09
Brno	325	46.89	52.96	49.92
České Budějovice	110	57.37	49.16	53.27
Frýdek-Místek	71	43.41	43.15	43.28
Hradec Králové	88	53.99	47.22	50.60
Jihlava	87	46.54	48.76	47.65
Liberec	74	47.46	49.34	48.40
Olomouc	168	35.74	40.66	38.20
Opava	51	34.31	38.86	36.59
Ostrava	244	37.68	43.70	40.69
Pardubice	81	48.77	48.95	48.86
Plzeň	133	40.29	44.08	42.18
Praha	701	55.35	54.23	54.79
Prostějov	26	41.73	42.42	42.08
Sokolov	36	36.81	39.36	38.08
Šumperk	26	32.08	39.65	35.87
Svitavy	21	44.48	47.00	45.74
Tábor	20	50.65	44.25	47.45
Teplice	51	51.76	47.45	49.61
Uherské Hradiště	85	39.39	39.48	39.44
Ústí nad Labem	72	41.17	45.46	43.31
Vsetín	87	32.26	38.39	35.33
Zlín	163	34.02	40.13	37.08
Slovakia				
City	**Participants**	**Verbal Part**	**Analytical Part**	**Overall Test**
Banská Bystrica	7	60.00	48.71	54.36
Bratislava	32	46.56	46.41	46.48
Košice	11	42.18	46.73	44.45
Martin	12	53.92	44.50	49.21
Michalovce	1	84.00	30.00	57.00
Nitra	15	38.80	40.80	39.80
Poprad	2	20.00	78.00	49.00
Prešov	10	50.40	47.40	48.90
Žilina	16	28.25	42.75	35.50
Summary Statistics				
Score Type	**Min**	**Median**	**Mean**	**Max**
Overall Test	0	46.50	46.51	100
Verbal Part	0	44.00	45.60	100
Analytical Part	0	49.00	47.43	100

3.2 Weather data

Two distinct weather datasets were collected to analyse the effect of weather on high-stakes exam performance in the Czech Republic and Slovakia. By providing different levels of granularity, these datasets enable a thorough examination of the potential influence of weather conditions on exam outcomes.

The first dataset was obtained from the Czech and Slovak Hydrometeorological Institutes, which are part of the Ministry of the Environment in their respective countries. These organisations provide measurements, forecasts, evaluations, and archived weather conditions across each country. While the data from Czech hydrometeorological stations is publicly accessible online, the data from Slovak ones require a formal request. As the number of stations from which it is possible to obtain information is limited, two cities in Slovakia (Michalovce and Poprad) had to be erased from our data set. However, this exclusion does not significantly affect the overall sample size, as these cities had only three participants: one in Michalovce and two in Poprad and we would omit them from the analysis due to low number of participants anyways. Moreover, stations in Banská Bystrica, Košice, Nitra do not measure the amount of rainfall, additionally, station in Prešov does not measure rainfall and sunshine hours. Therefore, the number of observations might differ across different models used for the analysis.

Weather information was collected for the specific exam day (March 7, 2020) from weather stations closest to each exam location. The weather variables used in the analysis are daily average measurements of temperature (measured in degrees Celsius), precipitation (in millimeters), sunshine duration (in hours), and wind speed (in kilometres per hour). The temperatures ranged from 2.3 to 7.4° Celsius, in some cases no rain was measured, compared to highest rainfall of 2.8 millimeters. Similarly, sunshine hours ranged from zero to 4.6 hours and wind from 0.4 to 6.3 kilometres per hour. The main weather statistics are summarised in Table 3.2.

While this dataset offers valuable information about weather conditions during the exam day, it has certain limitations. Relying on a daily average can obscure short-term fluctuations in weather that might have localised effects on students' performance. To address this limitation and obtain more time-relevant weather information, a second dataset was acquired from i-Weather, a global weather platform offering forecasts and historical data from certified

stations. This dataset allows for a closer examination of hourly weather conditions, particularly those experienced before and during the exam.

Hourly weather data was collected for the exam day (March 7, 2020) to capture conditions at key time intervals. Data from 8:30 a.m. to 9:00 a.m. reflects weather conditions students likely experienced while commuting to the exam venue, and data from 9:00 to 10:30 a.m. corresponds to the exam itself. To maintain the same approach and comparability of the results, the two cities - Michalovce and Poprad mentioned above were also deleted from this weather dataset. The final weather variables in this dataset include temperature (measured in degrees Celsius) and wind speed (in kilometres per hour), alongside categorical indicators for sky conditions, such as cloudy (reference category), partly cloudy, few clouds, rain, and snow.

For 8:30, time before the exam, the temperature ranged across the two countries from 1 to 7° Celsius and wind reached speed between 6 to 31 km per hour. When it comes to categorical indicators for sky conditions partly cloudy and few clouds were the most frequent sky conditions. For 9 to 10:30, the time of the exam the temperature ranged from 1 to 8.5° Celsius and wind speed from 4 to 38 kilometres per hour. The most frequent sky condition was partly cloudy. These summary statistics are showed in Table 3.2 for both hourly datasets.

Additionally, weather conditions from the night before the exam from i-Weather were collected to assess the potential indirect effects of weather on sleep quality and duration. This approach assumes that participants slept in close proximity to the exam location. Therefore, the weather data were obtained for the specific locations where participants took their exams. The weather conditions used to study the indirect effect of weather on sleep are temperature (measured in degrees Celsius) and wind speed (in kilometres per hour). The night weather statistics are summarised in Appendix in Table A.2.

This hourly dataset, however, also has its limitations. Unlike the daily dataset, the weather stations providing hourly data are not always located in immediate proximity to the exam venues. Nevertheless, the relatively uniform topography of the Czech Republic and Slovakia mitigates this limitation, and it is reasonable to assume that weather conditions recorded at nearby stations represent those experienced at the exam locations.

The combined use of these datasets provides a robust foundation for investigating the relationship between weather and exam performance. The daily dataset offers a macro-level perspective on general weather patterns, while the hourly dataset provides granular insights into short-term variations and their

potential effects during critical periods. Together, these datasets capture a broad range of weather variables, including temperature, wind speed, sunshine, precipitation and sky conditions. This dual approach strengthens the reliability of the findings and enables a comprehensive analysis of how weather may influence cognitive performance under high-stakes conditions.

Table 3.2: Weather conditions across Czech Republic and Slovakia

Weather data before exam

	Unit	**Min**	**Median**	**Mean**	**Max**
Temperature	°C	1	5	4.396	7
Wind	km/h	6	19	20.03	31
	Unit	**Frequency**	**Most Frequent**		
Cloudy	-	2	Partly cloudy and Few clouds		
Partly cloudy	-	10			
Few clouds	-	10			
Rain	-	7			
Snow	-	1			

Weather data during exam

	Unit	**Min**	**Median**	**Mean**	**Max**
Temperature	°C	1	6	5.636	8.5
Wind	km/h	4	25	24.32	38
	Unit	**Frequency**	**Most Frequent**		
Cloudy	-	3	Partly cloudy		
Partly cloudy	-	10			
Few clouds	-	8			
Rain	-	8			
Snow	-	1			

Weather data throughout the exam day

	Unit	**Min**	**Median**	**Mean**	**Max**
Temperature	°C	2.3	4.8	4.643	7.380
Precipitation	mm	0	0.2	0.485	2.8
Sunshine	hours	0	0.5	0.693	4.6
Wind	km/h	0.4	3	3.582	6.330

3.3 Other explanatory variables

Alongside exam performance, several individual characteristics about the students were collected from SCIO as well. This information was obtained either from data students filled in while signing up for the exam, such as gender age or preferred university, or through a supplementary questionnaire. The questionnaire included questions on sleep habits, and commuting time. It was neither part of the exam nor graded, and students completed it voluntarily. Together, these individual-level variables provide a rich foundation for exploring how personal characteristics intersect with weather to influence high-stakes exam performance. This section provides an overview of the additional explanatory variables, which add nuance to the dataset and enable a deeper understanding of the effects under investigation.

Gender information was recorded for all participants when they signed up for the exam. Out of the 2,864 participants, approximately 41% (1173) were female, while 59% (1691) were male. Even though, this proportion shows a larger representation of male participants, both groups are sufficiently represented to allow for meaningful analyses and comparisons. The age of participants ranged from fifty-three years old to eighteen years old, with the most frequent age being twenty. The summary statistics are showed in Table 3.3. Participants also filled out the most desired specialization of university which were chemical technology, economic sciences, humanities and social sciences, education, teaching and social welfare, legal sciences, natural sciences, technical sciences, cultural and artistic sciences, health, medicine and pharmaceutical sciences, agricultural, forestry and veterinary sciences. The most frequent specialization was humanities and social sciences. Even though information about age and the preferred field of study at university is not included in the regression models, as their effects on exam performance were consistently statistically insignificant, we still include them in this chapter as they provide relevant contextual information about the participating group.

The following information was obtained through self-reported questionnaire, which students filled out before the exam, offering additional insights into their pre-exam conditions. Out of the 2864 participants, 35% (994) answered all the questions, 52% (1491) answered at least some of the questions, and 13% (379) opted not to respond at all. Since the response rate varied across questions, the number of observations differs depending on the specific variable analysed.

The questionnaire included two questions about participants' sleep habits.

The first asked, "Approximately how many hours did you sleep last night?". Even though 85% (2442 out of 2864) of participants responded, not all answers were valid. Non-numerical responses or outliers deemed implausible, such as sleep durations of over 24 hours, were removed. After cleaning the data, 2434 valid responses remained, meaning eight answers were deleted. Reported sleep durations ranged from zero to twelve hours, with most participants sleeping between six and eight hours. The median sleep duration was 7 hours, and the average was 6.8 hours, as shown in Table 3.3 alongside with other statistics.

The second sleep-related question focused on the perceived quality of sleep: "Compared to a normal day, how would you rate the quality of your sleep today?" A total of 84% (2393 out of 2864) of participants responded, rating their sleep as significantly worse, slightly worse, very similar, slightly better, or significantly better. Responses were consolidated into three categories: worse than normal (consisting of participants answering significantly worse, slightly worse), same as normal (participants who answered very similar), and better than normal (participants who answered significantly better, slightly better). Out of the 2393 respondents, approximately 37% (893) reported worse sleep than usual, 46% (1,118) indicated their sleep quality was similar to a typical day, and 17% (413) reported better-than-normal sleep. These sleep-related responses provide valuable insights into the variability in students' rest before the exam, an important factor for understanding the indirect effects of weather.

Another important variable collected through the questionnaire was commuting time to the exam venue. Students reported the duration of their commute in minutes. Out of the 2864 participants, 84% (2,411) responded; however, 32 answers were removed due to implausible values such as excessively high commute times or non-numerical responses. After all, the final dataset includes 2379 valid responses. Commuting times ranged from 1 minute to over 11 hours with a median of 30 minutes and an average of 38.7 minutes. The most frequent commute time reported was 1 hour. This wide range reflects the geographic spread of students relative to the exam venues and the diversity of transportation methods used. The participants' commute time statistics are summarised in Table 3.3. Commuting data provides additional context to the dataset, offering an additional mediator to the effect of weather on exam performance through commuting.

