A Work Project, presented as part of the requirements for the Award of a Master's degree in Economics from the Nova School of Business and Economics.

# CLIMATE CHANGE IN THE CLASSROOM: TEMPERATURE AND EDUCATIONAL ACHIVEMENT IN AFRICA

ASHLYN NICOLE OSENDORF

Work project carried out under the supervision of:

Dr. Wayne Sandholtz

16/12/2022

## Abstract

Despite the fact that human capital accumulation is a key facilitator of economic growth and climate change is projected to increase temperatures, there has been little research into the contemporaneous effect of heat on educational achievement. Thus, I investigate how high temperatures affect high-stakes exam performance in a developing context, employing climate reanalysis data and test scores from Tanzania’s National Primary School Leaving Exam. Through the use of fixed effects models and the exploitation of quasi-random variations in outdoor temperatures during exams, I find that higher temperatures exert a negative impact on student performance.

Keywords: Education, Development, Africa, Climate Change, Human Capital

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

# 1 Introduction

Human-induced climate change has already begun to cause dangerous and widespread distortions in nature, most notably through increased heatwaves, floods, droughts, and weather variability, all of which are projected to worsen over the coming decades (Intergovernmental Panel on Climate Change 2022). As such, a significant body of economic research has become dedicated to exploring the impacts of temperature on a variety of outcomes, such as agriculture (Mendelsohn et al. 1994; Schlenker et al. 2006; Deschenes and Greenstone 2007), conflict (Burke et al. 2009), economic output (Hsiang and Solomon 2010; Dell et al. 2012), mortality (Deschenes and Greenston, 2011; Barreca et al. 2016), and crime (Ranson 2014). In particular, research on the intersection of temperature and development has become especially important, as it is developing nations which have been, and will continue to be, the most impacted by the climate crisis. One such area of interest is the impact of temperature on education, or the impact of temperature on cognitive performance in the educational setting, which has been comparatively underexplored. The importance of the subject is substantiated by the fact that education, or rather the accumulation of human capital, is a key driver of economic growth (Barro 2001; Nelson and Phelps 1966; Romer 1986). Thus, the aim of this work is to contribute knowledge to this area by investigating the impact of high temperatures on cognitive performance on a high-stakes exam taken across Tanzania.

In general, Tanzania is a fitting location of study given both its developing status and susceptibility to climate change. In 2021, Tanzania was ranked 160 out of 191 countries and territories on the Human Development Index (United Nations 2022), with a real GDP per capita of 1,135.5 US$ (World Bank 2021) and a poverty headcount ratio of 44.9% (World Bank 2018). The annual number of very hot days in Tanzania (those with a maximum temperature of at least 35 °C) is projected to rise substantially over the following century. In particular, eastern Tanzania is projected to experience about 100 very hot days per year by 2080 (Deutsche Gesellschaft für Internationale Zusammenarbeit 2021).

To examine the impact of heat on exam outcomes across Tanzania, I employ climate reanalysis data and the test scores of more than 3 million children from the high-stakes National Primary School Leaving Exam, exploiting quasi-random variations in temperature across exams. In particular, I use fixed effects models, which address the issue of endogeneity and allow students to be compared against themselves on hotter days, netting out variation in ability, geography, or socioeconomic status. I also employ fixed effects which account for the subject of a given exam and the day of the week on which it was taken.

Ultimately, I find that higher temperatures reduce both standardized exam scores and the likelihood of passing a given exam. This effect is significant at the 1% significance level and holds across various fixed effects specifications, alternative clustering of standard errors, and relaxing the assumption of linearity with both a piece-wise function and nonparametric binned approach. To my knowledge, this is the first work to provide evidence for the physiological effects of heat stress on the day of a high-stakes assessment in Africa.

The rest of this paper is organized according to the following. Section 1.2 discusses the relevant literature. Section 2 describes both the test and weather data employed. Section 3 explains the empirical strategy. Section 4 presents the results of the on the impact of temperature on educational achievement. Section 5 verifies key assumptions. Section 6 discusses implications and concludes.