Together, these information provide a comprehensive and multidimensional dataset that complements the exam performance and weather variables. The gender data ensures balanced representation, allowing for an exploration of

demographic patterns within the sample. Sleep-related data, including both duration and quality, offer detailed insights into participants' pre-exam rest, while the commuting data reflects the logistical aspects of the exam day. The sleep related and commuting variables also facilitate the analysis of indirect effects of weather on exam performance. By integrating these variables, the dataset provides a robust foundation for examining different set of factors related to the relationship between weather and exam performance in a highly standardized setting.

Table 3.3: Summary statistics for participants' personal information: Age, sleep duration and commuting time on the exam day

	Units	**Min**	**Median**	**Mean**	**Max**	**Most Frequent**
Age	Years	18	20	20.2	54	20
Sleep duration	Hours	0	7	6.78	12	7
Commuting time	Minutes	0	30	32.6	180	30

Chapter 4

Methodology

In this chapter, we describe the two methodological approaches used in the analysis. First, we employ multiple regression to examine the direct effect of weather on exam performance and the heterogeneous effect of weather across cognitive abilities and genders. Second, we apply mediation analysis to investigate if weather influences performance also indirectly through mediators.

4.1 Multiple regression

The goal of this thesis is to examine the effect of weather conditions on performance in high-stakes exams. To investigate such relationship, the following baseline specification was formulated:

$$Y_{ic} = \beta_0 + \boldsymbol{X}_c\Gamma + Z_{ic}\Lambda + \varepsilon_{ic}, \quad \text{where } i = 1, \ldots, n \quad \text{and} \quad c = 1, \ldots, m \quad (4.1)$$

Where:
Y_{ic}: Exam performance of student i in location c
$\boldsymbol{X}_c$: A vector of location-specific weather conditions
Z_{ic}: Gender of student i writing the exam in location c
ε_{ic}: Error term

The dependent variable Y_{ic} is a standardized exam score, meaning the exam performance is expressed as a percentile rank rather than raw points. SCIO sets up this approach for more straightforward interpretation and comparison across different test sessions while accounting for variations in test difficulty.

The studied independent variables are weather conditions $\boldsymbol{X}_c$, precisely

temperature, rain, wind speed, sunshine duration and sky conditions (categorized as cloudy, few clouds, partly cloudy, rain, and snow). Two weather datasets are utilized in the analysis: an hourly dataset, which records weather conditions thirty minutes before and throughout the exam period, and a daily dataset, which provides average weather conditions for the entire day. The hourly dataset offers more time-variant fluctuations in weather conditions that may directly influence cognitive performance. On the other hand, the daily dataset provides more precise spatial coverage, as the measurements are taken from meteorological stations that better represent the overall weather conditions in a given location. The robustness of the analysis is enhanced by incorporating both datasets and comparing the results. The primary parameter of interest is Γ, capturing the ceteris paribus effects of weather conditions on exam performance.

Estimation is conducted using Ordinary Least Squares (OLS) estimator. The validity of OLS estimates relies on several key assumptions, which are discussed in detail below. The first key assumption is linearity in parameters, which is satisfied by the way the model is specified, as all coefficients are simple scalars.

Secondly, the assumption of no perfect multicollinearity must be satisfied, meaning that no independent variable in the regression model is an exact linear combination of the others. This assumption was assessed using a correlation matrix, which did not indicate a perfect correlation among the independent variables. Furthermore, one category for categorical variables was always omitted to prevent perfect multicollinearity and ensure meaningful coefficient interpretation.

Another crucial assumption is the zero conditional mean of errors, which requires that the error term is uncorrelated with the independent variables. This implies that all explanatory variables must be exogenous, meaning they should not be systematically related to unobserved factors influencing student performance. In our case, this assumption may be violated if students in specific locations tend to perform better or worse due to unobserved local characteristics, such as school quality, socioeconomic background, or other region-specific factors that vary systematically with weather conditions. This might happen if, for example, more successful families tend to move away from regions with tough weather conditions, or if it is harder to find high-quality teachers to work in schools in such areas. By not accounting for these factors, the estimation could yield biased results by over- or underestimating the effect of weather on

exam performance. Ideally, this issue would be addressed using panel data, allowing for the inclusion of exam location fixed effects. However, only cross-sectional data was available. Therefore, we used the second-best approach - the inclusion of a control variable representing the average exam performance at each test location in March 2019, that is from the same term one year earlier. By doing so, the model accounts for persistent differences in students' ability and educational quality across locations, mitigating potential omitted variable bias and improving the validity of the estimated effect of weather conditions.

Homoscedasticity, which refers to the constant variance of the error term across observations, is another key assumption that must be met. To evaluate whether heterogeneity is present in our model, we performed the Breusch-Pagan. In all models, the test returned a p-value greater than 0.1, indicating that the null hypothesis of homoscedasticity cannot be rejected. Therefore, no evidence of heteroscedasticity was found in any of the models used in our study.

To account for potential violation of zero conditional mean a new baseline model incorporating the above mentioned control variable was used:

$$Y_{ic} = \beta_0 + \boldsymbol{X}_c\Gamma + Z_{ic}\Lambda + C_c\Omega + \varepsilon_{ic}, \quad \text{where } i = 1, \ldots, n \quad \text{and} \quad c = 1, \ldots, m \tag{4.2}$$

Where:
Y_{ic}: Exam performance of student i in location c
$\boldsymbol{X}_c$: A vector of location-specific weather conditions
Z_{ic}: Gender of student i writing the exam in location c
C_c: Average exam performance from previous year in location c
ε_{ic}: Error term

Given that the newly specified model most likely satisfies the key assumptions of linearity, no perfect multicollinearity, zero conditional mean of errors, and homoscedasticity, the OLS estimator in this analysis remains unbiased, consistent, and efficient.

Furthermore, the potential heterogeneity in the effect of weather on exam performance is examined. First, the heterogeneous effect of weather across cognitive skills is explored. Two additional models are estimated, each using either the verbal or analytical test scores as the dependent variable. Addi-

tionally, gender differences are analyzed by estimating two models - one with a subset of female students' test scores and another for male students' test scores.

The key OLS assumptions were also examined for each subgroup data to ensure the validity of the estimates. Linearity in parameters is satisfied as the models are correctly specified. No perfect multicollinearity was assessed through correlation matrix, which confirmed the satisfaction of this assumption. The zero conditional mean assumption is maintained by controlling for location-specific factors by including the previous year's average exam performance. Homoscedasticity is reassessed for each subsample's model and is confirmed by not rejecting the null hypothesis of the Breusch-Pagan test.

4.2 Mediation analysis

The above-described approach allows us to test whether there is a link between weather and exam performance. The natural follow-up question is through which channel weather affects exam performance. Thanks to the rich questionnaire, which included details of students' sleep and commuting times, we can investigate if weather affects exam performance through its effect on sleep quality, length, and commuting time. Previous studies have shown that longer commuting times can increase stress levels, which are negatively associated with exam performance, and that weather can result in sleep disruption, which is linked to worse cognitive outcomes. To investigate these relationships within a single context, we conduct a mediation analysis.

Mediation analysis allows us to test whether specific variables, referred to as predictors (weather conditions) affect the outcome (exam performance) through mediators (sleep quality, sleep duration, and commuting time), thereby creating a mediational effect. The methodological framework in this study is based on the principles outlined by Baron & Kenny (1986), supplemented by bootstrapping techniques.

Baron & Kenny (1986) introduced a four-step approach to mediation analysis, which assesses the relationships among the predictor, mediator, and outcome variable. The procedure begins by establishing the total effect of X (the predictor) on Y (the outcome). The next step involves estimating the effect of X on M (the mediator). Subsequently, a multiple regression analysis is conducted, incorporating both X and M to assess whether the mediator significantly predicts Y while controlling for X. Finally, the direct effect of X on Y in this final regression is compared to its total effect from the first step.

The econometric model for mediation analysis using Baron & Kenny (1986) method, is expressed as follows:

$$Y_{ic} = \alpha_0 + \boldsymbol{X}_c \Xi + \epsilon_{1ic}, \quad \text{where } i = 1, \ldots, n \quad \text{and} \quad c = 1, \ldots, m \tag{4.3}$$

$$M_{ic} = \beta_0 + \boldsymbol{X}_c \Upsilon + \epsilon_{2ic}, \quad \text{where } i = 1, \ldots, n \quad \text{and} \quad c = 1, \ldots, m \tag{4.4}$$

$$Y_{ic} = \delta_0 + \boldsymbol{X}_c \Theta + M_{ic} \Psi + \epsilon_{3ic}, \quad \text{where } i = 1 \ldots, n \quad \text{and} \quad c = 1, \ldots, m \tag{4.5}$$

Where:
Y_{ic}: Exam performance of student i in location c
M_{ic}: Mediator variable for student i in location c
X_{c}: A vector of location-specific weather conditions
Ξ: Total effect of X_{c} on Y_{ic} (step 1)
Υ: Effect of X_{c} on M_{ic} (step 2)
Θ: Direct effect of X_{c} on Y_{ic} controlling for M_{ic} (step 3)
Ψ: Effect of M_{ic} on Y_{ic} (step 3)
$\epsilon_{1ic}, \epsilon_{2ic}, \epsilon_{3ic}$: Independent error terms

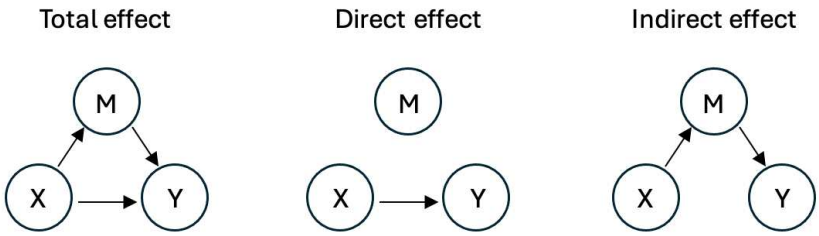
The indirect effect of X_{c} on Y_{ic} through the M_{ic} is calculated as the product of coefficients: Indirect effect = Υ*Ψ.

If the indirect effect is significant and the direct effect is reduced compared to the total effect, mediation is present. Complete mediation occurs if direct effect becomes insignificant, and partial mediation occurs if direct effect is reduced but remains significant. If there is no reduction in the direct effect and the indirect effect is not significant, it means no mediation is identified and the mediator does not explain the relationship between the predictor and the outcome.

The pathways described above are visually represented in the diagram Figure 4.1 where X is the independent variable, M is the mediator and Y is the outcome.

To obtain accurate results, the data used for mediation analysis must satisfy several assumptions. The causal order assumption requires that the independent variable influences the mediator before affecting the dependent variable. In our case, this requirement is satisfied, as all mediators in this study - night sleep and commuting occur before the exam. Weather conditions during the night may influence sleep quality and duration, which in turn can affect exam

Figure 4.1: Diagram for mediation analysis



performance. Similarly, weather before the exam may influence commuting time, subsequently impacting exam performance.

Several other assumptions come from the use of Ordinary Least Squares (OLS) estimator. Firstly, the linearity in parameters assumption is satisfied by the correct specification of the model. Multicollinearity was assessed using a correlation matrix to ensure that independent variables and mediators were not highly correlated. No evidence of perfect collinearity was found. The assumption of homoscedasticity was tested using the Breusch-Pagan test, which did not indicate the presence of heteroscedasticity.

To account for potential endogeneity, we include a control variable representing the average exam performance per location from 2019, thereby adjusting for historical differences in academic achievement across test centres. Additionally, we control for gender differences to account for individual-level variation. With these adjustments, Equation 4.3 becomes identical to the Equation 4.2 presented in Section 4.1, while Equation 4.5 corresponds to Equation 4.2 with one additional explanatory variable: the mediator.

The classical approach proposed by Baron & Kenny (1986) is to evaluate whether the indirect effect is significantly different from zero using the Sobel test (Sobel 1982). The indirect effect is given by the product of Υ and Ψ and the Sobel test assesses its significance based on the standard errors of Υ and Ψ.

The Sobel test, however, presents several limitations. It assumes that the indirect effect follows a normal distribution, which may not hold in small samples and is highly sensitive to measurement error. The normality assumption has been questioned by Kisbu-Sakarya *et al.* (2014), who suggested that the sampling distribution of the indirect effect is often skewed. We tested the normality of our residuals using the Shapiro-Wilk test, which showed a non-normal distribution. To address this issue, we followed the approach suggested

by Preacher & Hayes (2004) of using bootstrapping to retrieve the distribution of the indirect effect.

This technique resamples the dataset multiple times to approximate the sampling distribution of the estimator, providing a robust alternative when data distributions deviate from normality. It assumes that each iteration provides an unbiased estimate, allowing the retrieval of the theoretical estimator's distribution, including its mean and standard deviation. In our study, bootstrapping was performed with 5000 iterations, a commonly used threshold balancing computational efficiency and precision. The indirect effect from each iteration was stored, and confidence intervals were constructed to assess significance.

Judd (1981) introduced an alternative approach to mediation analysis, but it is primarily designed for experimental studies and was therefore not applied in this research.

Structural Equation Modeling (SEM) is another method used for mediation analysis, allowing for the simultaneous estimation of direct and indirect effects within a single framework. However, this model was not used in our study due to several considerations. SEM relies on strong theoretical foundations to specify relationships among latent constructs. However, in the case of our research, there is no established theoretical framework linking exam performance, weather, and the mediators under consideration, making this approach unsuitable for our study. Additionally, the method also requires specific data structure assumptions, including multivariate normality, which may not hold in the present dataset. Given these limitations, the chosen mediation analysis approach ensures compatibility with the studied data and aligns with the study's research objectives.

Chapter 5

Results

In this chapter, we discuss the results of the above-described methodological approaches. First, we present the findings from multiple regression analyses assessing the direct relationship between weather variables and exam performance, followed by examining the heterogeneous effects of weather on exam performance across cognitive skills and genders. Second, we describe the mediation analysis results, which investigates the indirect effects of weather through sleep and commuting time. Next, we assess the robustness of the estimates through various checks. Finally, we compare the results with the previously stated hypotheses.

5.1 Results multiple regression

In this section we comment on the results of the estimated direct relationship between weather variables and exam performance. Firstly, a model specified by Equation 4.1 using only weather variables and gender was estimated to assess the effect of weather conditions on exam performance. However, as discussed in Chapter 4, this model may face challenges in fulfilling the zero conditional mean assumption. To address this issue, a second estimation following the Equation 4.2 was conducted, incorporating a control variable representing the average exam scores from the previous year 2019 at each test location.

Table 5.1 presents the results for the models estimated with and without the additional control variable (average results from 2019 per location). Specifically, the results include specification (1) with hourly weather data before the exam, specification (2) with hourly weather data during the exam, and specification (3) with weather conditions throughout the exam day.

After the inclusion of the control variable, several notable changes in the results can be observed. The significance of some weather variables diminishes, indicating that part of their initial effect may have been confounded by differences in prior performance across locations. Such findings align with our assumption that people tend to perform better in particular places, for instance, due to the quality of education or higher competitiveness between students.

The effect of wind speed and sunshine loses significance after controlling for prior exam scores. These findings suggest that location-specific factors may account for some of the relationships previously attributed to direct meteorological conditions. The dummy variable for a few clouds retains its initial negative effect, although with weak significance when weather during the exam is taken into account and loses significance when weather conditions before the exam are taken into account. The effect of snow remains unchanged, continuing to be statistically insignificant across all specifications.

In contrast, the effect of temperature remains consistently positive and significant across specifications, reinforcing its role as an influential factor during exams. Furthermore, the effect of rain before and during the exam also remains positive and significant, while average rain during the exam day gains significance. The control variable itself is highly statistically significant and positive in all specifications, indicating that locations with historically higher scores tend to maintain this advantage, underscoring the importance of its inclusion in the model.

Hereafter, we focus on describing the results of the specifications, including the control variable (right panel in Table 5.1), as we believe they produce more reliable results. The positive effect of temperature is evident for all specifications. Temperature during the exam period suggests an increase in exam performance by 19.8 points if the temperature rises from 1°C (minimum in our data) to 8.5°C (maximum in our data), on average, c.p. Similarly, if daily average temperature rises from 2.3°C to 7.4°C, the exam performance increases by 13.94 points, on average, c.p. These findings confirm our assumption based on previous research of the importance of temperature in the context of exam performance. On the other hand, in specification (1), the effect of temperature before the exam is weaker and statistically insignificant, suggesting that immediate environmental conditions during the test may be more important for cognitive performance than the weather experienced before. Our findings are in line with previous literature, however, due to insufficient variation in our temperature data we were not able to confirm the potential bell-shaped effect

of temperature.

Wind speed is statistically insignificant across all specifications, meaning that we cannot conclude, at a conventional confidence level, that it has the proposed effect. The same result was found in the case of sunshine duration.

Cloud cover in the hourly data specifications presents notable variations in its impact on exam scores. Few clouds are associated with a weakly statistically significant negative effect on exam performance during the test period. This implies that partly obscured sky conditions can cause visual or psychological disruptions that negatively affect cognitive function. On the other hand, partly cloudy conditions do not exhibit a statistically significant impact.

Rain maintains a positive and statistically significant association with exam scores across multiple specifications. It is important to note that the rain variable in regressions taking into account rain before and during the exam is categorical - rain or other sky conditions, whereas it is measured as a continuous variable - millimetres dropped - for the rain throughout the exam day. The variable definition may contribute to some of the observed differences in the size of the effect. The coefficients for rain are significant both before and during the exam, with values of 5.852 and 5.354, respectively, meaning if it rains before the exam, the test score increases on average by 5.8 points, c.p. and on average by 5.4 points, c.p. if it rains during the exam. The effect is slightly weaker for the total amount of rain during the exam day (1.856). This means that if it starts raining, meaning the amount of rainfall during the day increases from zero (minimum in our data) to 2.8 millimetres (maximum in our data), the exam score increases by approximately 5.19 points, on average, c.p. The positive impact suggests that rain may improve focus by reducing external noise and distractions which is in line with previous findings of Proverbio *et al.* (2018); Lee *et al.* (2014).

Since snow was recorded in only one city, the results should be interpreted cautiously. Removing this variable might reduce the risk of unreliable estimates but also lower the number of observations, leading to potential biases in other coefficient estimates. Thus, we retain snow in the final model. To ensure valid results, we perform a robustness check by excluding observations with snow from our dataset and comment on the results in the Section 5.3 below.

Overall, the results for both hourly and daily weather conditions show similar patterns. Temperature and rain were found to be the key factors in exam performance. On the other hand, wind, sunshine hours, sky coverage (except for a few clouds), and snow do not seem to play a significant role. The con-

sistency of findings between the two datasets strengthens the reliability of the conclusions that some weather conditions on the exam day play a measurable role in exam performance.

It is important to note that the exam performance used in this study is measured using percentile scores, which reflect students' relative standing within the full distribution of test takers. This means that the outcomes do not capture absolute performance, such as the number of correct answers, but instead indicate how well student performed compared to others. As a result, any improvement in one student's percentile ranking necessarily corresponds to a shift in the rankings of others.

Table 5.1: Regression results for exam score with and without control variables and weather conditions

	Without control variable			**With control variable**		
Variable	(1) Before	(2) During	(3) Day	(1) Before	(2) During	(3) Day
Temp	$1.928^{\cdot}$ (1.151)	2.620^{**} (0.899)	3.359^{***} (0.823)	1.003 (1.147)	2.640^{**} (0.895)	2.738^{***} (0.812)
Wind	0.442^{***} (0.114)	$0.186^{\cdot}$ (0.108)	0.673^{***} (0.073)	0.098 (0.122)	0.112 (0.108)	−0.019 (0.102)
Sun	–	–	-2.560^{**} (0.812)	–	–	0.071 (0.845)
Few Clouds	-6.803^{*} (2.805)	-14.778^{***} (3.557)	–	2.056 (3.027)	$-7.306^{\cdot}$ (3.788)	–
Partly Cloudy	−0.714 (3.246)	−2.457 (2.860)	–	4.372 (3.289)	−2.028 (2.846)	–
Rain	7.878^{**} (2.517)	8.805^{***} (2.578)	1.156 (0.953)	5.852^{*} (2.509)	5.354^{*} (2.639)	1.856^{*} (0.940)
Snow	−0.434 (5.649)	2.808 (5.633)	–	−2.109 (5.601)	2.768 (5.603)	–
Results 2019	–	–	–	0.797^{***} (0.108)	0.745^{***} (0.135)	0.903^{***} (0.094)
Men	8.607^{***} (0.928)	8.641^{***} (0.922)	8.635^{***} (0.941)	8.499^{***} (0.919)	8.508^{***} (0.918)	8.356^{***} (0.926)
Constant	25.496^{***} (5.661)	26.081^{***} (4.883)	20.294^{***} (3.939)	−3.976 (6.887)	−8.116 (7.867)	-11.960^{*} (5.132)
Observations	2,818	2,818	2,775	2,818	2,818	2,775
R^2	0.082	0.090	0.069	0.099	0.100	0.099
Adj. R^2	0.079	0.088	0.067	0.096	0.098	0.097
Res. Std. Err	24.225	24.108	24.402	23.998	23.982	24.011
F Statistic	35.656^{***}	39.916^{***}	40.797^{***}	38.585^{***}	39.110^{***}	50.429^{***}

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Standard errors are shown in parentheses.
The results on the left panel are based on Equation (4.1) and on the right panel on Equation (4.2).
Columns represent separate specifications using (1) weather conditions before the exam, (2) during the exam, and (3) average weather conditions throughout the exam day.