## 1.1 Literature Review

Throughout the world, standardized exams have become increasing prevalent, given that they are a cost-effective measure of students' ability and knowledge acquisition, especially for students who are high-achieving but may be economically disadvantaged

(Hyman, 2017). They also allow for the comparison of test-takers around the world and at different points in time. These exams are often high-stakes, dictating government financial transfers to schools, teacher compensation, student placement, admission to university, and financial aid. Therefore, it is valuable to understand the impact of physical conditions on key assessments, particularly when they may be in the process of shifting. As such, temperature stress has become a particular condition of interest, given the projected increase in both average and extreme temperatures.

The impact of temperature on exam performance has been studied in both the long and short run. In the long run, Cho (2017) finds a negative effect of exposure to high temperatures in the summertime on exam scores in South Korea. Park et al. (2021) find that heat exposure during learning periods negatively impacts student performance on the PISA across more than 50 countries. Roach and Whitney (2022) show that when cooler areas of the United States experience unusually hot years, student achievement on standardized exams is reduced. One explanation for this phenomenon may be that temperature impacts time allocation. In fact, Alberto et al. (2021) show that in response to extreme temperatures (both cold and hot), students' reduce time allocation for studying and increase leisure. In the developing context, Garg, Jagnani, and Taraz (2020) show additional hot days throughout the year prior to an exam reduce exam scores in India. The authors argue this is primarily due to temperature's negative effects on already low agricultural productivity and incomes, which are especially vulnerable to variations in precipitation and temperature. However, other theories surrounding the long run effect have not yet been ruled out, such as heat stress affecting school closures, teacher absenteeism, and the prevalence of diseases which thrive in hot climates.

In the short run, however, high temperatures are thought to impact academic performance through a physiological mechanism. This is because the chemistry, electrical properties, and functions of the brain are all known to be sensitive to temperature (Bowler and Tirri 1974; Deboer 1998; Hocking et al. 2001; Schiff and Somjen 1985; Yablonskiy, Ackerman and Raichle 2000). At the same time, the temperature of the brain itself rises as a result of both external temperature and cognitive demands. In fact, labratory experiments have shown that heat stress diminishes attention, memory, the retention and processing of information, and the ability to perform psycho-perceptual tasks (Vasmatzidis et al. 2002). Additionally, performance on both cognitive and physical tasks has also been shown to be negatively impacted by temperature (Seppanen et al. 2006). However, it was previously not certain if these findings, which took place in artificial, experimental environments, would translate into visible effects in the real world. Additionally, the stakes of these studies were low, making it difficult disentangle the effect of temperature on performance from participants simply reducing effort or concentration, given that the disutility of exerting effort is likely to increase as temperatures get hotter.

Graff Zivin et al. (2018) overcame the first issue by studying cognitive performance on a survey of youth in the United Sates. Here, they found that hot temperature on the day of the survey reduced math (but not reading) performance. Similarly, Garg, Jagnani, and Taraz (2020) also examined the short run effect of temperature on the surveys administered to Indian students. While both these studies confirmed the negative impact of heat on cognitive performance in the real world, they both took place in low-stakes environments, rendering the findings difficult to generalize. However, by studying the impact of temperature on the outcomes of a key exam for high school students in New York City, Park (2022) proved the contemporaneous impact of heat stress on performance in a setting with economically meaningful stakes. The magnitude of the effect, though, was much smaller than Garg, Jagnani, and Taraz (2020), further indicating that the stakes of an assessment may be important in capturing the true effect of temperature. Next, Graff Zivin et al. (2020)

investigated exam-day temperature on performance on China's National College Entrance Examination (NCEE), becoming the first to provide nation-wide estimates for a high-stakes environment within a developing country. They found a similar effect to Park (2022). Finally, by exploiting a change in the stakes of a key Brazilian exam, Melo and Suzuki (2022) verify the theory that when the stakes of an exam are increased, the negative effect of temperature on exam scores is reduced.