Subsequently, the heterogeneous effects of weather conditions on exam performance were explored by analyzing their impact on different cognitive abilities to examine if they respond differently to environmental factors. More specifically, we estimated the effect of weather conditions on verbal and analytical test scores with the inclusion of control variable of average exam performance for each type of the test from 2019 in each test location. The estimation outcomes for verbal and analytical test scores, separately are displayed in Table 5.2. Specifications (1) use weather data before the exam, specifications (2) during and specifications (3) weather data observed throughout the exam day.

To be able to determine whether the effect of weather differs by cognitive load, we examine whether the differences in estimated coefficients are statistically significant. For this purpose, we compute 95% confidence intervals using the formula $\hat{\Gamma} \pm 1.96 \times \text{s.e.}(\hat{\Gamma})$ and compare them. If the cofidence intervals overlap, we cannot conclude that the estimated effects differ significantly between groups. The confidence intervals for each weather variable are presented in the appendix Table A.3.

Temperature during the exam and daily average temperature were found to be significant factors for both verbal and analytical scores. Temperature during the exam appears to have a slightly stronger estimated effect on analytical scores than verbal scores, whereas the daily average temperature shows a higher influence for verbal performance than for analytical performance. This suggests that verbal reasoning may be more influenced by consistent temperature conditions throughout the exam day rather than by short-term fluctuations. However, after comparing the confidence intervals of temperature across groups, they overlap, meaning these differences are not statistically significant, and we cannot conclude that the effect of temperature varies across the two cognitive domains.

Rain also emerges as a relevant weather factor, with statistically significant estimates in several specifications for both test types. The estimated effect of rain before and during the exam is somewhat stronger for verbal scores than for analytical scores. Yet again, the confidence intervals for both sets of coefficients overlap, and the differences cannot be interpreted as evidence of a heterogeneous effect.

The estimated influence of wind speed is positive and marginally significant for verbal scores before and during the exam (0.242 and 0.270), while remaining small and statistically insignificant for analytical reasoning. Despite this contrast, the difference in estimates for verbal and analytical parts were not

statistically different, suggesting no meaningful difference in how wind speed affects each test component.

For sky conditions, the presence of few clouds during the exam is significantly associated with lower scores for both verbal and analytical reasoning. The coefficients are remarkably similar in magnitude (-8.357 and -8.577), however, again the estimations were not statistically different and the heterogeneous effect cannot be confirmed. No significant effects were found for partly cloudy conditions, sunshine, or snow.

In summary, although estimates differ slightly across the test components and weather variables, our analysis provides no statistical evidence of a heterogeneous effect of weather on verbal and analytical reasoning. While previous studies have suggested that weather may affect distinct cognitive domains differently (Cho 2017; Zivin *et al.* 2015), we do not find support for such effects in the context of high-stakes exams, which might be caused by low variability in weather conditions in our data that causes relatively high standard errors of the estimates.

The effect of weather on exam performance was further analyzed concerning gender differences by estimating separate models for men and women. The results are presented in Table 5.3 for men and women, separately, where specifications (1) use weather data before the exam, specifications (2) during the exam, and specifications (3) weather data observed throughout the exam day.

Again, to be able to determine whether the effect of weather differs by gender, we calculate the confidence intervals for each weather variable, which are presented in the Appendix Table A.4 and compare them.

Temperature is positively associated with exam performance in both gender groups. During the exam period as well as when using daily average temperature, the estimates are statistically significant for women but not for men. However, a comparison of the confidence intervals indicates substantial overlap, and we therefore cannot conclude that the effect of temperature differs significantly between men and women.

Similarly, the estimated effects of rain before and during the exam are larger and statistically significant in the male subsample, while they are smaller and statistically insignificant in the female sample. Yet, the difference was not found to be statistically significant, meaning we cannot conclude the effect differs.

Other weather variables, including wind speed, sunshine duration, and snow, do not exhibit statistically significant effects in either group. The only exception is cloud cover few clouds during the exam is significantly associated

with lower performance for women, while the effect is smaller and not statistically significant for men. However, the estimates are not statistically different and the heterogeneous effect cannot be concluded.

Overall, while patterns in the estimated coefficients suggest potential differences in how weather conditions relate to exam performance across genders, these differences are not statistically significant. Previous research has suggested that women may be more sensitive to weather conditions (Karjalainen 2007), while others have shown that men tend to outperform women in high-stakes testing more generally. (Cai *et al.* 2019) However, these aspects have typically been examined in isolation. Although we found evidence that men perform better in our assessed high-stakes exams, we did not find evidence that gender moderates the effect of weather on exam performance. Our findings highlight the importance of studying the role of gender and weather sensitivity within a single context, as such effects may be dependent on the specific setting, test type, and population.

Table 5.2: Regression results for verbal and analytical exam scores and weather conditions

	Verbal Score			**Analytical Score**		
Variable	(1) Before	(2) During	(3) Day	(1) Before	(2) During	(3) Day
Temp	1.820 (1.378)	2.591* (1.081)	3.211** (0.978)	0.399 (1.241)	2.780** (0.965)	2.348** (0.874)
Wind	0.242· (0.144)	0.270* (0.130)	0.081 (0.128)	0.018 (0.134)	−0.018 (0.118)	−0.032 (0.102)
Sun	–	–	−0.229 (0.824)	–	–	0.039 (0.892)
Few Clouds	2.427 (3.649)	−8.357· (4.420)	–	0.315 (3.216)	−8.577* (4.227)	–
Partly Cloudy	4.762 (3.995)	−3.112 (3.437)	–	3.385 (3.497)	−1.469 (3.093)	–
Rain	8.175** (3.015)	5.737· (3.229)	2.030· (1.133)	3.952 (2.711)	5.482* (2.783)	1.590 (1.011)
Snow	0.907 (6.744)	2.125 (6.755)	–	−4.617 (6.039)	3.414 (6.027)	–
Results 2019	0.794*** (0.108)	0.735*** (0.133)	0.854*** (0.099)	0.648*** (0.135)	0.549*** (0.161)	0.832*** (0.112)
Men	6.145*** (1.108)	6.142*** (1.107)	5.984*** (1.116)	10.876*** (0.989)	10.913*** (0.988)	10.764*** (0.996)
Constant	−10.262 (7.905)	−9.671 (8.576)	−11.768* (5.675)	7.776 (7.744)	2.690 (8.924)	−7.809 (6.018)
Observations	2,818	2,818	2,775	2,818	2,818	2,775
R^2	0.083	0.083	0.082	0.078	0.080	0.078
Adj. R^2	0.081	0.081	0.080	0.076	0.078	0.076
Res. Std. Err	28.917	28.914	28.937	25.827	25.797	25.825
F Statistic	31.873***	31.954***	41.080***	29.808***	30.709***	38.988***

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '·' 0.1 ' ' 1
Standard errors are shown in parentheses.
The results are based on Equation (4.2).
Columns represent separate specifications using (1) weather conditions before the exam, (2) during the exam, and (3) average weather conditions throughout the exam day.

Table 5.3: Regression results for exam score of men and women and weather conditions

	Men			**Women**		
Variable	(1) Before	(2) During	(3) Day	(1) Before	(2) During	(3) Day
Temp	2.139 (1.841)	2.309 (1.456)	2.209$^{\cdot}$ (1.049)	0.312 (1.462)	2.926* (1.135)	3.048** (1.049)
Wind	−0.075 (0.196)	0.043 (0.179)	0.12 (0.132)	0.233 (0.156)	0.163 (0.136)	−0.108 (0.132)
Sun	–	–	0.378 (1.066)	–	–	−0.142 (1.066)
Few Clouds	2.810 (4.900)	−4.142 (6.140)	–	1.355 (3.839)	−10.199* (4.820)	–
Partly Cloudy	1.557 (5.296)	−1.750 (4.561)	–	6.420 (4.185)	−2.615 (3.662)	–
Rain	10.569** (4.073)	8.162* (4.089)	2.051 (1.244)	2.811 (3.177)	3.318 (3.477)	1.617 (1.244)
Snow	−0.128 (8.868)	−0.008 (8.741)	–	−3.108 (7.209)	4.649 (7.306)	–
Results 2019	0.585*** (0.176)	0.547* (0.225)	0.654*** (0.121)	0.918*** (0.137)	0.852*** (0.168)	1.065*** (0.121)
Constant	11.476 (11.066)	11.015 (12.889)	−19.532*** (6.554)	−8.170 (8.749)	−14.072 (9.942)	−19.532** (6.554)
Observations	1,170	1,170	1,617	1,648	1,648	1,617
R^2	0.053	0.053	0.088	0.091	0.090	0.088
Adj. R^2	0.048	0.048	0.085	0.087	0.086	0.085
Res. Std. Err	24.077	24.076	23.958	23.873	23.876	23.958
F Statistic	9.348***	9.354***	31.015***	23.327***	23.269***	31.015***

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Standard errors are shown in parentheses.
The results are based on Equation (4.2).
Columns represent separate specifications using (1) weather conditions before the exam, (2) during the exam, and (3) average weather conditions throughout the exam day.

5.2 Results mediation analysis

This section presents the mediation analysis results examining whether commuting time, sleep length, or sleep quality mediate the relationship between weather conditions and test performance. The mediation analysis focuses only on those weather variables that were found in Section 5.1 to have a direct effect on exam performance. Specifically, we analyse the effect of average temperature and the total amount of rainfall during the exam day, which are measured from midnight to midnight, making them well suited for capturing conditions that could influence both sleep and commuting on the day of the exam. In addition, we examine the effect of rain before the exam, which is particularly relevant for understanding the role of commuting. Lastly, we introduce an additional weather dataset that includes nighttime measurements of temperature and wind, allowing us to assess their potential impact through sleep.

The outputs of the mediation analysis show several different metrics. The *average causal mediated effect*, *ACME*, shows the indirect effect (Υ*Ψ) of the independent variable, weather, on test performance through a mediator. The *average direct effect* (Θ), *ADE*, shows the independent variable's unmediated effect (not passing through the mediator). Based on the significance of these results, one can conclude whether the effect of weather on test performance is direct or indirect. The *total effect* (Ξ) shows the sum of the average causal mediated and unmediated effects. Lastly, *proportion mediated* is the fraction of the total effect that is mediated. If this value is close to zero, it indicates that the mediator explains very little or none of the relationship between the independent variable and the outcome. On the other hand, if this value is closer to 1, it would indicate strong mediation.

Three weather variables are considered: rain before and during the exam, total rainfall during the exam day, and average temperature during the exam day. The results presented in Table 5.4 indicate no evidence of mediation. The *ACME* estimates for rain one hour prior to the exam ($p = 0.53$), average temperature during the exam day ($p = 0.15$), and total rainfall during the exam day ($p = 0.22$) are all statistically insignificant, suggesting that commuting time does not mediate the relationship between weather conditions and exam performance. It should be emphasized that the *ACME* for average temperature during the day is close to being significant. Perhaps an indirect effect of temperature through commuting would be present if more variation in the weather conditions would be observed.

Table 5.4: Mediation analysis results: Commuting time and weather conditions

Morning weather conditions before the exam

Weather Condition	Metric	Estimate	95% CI Lower	95% CI Upper	p-value
Rain	ACME	-0.0466	-0.2677	0.080	0.530
	ADE	3.1305	-2.4359	8.400	0.270
	Total Effect	3.0839	-2.4750	8.370	0.280
	Prop. Mediated	-0.0151	-0.2307	0.180	0.640

Sample size: 2375

Average weather conditions throughout the day

Weather Condition	Metric	Estimate	95% CI Lower	95% CI Upper	p-value
Temperature	ACME	-0.0483	-0.1566	0.010	0.1500
	ADE	2.7749	1.0728	4.560	0.0012 **
	Total Effect	2.7266	1.0204	4.510	0.0016 **
	Prop. Mediated	-0.0177	-0.0775	0.010	0.1516
Rain	ACME	-0.0354	-0.1343	0.020	0.220
	ADE	1.6239	-0.3205	3.580	0.110
	Total Effect	1.5885	-0.3671	3.540	0.120
	Prop. Mediated	-0.0223	-0.2036	0.140	0.310

Sample size: 2338

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
The results are based on Equations (4.3), (4.4) and (4.5)

Secondly, the potential mediation effect of sleep length on the relationship between weather variables and exam performance was analysed. Sleep duration on the night before the exam may not be directly influenced by weather conditions on the day of the exam. However, since weather on the exam day could be correlated with weather during the preceding night, we account for the mediating role of sleep by considering weather conditions during the night before the exam. Therefore, a reduced set of weather variables compared to the ones used in Section 5.1 for night-time weather conditions was used. Certain categorical variables, such as cloud cover, were excluded from the estimation, as individuals do not directly experience differences in sky conditions while sleeping indoors. Additionally, since night-time weather data represent averages over the entire night, calculating an average from categorical variables would not be meaningful. Therefore, temperature and wind were assessed for night-time weather data, and temperature and rain were examined for daily average data concerning exam performance through sleep duration. It is important to note that this analysis assumes that individuals slept near the exam location from which the weather data were collected as we do not have information on participant's place of living.

Table 5.5 shows the outputs of mediation through sleep length. *ACME* estimates for temperature ($p = 0.87$) and wind ($p = 0.52$) for nigh-time data and temperature ($p = 0.86$) and rain ($p = 0.57$) for daily average data are statistically insignificant. Also, the proportion of the effect mediated remains close to zero.

Table 5.5: Mediation analysis results: Sleep length and weather conditions

Night weather conditions before the exam day

Weather Condition	Metric	Estimate	95% CI Lower	95% CI Upper	p-value
Temperature	ACME	0.00413	-0.03331	0.050	0.867
	ADE	1.40393	-0.04183	2.820	0.054 .
	Total Effect	1.40806	-0.05015	2.810	0.054 .
	Prop. Mediated	0.00293	-0.04246	0.060	0.872
Wind	ACME	-0.00322	-0.01523	0.010	0.520
	ADE	-0.02849	-0.22642	0.170	0.780
	Total Effect	-0.03172	-0.23008	0.160	0.750
	Prop. Mediated	0.10167	-0.73392	0.800	0.880

Sample size: 2397

Average weather conditions throughout the day

Weather Condition	Metric	Estimate	95% CI Lower	95% CI Upper	p-value
Temperature	ACME	0.00329	-0.04185	0.060	0.8596
	ADE	2.34704	0.60867	4.070	0.0092 **
	Total Effect	2.35034	0.60246	4.080	0.0092 **
	Prop. Mediated	0.00140	-0.02369	0.030	0.8584
Rain	ACME	-0.0207	-0.1193	0.050	0.570
	ADE	1.3474	-0.5681	3.320	0.180
	Total Effect	1.3268	-0.5831	3.310	0.180
	Prop. Mediated	-0.0156	-0.2336	0.180	0.650

Sample size: 2360

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
The results are based on Equations (4.3), (4.4) and (4.5)

The mediation analysis was also conducted with sleep quality as a categorical mediator, distinguishing between better, worse, and normal sleep (reference category). Similarly, categorical weather variables for sky conditions were excluded, as the effect indoors is not directly noticeable and averaging categorical variables would not provide a meaningful representation of exposure. Therefore, temperature and wind were examined from the night-time dataset, and temperature and rain were assessed from the daily average dataset. The analysis was conducted under the same assumption as with sleep length - individuals slept near the exam location.

The results for sleep quality, presented in Table 5.6, further confirm the

absence of significant indirect effects. The *ACME* estimates remain small and statistically insignificant for temperature and wind in both cases for nighttime data - better sleep ($p = 0.31$ and $p = 0.21$, respectively) and worse sleep ($p = 0.95$ and $p = 0.37$, respectively). The same results were found for temperature and rain using daily average data - better sleep ($p = 0.38$ and $p = 0.96$, respectively) and worse sleep ($p = 0.78$ and $p = 0.62$, respectively). These findings suggest that self-reported sleep disturbances before the exam do not mediate the relationship between weather and test performance.

We can observe that the direct effect of rain in Table 5.4, Table 5.5 and in Table 5.6 is not statistically significant in the mediation analysis, what is in contrast to the multiple regression model in Section 5.1. This can result from including mediator in the model, which may alter the estimation of coefficients and increase the standard errors. However, the direct effect of rain is on the edge of significance and its coefficient remains close to the one in the multiple regression estimation. Therefore, we do not conclude any change in relevance of rain in our analysis.

Overall, the mediation analysis provides no evidence that commuting time, sleep length, or sleep quality explains the observed relationship between weather and exam performance. Across all specifications, the *average causal mediation* effects remain small and statistically insignificant, indicating that these potential mediators do not explain the observed relationship between weather conditions and exam performance.

While previous research has suggested that weather may disrupt sleep patterns (Zhu *et al.* 2016; Lan *et al.* 2017) and that sleep disturbances can impair cognitive performance (Fakhari *et al.* 2016), or that longer commuting has been linked to increased stress levels, a factor negatively associated with exam performance in high-stakes settings (Heissel *et al.* 2021), our results of examining these in a single context do not provide supporting evidence of such findings. A likely explanation is that the weather conditions during the study period were not sufficiently adverse to meaningfully disrupt participants' sleep or commuting. This might be particularly true for the effect of average daily temperature through commuting, where the indirect effect is on the verge of being significant and more adverse conditions could yield supporting evidence of mediation.

Secondly, all mediators were self-reported, subjecting them to recall bias and individual perception. Since students filled out the questionnaire after the exam, they may not have recalled details accurately, and factors such as exam stress could have influenced their responses. Finally, the findings suggest that

the effect of weather on cognitive performance is direct rather than through changes in sleep patterns or commuting duration.

Table 5.6: Mediation analysis results: Sleep quality and weather conditions

Night weather conditions before the exam day

Better Sleep

Weather Condition	Metric	Estimate	95% CI Lower	95% CI Upper	p-value
Temperature	ACME	-0.03190	-0.12120	0.030	0.309
	ADE	1.35710	-0.05190	2.710	0.058 .
	Total Effect	1.32520	-0.07930	2.680	0.064 .
	Prop. Mediated	-0.02410	-0.20050	0.080	0.359
Wind	ACME	0.00586	-0.00237	0.020	0.210
	ADE	-0.04718	-0.22785	0.140	0.630
	Total Effect	-0.04133	-0.22216	0.140	0.680
	Prop. Mediated	-0.14169	-0.89304	0.950	0.760

Worse Sleep

Weather Condition	Metric	Estimate	95% CI Lower	95% CI Upper	p-value
Temperature	ACME	0.00185	-0.04751	0.050	0.947
	ADE	1.35714	-0.10261	2.750	0.066 .
	Total Effect	1.35900	-0.09640	2.750	0.066 .
	Prop. Mediated	0.00136	-0.07184	0.070	0.951
Wind	ACME	-0.00364	-0.01459	0.000	0.370
	ADE	-0.04718	-0.23251	0.140	0.640
	Total Effect	-0.05082	-0.23575	0.140	0.610
	Prop. Mediated	0.07154	-0.56756	0.620	0.760

Sample size: 2392

Average weather conditions throughout the day

Better Sleep

Weather Condition	Metric	Estimate	95% CI Lower	95% CI Upper	p-value
Temperature	ACME	-0.03530	-0.15040	0.040	0.3768
	ADE	2.48690	0.78810	4.230	0.0044 **
	Total Effect	2.45160	0.76420	4.190	0.0052 **
	Prop. Mediated	-0.01440	-0.08720	0.020	0.3804
Rain	ACME	0.00182	-0.10551	0.110	0.960
	ADE	1.39861	-0.53827	3.320	0.170
	Total Effect	1.40043	-0.54013	3.310	0.170
	Prop. Mediated	0.00130	-0.23298	0.190	0.960

Worse Sleep

Weather Condition	Metric	Estimate	95% CI Lower	95% CI Upper	p-value
Temperature	ACME	0.00906	-0.04967	0.080	0.7780
	ADE	2.48690	0.78021	4.210	0.0060 **
	Total Effect	2.49596	0.78015	4.240	0.0056 **
	Prop. Mediated	0.00363	-0.02623	0.040	0.7764
Rain	ACME	-0.01660	-0.10270	0.050	0.620
	ADE	1.39860	-0.50040	3.440	0.150
	Total Effect	1.38200	-0.51240	3.410	0.160
	Prop. Mediated	-0.01200	-0.20440	0.160	0.670

Sample size: 2354

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
The results are based on Equations (4.3), (4.4) and (4.5)

5.3 Robustness check

Robustness checks were conducted to ensure the reliability of our results. Firstly, outliers were detected using the interquartile range (IQR) method, where observations falling below the lower bound (Q1 - 1.5 x IQR) or above the upper bound (Q3 + 1.5 x IQR) were considered extreme values. Q1 signifies the first quintile, Q3 signifies the third, and IQR represents the difference between Q3 and Q1. This approach was applied separately to each weather variable, which showed up to be statistically significant in our main regressions presented in Table 5.1 as well as to the average test score of each exam location. The goal was to assess whether the estimated effects were driven by a few extreme observations. By focusing only on significant variables, we ensured that the robustness checks targeted the core findings of the analysis, avoiding additional variation from variables that did not contribute meaningfully to the main results. Moreover, if we would consider outliers for all weather variables at once, the number of observations would significantly decrease, which could yield inaccurate results.