Ultimately, this work will contribute to the literature by offering two unexplored perspectives. Firstly, it will investigate the contemporaneous impact of high temperatures on exam performance within a high-stakes environment in Africa, which has not yet been done. It will also be the first to examine the impact of temperature on a high-stakes national assessment given exclusively to primary school children. While Roach and Whitney (2022) do study the impact of temperature on American elementary and middle school students, the exams used in their study do not necessarily have stakes which are clear and consistent for students across the country. Additionally, Roach and Whitney (2022) study the impact of hot years on exam performance rather than the impact of exam-time temperature, thus making another key distinction between the two works.

# 2 Data

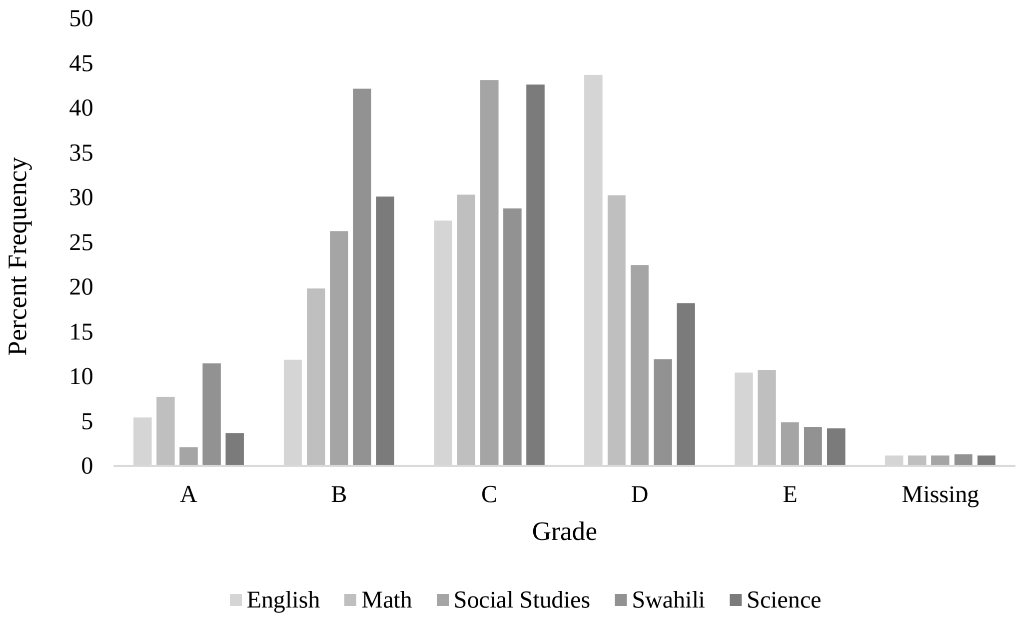
## 2.1 Student Data

Following the completion of 7 years of primary education, students across Tanzania take the Primary School Leaving Examination (PSLE). This is a national standardized test, and passing it is a requirement for admission into public secondary school. For reference, primary net enrollment in Tanzania is 81%, while secondary net enrollment is only 27% (The World Bank 2022). The dates of the PSLE are fixed by the National Examinations Council of Tanzania (NECTA), eliminating the possibility of self-selection on test dates. While the exams always take place during the month of September, the specific dates vary by year. It is also worth noting that exams are brought to a central location for grading, eliminating the possibility of systematic manipulation in grading, such as that seen in Park (2022).

The PSLE consists of five sections: Mathematics, Science, English, Kiswahili, and Social Studies. Each section has a maximum score of 50 points, with the letter grades A, B, C, D, and E corresponding to minimum points of 40, 30, 20, 10, and 0 respectively. After taking all five exams, the points are added up for an overall score of 250, and students must earn 100 points or more to pass. In other words, students must receive at least a C on average across exams. For the ease of interpretation, I follow the strategy of Sandholtz (2020), where the scores are converted to a 100-point scale.