Since some variables in the dataset are dummy variables, traditional outlier detection methods do not apply. Instead, their distribution was examined to ensure that extreme imbalances or inconsistencies did not influence the results. Both significant dummy variables - rain and few clouds - were recorded in multiple exam locations, suggesting that their representation in the dataset was sufficient for reliable analysis.

No outliers were detected for the average test scores across exam locations. Only temperature during the exam (along with the two dummy variables - few clouds and rain) was statistically significant. Temperature values did not exhibit any outliers. However, the presence of outliers was identified in Košice and Svitavy for average temperature during the day and average amount of rainfall during the day in Liberec.

To assess if the discovered outliers alter the main findings, we re-estimated the primary models with their exclusion. The results are presented in the appendix in Table A.5. The model was first estimated separately without outliers for each weather variable: specification (1) excludes temperature outliers, specification (2) excludes rain outliers, while specification (3) excludes all identified outliers simultaneously.

Focusing on specification (3), where all extreme values are removed simultaneously, we observe that the effect of temperature remains positive and sig-

nificant, with an increase in magnitude compared to the original specification. Similarly, rain remains positive and significant, with a slightly stronger effect, reinforcing its role in exam performance. Wind and sunshine remain their signs and insignificance, which is consistent with the main findings. Overall, excluding extreme values does not change the core findings, supporting the robustness of the identified relationships.

To further test the robustness of our findings, the main specifications presented in Table 5.1 with control variable were re-estimated, excluding Slovakia as a country with fewer observations, considering only Czech exam locations. The results, presented in the appendix in Table A.6, are generally consistent with the full-sample estimates, with slight differences in the significance of specific weather effects.

For specification (1) across both tables, which examines weather conditions before the exam, rain remains positive and significant with a slightly stronger effect. The main difference in significance is observed for the cloud cover dummy variable partly cloudy, which becomes significant at the 10% level and shows a more substantial effect on exam performance. Temperature, wind and snow remain insignificant.

For specification (2), which reflects weather during the exam, the temperature remains positive and significant with a slightly stronger effect. The effect of rain weakens and loses significance, implying that rain's influence may have been more pronounced in the Slovak dataset. The dummy variable few clouds also exhibits minor differences, as it loses its statistical significance (previously only at the 10% level) but retains the same negative estimated effect on exam performance.

Lastly, for specification (3), which use the daily average weather data, the temperature remains positive and significant, with little change in magnitude. Rain keeps its positive sign and magnitude while its significance slightly weakens. The effect of wind and sunshine remains insignificant.

Overall, Slovakia's exclusion does not substantially alter the main conclusions. The robust findings for temperature suggest that it consistently influences exam performance, regardless of regional differences. The impact of rain needs to be interpreted with caution as it seems to be more sensitive to regional differences. Cloud cover effects exhibit minor variations in significance but do not substantially alter the interpretation of results. These findings confirm that regional differences do not strongly drive the estimated weather effects.

The last robustness check performed was re-estimating the main regressions

from Table 5.1 without snow, which was measured in only one exam location (Sokolov) across hourly weather datasets. The results of the sub-sample estimations are displayed in appendix in Table A.7. Comparing the impact of weather before the exam, specification (1), rain, remains positive and significant, with a minor increase in the effect, confirming its robustness. Wind speed, cloud cover, and temperature remain insignificant, with the same sign.

For weather during the exam, specification (2), the effect of temperature remains positive and significant, though slightly lower, confirming its stability. Rain remains consistently positive and significant, with a slightly higher magnitude. Wind remains positive and insignificant. The main change is in a few clouds, which shifts from a weakly significant negative effect to a more robust negative effect, suggesting its previous estimate may have been influenced by snow.

Overall, excluding snow does not alter the main findings. Temperature and rain remain positively associated with performance, while few clouds maintain their negative effect.

The robustness checks confirm the stability and reliability of the estimated weather effects on exam performance. The exclusion of outliers strengthens the significance of temperature and rain, suggesting that extreme values may have slightly dampened their estimated effects but did not alter their direction. The exclusion of Slovakia yields results that are broadly consistent with the entire sample, with minor regional differences, particularly in the significance of rain. Similarly, the exclusion of snow, which was measured in only one exam location, does not meaningfully change the results, apart from a slightly more substantial negative effect of cloud cover.

5.4 Hypothesis discussion

The results provide partial support for the proposed hypotheses. While some weather variables show a significant relationship with exam performance, others do not. Hypothesis one is partially supported, as certain weather conditions influence test scores while others remain insignificant. Hypothesis two is not supported, as no evidence was found for the indirect effect of weather through sleep or commuting time. Hypothesis three and four were not confirmed, as the estimates were not found to be statistically different from each other.

These findings should be considered within the context of the Czech Republic and Slovakia and the specific nature of the SCIO exam. The standardized

format and high-stakes nature may limit the impact of external environmental factors compared to other educational settings used in the previous studies. Additionally, weather conditions on the exam day were not extreme and did not vary enough, which may explain the lack of significant effects for some variables. Finally, individual adaptation to local climate patterns and the consistency of exam preparation across locations could also contribute to the observed results. More detailed and hypothesis-specific reasoning is discussed in the respective sections.

Hypothesis 1: *Different weather variables (temperature, rain, sunshine, wind and sky conditions) have a direct significant effect on exam performance.*

The results partially support the hypothesis, showing that some weather conditions influence exam performance while others do not. Temperature and rain were positively associated with test scores, suggesting that moderate warmth may enhance cognitive functioning, while rain might reduce distractions and improve focus. Additionally, few clouds were linked to lower scores, which may be related to fluctuations in natural light exposure.

In contrast, wind, sunshine hours, sky conditions - partly cloudy, and snow did not show significant effects, indicating that these conditions did not substantially alter cognitive performance in this setting.

Hypothesis 2: *Different weather variables (temperature, rain, sunshine, wind and sky conditions) have an indirect effect on exam performance through sleep length, sleep quality and commuting time.*

The findings of our study do not support the hypothesis that different weather variables indirectly affect exam performance, as no significant mediation effects were discovered for sleep length, sleep quality, and commuting time.

Several factors may explain why no significant mediation effects were found. One possible reason is that the weather impacts exam performance through direct psychological pathways rather than through sleep or commuting time. Moreover, weather conditions on the exam day were not extreme enough to substantially affect sleep or commuting time. This point is supported mainly by the average temperature during the day of the exam for which the indirect effect through commuting is close to being significant and more adverse weather

could yield different results. Lastly, the mediators analyzed were based on self-reported data, which introduces potential recall bias and subjective variation.

Hypothesis 3: *Weather has a stronger effect on activities with higher cognitive load, meaning the effects proposed in H1 are stronger for the analytical part of the test than the verbal one.*

The results do not confirm the hypothesis that weather has a stronger effect on activities with higher cognitive load. While the estimates suggested some differences, these findings cannot be confirmed as the estimates were not found to be statistically different from each other. While previous studies suggested the heterogeneous effect on normal exams, we were not able to confirm it in high-stakes settings.

Hypothesis 4: *Women are more sensitive to weather than men, meaning the effects proposed in H1 are stronger for women.*

The results do not support the hypothesis that women are more sensitive to weather than men. The estimates were not found to be statistically different from each other, even though they suggested some varying effects across weather variables. Even though previous studies indicated that men tend to perform better in high-stakes exams and that women are more sensitive to weather, we were not able to confirm these findings within a single context.

Chapter 6

Conclusion

High-stakes exams, such as school admission tests or final school leaving exams, represent critical milestones in students' lives, shaping their future educational and career opportunities. Despite their significance, research addressing the impact of external factors on exam performance remains limited, particularly regarding the influence of weather. This gap in the literature should be addressed especially given the increasing unpredictability and extremity of weather patterns in recent years.

Our study makes a valuable contribution to this area of research by being, to the best of our knowledge, the first to examine the effect of weather on high-stakes exam performance in the Czech Republic and Slovakia. The analysis is conducted using a unique and rich dataset of exam performance in university admission testing that is not publicly available. We employ a wide range of weather variables and investigate also the heterogeneous effects across cognitive abilities and genders. Furthermore, we explore whether the effect of weather is direct or mediated through factors such as sleep length and quality or commuting time. Our results were also subjected to robustness checks to ensure reliability.

The key findings of our study indicate that temperature and rain are the most influential weather factors affecting positively exam performance. Few clouds, on the other hand, showed the opposite impact. Even though the estimates suggested contrasting results across cognitive demands and gender, they did not show up to be statistically different from each other to confirm the heterogeneous effects of weather. Mediation analysis suggested that the impact of weather on exam performance may be rather direct or that the weather conditions observed on the exam day were not extreme enough to activate the

indirect mechanisms as our results did not identify significant mediating effects of sleep or commuting.

Our study emphasizes the importance of external factors when evaluating high-stakes exams, particularly as the weather conditions observed on the exam day were relatively mild and in line for the time of the year, yet still demonstrated significant impact on performance. In the context of increasingly volatile weather patterns, understanding and addressing these influences becomes more important.

While our findings do not lead to straightforward policy recommendations, such as shifting exam periods to seasons with more stable weather, they do suggest several practical considerations. For instance, maintaining optimal indoor temperature conditions is crucial, as reports indicate that schools fail to consistently regulate classroom environments. Additionally, increasing the frequency of high-stakes exams could offer students more flexibility and reduce the impact of adverse external conditions. Policymakers might also consider incorporating adaptive measures, such as allowing for more flexible exam scheduling during periods of extreme weather, to mitigate potential disadvantages faced by affected students.

Our study also presents several limitations. The weather conditions observed were relatively mild, leaving the possibility that more extreme weather could yield different results, especially in the case of mediation analysis which showed some results on the edge of becoming significant. We also suspect that some weather variables, particularly temperature, may exhibit a non-linear bell-shaped relationship with performance, but our dataset lacked the necessary variability to capture this pattern. Including additional variables, such as classroom temperature, lighting conditions, or students' stress levels, would likely provide a more nuanced understanding of how weather affects cognitive performance.

Despite these limitations, we believe that our study offers valuable insights into the relationship between weather and exam performance and highlights the importance of further research on this topic.

Bibliography

Ahrberg, K., M. Dresler, S. Niedermaier, A. Steiger, & L. Genzel (2012): "The interaction between sleep quality and academic performance." *Journal of Psychiatric Research* **46(12)**: pp. 1618–1622.

Alberto, I. C., Y. Jiao, & X. Zhang (2021): "Too hot or too cold to study? the effect of temperature on student time allocation." *Economics of Education Review* **84**: p. 102152.

Arthur, N. & P. Everaert (2012): "Gender and performance in accounting examinations: Exploring the impact of examination format." *Accounting Education* **21(5)**: pp. 471–487.