In this work, I use the PSLE test data of over 3 million students from more than 17,000 schools across Tanzania from 2016-2019, which are the years in which information regarding exact days and times of exams is available. Additionally, the sample is restricted to students with unique names within their exam year. Around 98% of students satisfy this condition, so this does not significantly reduce the sample and should not bias the results. The student data is reshaped so that each student has an observation for each of their exams, allowing me to connect each observation to the temperature at the beginning of the exam. This results in a final sample of almost 17 million observations. Overall, the pass rate for the sample is 75.12%, with an average of 49.52 total points earned. Figure 1 shows the distribution of exam scores by subject across the sample.

**Figure 1:** *Distribution of Grades by Subject*



## 2.2 Temperature Data

Ideally, temperature data would be observational, taken from actual ground stations across Tanzania. However, the spatial and temporal coverage of ground stations across developing countries is often poor, and Tanzania is no exception. Therefore, I instead use hourly temperature reanalysis data generated by the European Centre for Medium Term Weather Forecasting, which is accessible through the ERA-Interim archive. In essence, climate reanalysis provides estimates of atmospheric parameters through the combined use of climate models and real observations. According to the literature, reanalysis data consistently estimates weather in a grid-cell; it has also been widely used within economics literature[1] (Auffhammer et al. 2013; Schlenker and Lobell 2010). In fact, reanalysis

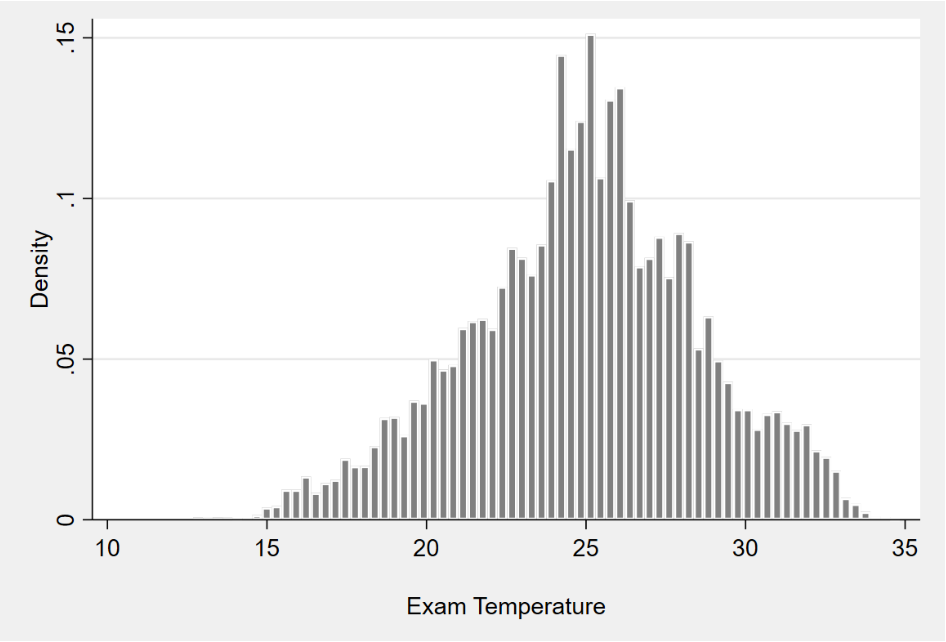
---

[1] For reference, Dee et al. (2011) discusses the methodology and construction of data from the ERA-Interim.

temperature data from ERA-Interim archive was also used by Garg, Jagnani, and Taraz (2020) in India.