Baron, R. & D. Kenny (1986): "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations." *Journal of Personality and Social Psychology* **51**: pp. 1173–1182.

Beebe, D. W., D. Rose, & R. Amin (2010): "Attention, learning, and arousal of experimentally sleep-restricted adolescents in a simulated classroom." *Journal of Adolescent Health* **47(5)**: pp. 523–525.

Bélanger, M., K. Gray-Donald, J. O'loughlin, G. Paradis, & J. Hanley (2009): "Influence of weather conditions and season on physical activity in adolescents." *Annals of Epidemiology* **19**: pp. 180–186.

Bundo, M., M. Preisig, K. Merikangas, J. Glaus, J. Vaucher, G. Waeber, P. Marques-Vidal, M.-P. F. Strippoli, T. MĂĽller, O. Franco, & A. M. Vicedo-Cabrera (2023): "How ambient temperature affects mood: an ecological momentary assessment study in switzerland." *Environmental Health* **22**: pp. 1–11.

Cai, X., Y. Lu, J. Pan, & S. Zhong (2019): "Gender Gap under Pressure: Evidence from China's National College Entrance Examination." *The Review of Economics and Statistics* **101(2)**: pp. 249–263.

Cho, H. (2017): "The effects of summer heat on academic achievement: A cohort analysis." *Journal of Environmental Economics and Management* **83**: pp. 185–196.

Cook, N. & A. Heyes (2020): "Brain freeze: outdoor cold and indoor cognitive performance." *Journal of Environmental Economics and Management* **101**: p. 102318.

Copernicus (2024): "Copernicus: Summer 2024 - hottest on record globally and for europe." https://climate.copernicus.eu/copernicus-summer-2024-hottest-record-globally-and-europe, accessed 2024-12-17.

Denissen, J., L. Butalid, L. Penke, & M. Aken (2008): "The effects of weather on daily mood: A multilevel approach." *Emotion (Washington, D.C.)* **8**: pp. 662–7.

Engle-Friedman, M., S. Riela, R. Golan, A. Ventuneac, C. Davis, A. Jefferson, & D. Major (2003): "The effect of sleep loss on next day effort." *Journal of sleep research* **12**: pp. 113–24.

Fakhari, A., N. Kheradmand, & N. Dolatkhah (2016): "Sleep duration the night before an exam and its relationship to students' exam scores." *British Journal of Medicine and Medical Research* **15**: pp. 1–6.

Falla, M., A. Micarelli, K. Hüfner, & G. Strapazzon (2021): "The effect of cold exposure on cognitive performance in healthy adults: A systematic review." *International Journal of Environmental Research and Public Health* **18**: p. 9725.

Finlay, J., A. Khan, C. Gronlund, K. Sol, J. Jang, R. Melendez, S. Judd, & P. Clarke (2020): "Weather woes? exploring potential links between precipitation and age-related cognitive decline." *International Journal of Environmental Research and Public Health* **17**: p. 9011.

Funk, P. & H. Perrone (2023): "Gender differences in academic performance: The role of exam design in multiple-choice tests." `https://ssrn.com/abstract=4401568`. Preprint, not peer-reviewed.

Gao, Q., D. chun Luan, X. Wang, S. Xin, Y. Liu, & J. Li (2018): "Effect of sun exposure on cognitive function among elderly individuals in northeast china." *Clinical Interventions in Aging* **13**: pp. 2075–2082.

Graff Zivin, J., Y. Song, Q. Tang, & P. Zhang (2020): "Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in china." *Journal of Environmental Economics and Management* **104**: p. 102365.

Haverinen-Shaughnessy, U. & R. Shaughnessy (2015): "Effects of classroom ventilation rate and temperature on studentsâ€™ test scores." *PLoS ONE* **10**: pp. 1–14.

Heissel, J. A., E. K. Adam, J. L. Doleac, D. N. Figlio, & J. Meer (2021): "Testing, stress, and performance: How students respond physiologically to high-stakes testing." *Education Finance and Policy* **16(2)**: pp. 183–208.

Herawati, K. & D. Gayatri (2019): "The correlation between sleep quality and levels of stress among students in universitas indonesia." *Enfermería Clínica* **29**: pp. 357–361.

Heshmati, K., E. Shahabpoor, A. Darby, I. Walker, & S. Ghanbari (2024): "Effects of wind-induced vibrations in tall buildings on cognitive work performance, comfort, and wellbeing of the occupants." *Journal of Physics: Conference Series* **2647**: p. 252032.

Howarth E, H. M. (1984): "A multidimensional approach to the relationship between mood and weather." *Br J Psychol* **75**: pp. 15–23.

Judd, C. M., . K. D. A. (1981): "Process analysis: Estimating mediation in treatment evaluations." *Evaluation Review* **5**: pp. 602–619.

Karjalainen, S. (2007): "Gender differences in thermal comfort and use of thermostats in everyday thermal environments." *Building and Environment* **42(4)**: pp. 1594–1603.

Keller, M., B. Fredrickson, O. Ybarra, S. Cote, K. Johnson, J. Mikels, A. Conway, & T. Wager (2005): "A warm heart and a clear head the contingent effects of weather on mood and cognition." *Psychological science* **16**: pp. 724–31.

Kent, S., L. McClure, W. Crosson, D. Arnett, V. Wadley, & N. Sathiakumar (2009): "Effect of sunlight exposure on cognitive function among depressed and non-depressed participants: A regards cross-sectional study." *Environmental health : a global access science source* **8**: p. 34.

Khan, A. M., J. M. Finlay, P. Clarke, K. Sol, R. Melendez, S. Judd, & C. J. Gronlund (2021): "Association between temperature exposure and cognition: a cross-sectional analysis of 20,687 aging adults in the united states." *BMC Public Health volume* **21**: pp. 1471–2458.

Kisbu-Sakarya, Y., D. MacKinnon, & M. Miočevič (2014): "The distribution of the product explains normal theory mediation confidence interval estimation." *Multivariate behavioral research* **49**: pp. 261–268.

Lan, L., K. Tsuzuki, Y. Liu, & Z. Lian (2017): "Thermal environment and sleep quality: A review." *Energy and Buildings* **149**: pp. 101–113.

Lee, D., A. Tajar, A. Ulubaev, N. Pendleton, T. O'Neill, D. O'Connor, G. Bártfai, S. Boonen, R. Bouillon, F. Casanueva, J. Finn, G. Forti, A. Giwercman, T. Han, I. Huhtaniemi, K. Kula, M. Lean, M. Punab, A. Silman, & F. Wu (2009): "Association between 25-hydroxyvitamin d levels and cognitive performance in middle-aged and older european men." *Journal of neurology, neurosurgery, and psychiatry* **80**: pp. 722–9.

Lee, J., F. Gino, & B. Staats (2014): "Rainmakers: Why bad weather means good productivity." *Journal of Applied Psychology* **99**: pp. 504–513.

Leiner, J. E. M., T. Scherndl, & T. M. Ortner (2018): "How do men and women perceive a high-stakes test situation?" *Frontiers in Psychology* **9**: pp. 1664–1078.

Li, X. & P. C. Patel (2021): "Weather and high-stakes exam performance: Evidence from student-level administrative data in brazil." *Economics Letters* **199**: p. 109698.

Llewellyn, D., I. Lang, K. Langa, & D. Melzer (2010): "Vitamin d and cognitive impairment in the elderly u.s. population." *The journals of gerontology. Series A, Biological sciences and medical sciences* **66**: pp. 59–65.

Okano, K., J. R. Kaczmarzyk, N. Dave, J. D. E. Gabrieli, & J. C. Grossman (2019): "Sleep quality, duration, and consistency are associated with better academic performance in college students." *NPJ Science of Learning* **4**: p. 16.

Park, R. J. (2022): "Hot Temperature and High-Stakes Performance." *Journal of Human Resources* **57(2)**: pp. 400–434.

Pazos, N., M. Favara, A. Sánchez, D. Scott, & J. Behrman (2024): "Long-term effects of early life rainfall shocks on foundational cognitive skills: Evidence from peru." *Economics Human Biology* **54**: p. 101407.

Preacher, K. & A. Hayes (2004): "Spss and sas procedures for estimating indirect effects in simple mediation models." *Behavior research methods, instruments, computers : a journal of the Psychonomic Society, Inc* **36**: pp. 717–31.

Prenhaca, I. G. (2023): "Here comes the rain: Impacts of weather shocks on higher education decisions." Working paper, Insper Institute of Education and Research. Unpublished Manuscript.

Proverbio, A. M., F. D. Benedetto, M. V. Ferrari, & G. Ferrarini (2018): "When listening to rain sounds boosts arithmetic ability." *PLoS One* **13**: pp. 1–19.

Pun, M., V. Guadagni, K. M. Bettauer, L. L. Drogos, J. Aitken, S. E. Hartmann, L. M. Michael Furian, M. Lichtblau, P. R. Bader, J. M. Rawling, A. B. Protzner, S. Ulrich, K. E. Bloch, B. Giesbrecht, & M. J. Poulin (2018): "Effects on cognitive functioning of acute, subacute and repeated exposures to high altitude." *Front Physiol* **9**: p. 1131.

Raymond-Lezman, J. & S. Riskin (2023): "Benefits and risks of sun exposure to maintain adequate vitamin d levels." *Cureus* **15**: p. 38578.

Roach, T. & J. Whitney (2021): "Heat and learning in elementary and middle school." *Education Economics* **30**: pp. 1–18.

S. Lamb, K.C.S. Kwok, D. W. (2014): "A longitudinal field study of the effects of wind-induced building motion on occupant wellbeing and work performance." *Journal of Wind Engineering and Industrial Aerodynamics* **133**: pp. 39–51.

SCIO (2023): "Obecné studijní předpoklady - testová úloha."

Sharma VM, Pichan G, P. M. (1983): "Differential effects of hot-humid and hot-dry environments on mental functions." *International Archives of Occupational and Environmental Health* **52**: pp. 315–327.

Sobel, M. E. (1982): "Asymptotic confidence intervals for indirect effects in structural equation models." *Sociological Methodology* **13**: p. 290.

Spasova, Z. (2011): "The effect of weather and its changes on emotional state - individual characteristics that make us vulnerable." *Advances in Science and Research* **6**: pp. 281–290.

Srivastava, P., T.-A. Trinh, & X. Zhang (2022): "Weather effects on academic performance: An analysis using administrative data." *Discussion Papers 2207*, University of Exeter, Department of Economics.

Suzuki, M. (2024): "Winter weather on exam dates and matriculation for a prestigious university in japan." *Economics Letters* **237**: p. 111631.

Tian, X., Z. Fang, & W. Liu (2020): "Decreased humidity improves cognitive performance at extreme high indoor temperature." *Indoor Air* **31**: pp. 608–627.

Tsutsumi, H., S. ichi Tanabe, J. Harigaya, Y. Iguchi, & G. Nakamura (2007): "Effect of humidity on human comfort and productivity after step changes from warm and humid environment." *Building and Environment* **42(12)**: pp. 4034–4042.

Weilnhammer, V., J. Schmid, I. Mittermeier, F. Schreiber, L. Jiang, V. Pastuhovic, C. Herr, & S. Heinze (2021): "Extreme weather events in europe and their health consequences - a systematic review." *International Journal of Hygiene and Environmental Health* **233**: p. 113688.

Wolfson, A. R. & M. A. Carskadon (2003): "Understanding adolescent's sleep patterns and school performance: a critical appraisal." *Sleep Medicine Reviews* **7(6)**: pp. 491–506.

World Meteorological Organization (2024): "2024 is on track to be hottest year on record as warming temporarily hits 1.5° c." https://wmo.int/news/media-centre/2024-track-be-hottest-year-record-warming-temporarily-hits-15degc, accessed 2024-12-17.

Yeganeh, A., G. Reichard, A. Mccoy, T. Bulbul, & F. Jazizadeh (2018): "Correlation of ambient air temperature and cognitive performance: A systematic review and meta-analysis." *Building and Environment* **143**: pp. 701–716.

Yin, B., W. Fang, L. Liu, Y. Guo, X. Ma, & Q. Di (2024): "Effect of extreme high temperature on cognitive function at different time scales: A national difference-in-differences analysis." *Ecotoxicology and Environmental Safety* **275**: p. 116238.

Yoorim Choi and Minjung Kim and Chungyoon Chun (2019): "Effect of temperature on attention ability based on electroencephalogram measurements." *Building and Environment* **147**: pp. 299–304.

Zheng, G., K. Li, & Y. Wang (2019): "The effects of high-temperature weather on human sleep quality and appetite." *International Journal of Environmental Research and Public Health* **16**: p. 270.

Zhu, M., Q. Ouyang, H. Shen, & Y. Zhu (2016): "Field study on the objective evaluation of sleep quality and sleeping thermal environment in summer." *Energy and Buildings* **133**: pp. 843–852.

Zivin, J. S. G., S. M. Hsiang, & M. J. Neidell (2015): "Temperature and Human Capital in the Short- and Long-Run." *NBER Working Papers 21157*, National Bureau of Economic Research, Inc.

Appendix A

List of tables

Table A.1: Average test scores per each place of the exam in Czech Republic and Slovakia March 9, 2019

Czech Republic			
City	**Verbal part**	**Analytical part**	**Overall test**
Břeclav	35.91	32.00	33.96
Brno	49.94	45.81	47.88
České Budějovice	54.59	54.09	54.34
Frýdek-Místek	43.07	31.45	37.26
Hradec Králové	52.29	51.69	51.99
Jihlava	52.19	45.73	48.96
Liberec	50.69	52.10	51.39
Olomouc	40.05	37.47	38.76
Opava	36.16	33.13	34.64
Ostrava	43.55	37.42	40.48
Pardubice	50.72	44.61	47.67
Plzeň	46.88	47.61	47.25
Praha	53.77	55.54	54.65
Prostějov	43.02	41.64	42.33
Sokolov	47.77	47.43	47.60
Šumperk	38.80	32.00	35.40
Svitavy	46.50	37.44	41.97
Tábor	54.04	59.98	57.01
Teplice	48.28	48.76	48.52
Uherské Hradiště	42.66	36.16	39.41
Ústí nad Labem	46.96	47.51	47.23
Vsetín	36.00	31.11	33.56
Zlín	39.17	33.12	36.14

Slovakia			
City	**Verbal Part**	**Analytical Part**	**Overall Test**
Banská Bystrica	44.39	40.00	42.20
Bratislava	48.83	45.04	46.93
Košice	47.48	47.96	47.72
Martin	59.60	50.20	54.90
Nitra	38.93	50.29	44.61
Poprad	20.00	78.00	49.00
Prešov	48.36	45.18	46.77
Žilina	47.54	47.39	47.46

Table A.2: Summary statistics for weather conditions during the night before the exam across Czech Republic and Slovakia

	Units	**Min**	**Median**	**Mean**	**Max**
Temperature	Degrees Celsius	1.4	4.63	4.87	5.75
Wind	km/h	4.18	18.63	18.42	24.67

Table A.3: 95% confidence intervals for weather variables by exam part and specification

Weather before the exam

	Verbal	**Analytical**
Temperature	[–0.881, 4.521]	[–2.033, 2.831]
Wind	[–0.040, 0.524]	[–0.245, 0.281]
Few Clouds	[–4.725, 9.579]	[–5.988, 6.618]
Partly Cloudy	[–3.068, 12.592]	[–3.469, 10.239]
Rain	[2.580, 18.550]	[–1.217, 9.121]
Snow	[–11.373, 13.187]	[–16.274, 7.040]

Weather during the exam

	Verbal	**Analytical**
Temperature	[0.472, 4.710]	[0.889, 4.671]
Wind	[0.015, 0.525]	[–0.249, 0.213]
Few Clouds	[–17.020, 0.306]	[–16.862, –0.292]
Partly Cloudy	[–9.849, 3.625]	[–7.531, 4.593]
Rain	[–0.576, 12.050]	[0.015, 10.949]
Snow	[–11.088, 15.338]	[–8.341, 15.169]

Daily average weather

	Verbal	**Analytical**
Temperature	[1.294, 5.128]	[0.635, 4.061]
Wind	[–0.170, 0.332]	[–0.232, 0.168]
Sunshine	[–1.844, 1.386]	[–1.709, 1.787]
Rain	[–0.191, 4.251]	[–0.391, 3.571]

Note: Confidence intervals are calculated as $\hat{\Gamma} \pm 1.96 \times \text{s.e.}(\hat{\Gamma})$.

Table A.4: 95% confidence intervals for weather variables by gender and specification

Weather before the exam

Variable	**Men**	**Women**
Temperature	[–1.47, 5.75]	[–2.55, 3.18]
Wind	[–0.46, 0.31]	[–0.07, 0.54]
Few Clouds	[–6.80, 12.42]	[–6.16, 8.87]
Partly Cloudy	[–8.82, 11.93]	[–1.78, 14.62]
Rain	[2.58, 18.55]	[–3.42, 9.04]
Snow	[–17.51, 17.25]	[–17.24, 11.02]

Weather during the exam

	Men	**Women**
Temperature	[–0.55, 5.16]	[0.70, 5.15]
Wind	[–0.31, 0.40]	[–0.10, 0.44]
Few Clouds	[–16.18, 7.89]	[–19.65, –0.75]
Partly Cloudy	[–10.69, 7.19]	[–9.79, 4.56]
Rain	[0.15, 16.18]	[–3.50, 10.13]
Snow	[–17.14, 17.12]	[–9.67, 18.97]

Daily average weather

	Men	**Women**
Temperature	[0.15, 4.27]	[0.99, 5.11]
Wind	[–0.14, 0.38]	[–0.37, 0.15]
Sunshine	[–1.71, 2.47]	[–2.23, 1.95]
Rain	[–0.39, 4.49]	[–0.83, 4.07]

Note: Confidence intervals are calculated as $\hat{\Gamma} \pm 1.96 \times \text{s.e.}(\hat{\Gamma})$.

Table A.5: Regression for total exam score for daily average weather conditions without outliers

	Dependent variable: Overall score		
	(1)	(2)	(3)
Temp	3.343***	2.539**	3.162***
	(0.858)	(0.852)	(0.901)
Wind	−0.065	−0.042	−0.084
	(0.104)	(0.106)	(0.108)
Sun	0.147	0.241	0.293
	(0.845)	(0.873)	(0.873)
Rain	2.344**	2.357*	2.770*
	(0.967)	(1.140)	(1.156)
Results 2019	0.940***	0.919***	0.953***
	(0.096)	(0.097)	(0.098)
Men	8.331***	8.370***	8.344***
	(0.929)	(0.940)	(0.943)
Constant	−16.244***	−11.792*	−16.033**
	(5.502)	(5.142)	(5.517)
Observations	2,754	2,701	2,680
R^2	0.100	0.100	0.102
Adjusted R^2	0.098	0.098	0.100
Residual Std. Error	24.002	24.037	24.029
F Statistic	51.110***	50.036***	50.687***

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
The results are based on Equation (4.2).
Columns represent separate specifications using (1) weather conditions before the exam, (2) during the exam, and (3) average weather conditions throughout the exam day.

Table A.6: Regression results for total exam for hourly and daily average weather conditions excluding Slovakia

	Dependent variable: Overall score		
	(1)	(2)	(3)
Temp	0.407	2.841**	2.762***
	(1.410)	(1.078)	(0.830)
Wind	−0.039	0.046	−0.047
	(0.153)	(0.160)	(0.104)
Sun	-	-	0.151
			(1.089)
Few Clouds	4.195	−7.871	-
	(3.365)	(4.958)	
Partly Cloudy	6.948·	−2.579	-
	(3.917)	(3.799)	
Rain	6.166*	4.660	1.829·
	(2.567)	(3.212)	(0.942)
Snow	−3.383	2.460	-
	(6.054)	(5.726)	
Results 2019	0.874***	0.786***	0.926***
	(0.120)	(0.148)	(0.096)
Men	8.594***	8.638***	8.523***
	(0.936)	(0.936)	(0.937)
Constant	−3.474	−8.957	−12.800*
	(7.924)	(8.461)	(5.155)
Observations	2,707	2,707	2,715
R^2	0.1027	0.1031	0.1013
Adjusted R^2	0.096	0.1005	0.0993
Residual Std. Error	24.02	24.01	24.03
F Statistic	38.71***	38.89***	50.88***

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '·' 0.1 ' ' 1
The results are based on Equation (4.2).
Columns represent separate specifications using (1) weather conditions before the exam, (2) during the exam, and (3) average weather conditions throughout the exam day.

Table A.7: Regression results for total exam score for daily average weather conditions excluding snow

	Dependent variable: Overall score	
	(1)	(2)
Temp	1.257	2.044***
	(0.928)	(0.597)
Wind	0.103	0.131
	(0.122)	(0.106)
Few Clouds	2.008	−4.651*
	(3.024)	(2.019)
Partly Cloudy	4.204	17.357
	(3.258)	(24.023)
Rain	6.317**	5.818***
	(2.183)	(1.679)
Results 2019	0.796***	0.751***
	(0.108)	(0.135)
Men	8.504***	8.535***
	(0.919)	(0.918)
Constant	−5.270	−6.999
	(5.967)	(7.171)
Observations	2,818	2,818
R^2	0.099	0.100
Adjusted R^2	0.097	0.098
Residual Std. Error	23.99	23.98
F Statistic	44.09***	44.66***

Significance: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
The results are based on Equation (4.2).
Columns represent separate specifications using (1) weather conditions before the exam, (2) during the exam, and (3) average weather conditions throughout the exam day.