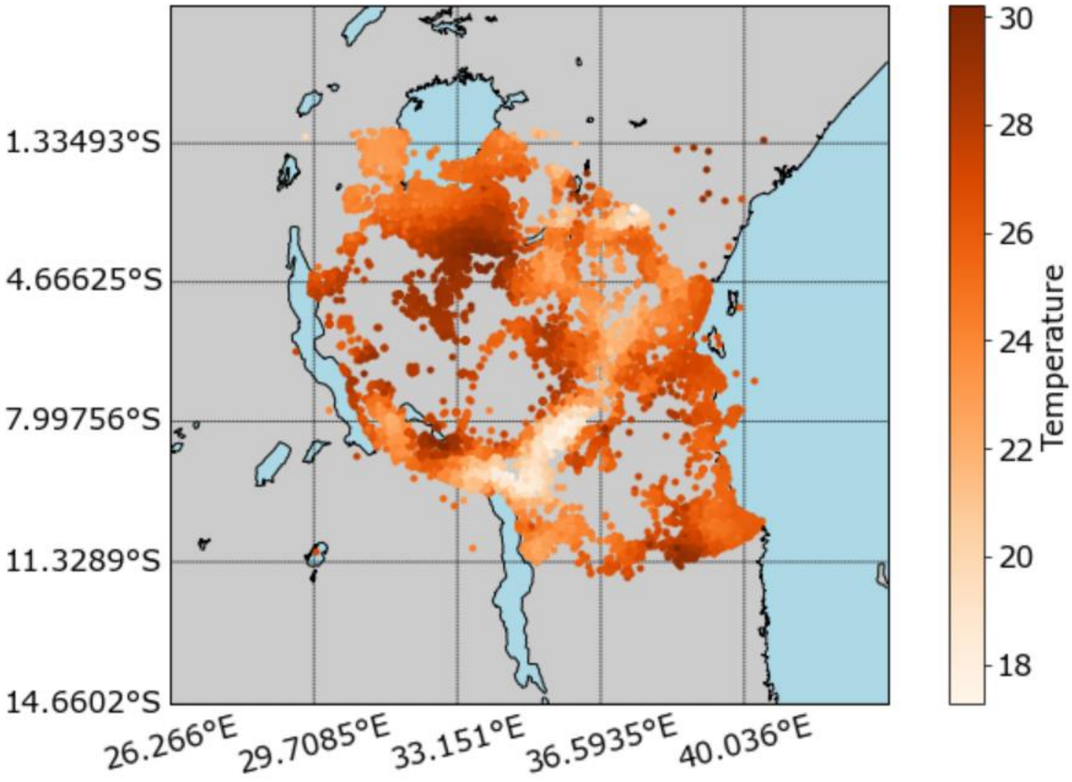
I specifically use the ERA-Interim hourly temperature and control data on a 1 x 1 degree latitude-longitude grid for the day and hour of each subject exam across the sample years. The weather and air quality controls employed are precipitation, dew point temperature, relative humidity, air pressure, wind speed, and ozone levels, following their use in key papers such Park (2022) and Zivin et al. (2020). The compiled dataset provides information from around 2000 points across Tanzania. Students are assigned a point which is closest to the coordinates of their school. On average, the distance from a school to its closest weather point is 11km. Figure 2 shows the distribution of temperatures at the start times of PSLE exams from 2016-2019. The temperatures in the sample range from around 12 ºC to 35 ºC, with an average temperature of 25ºC.

**Figure 2:** *Temperature Distribution*



Once students are assigned a weather point based on the location of their school, they are matched with the appropriate temperature and weather controls based on the year, day, and time of each of their exams. Figure 3 shows the average temperature at each of the more than 17,000 schools over the 2016-2019 period based on the matching technique.

**Figure 3:** *Average Temperature at Schools*



## 3 Empirical Strategy

The empirical strategy aims to uncover a causal relationship between temperature and exam performance. Importantly, the strategy relies on the PSLE being a high-stakes assessment for the students, rendering the environment economically meaningful. As such, I make the assumption that students' disutility from higher temperatures and effort exertion during the exam is offset by the high-stakes nature of the PSLE This should eliminate any extensive margin response and limit selection bias. Therefore, temperature stress should not simply reduce effort or concentration, and the estimates should represent the true effect of temperature on performance. It is also important to note that temperature is plausibly exogenous given that it varies across the different exam dates and times, which were defined ahead of time by the examination board.

To explore the effect of temperature, I estimate a series of linear fixed effects regression models. In this case, the use of a naive OLS model would be inappropriate, producing biased and inconsistent parameter estimates. This is because of unmeasured differences among students which may be associated with exam outcomes. Fixed effects models, however, allow individual-specific effects to be correlated with the independent variables. In essence, the fixed effects strategy will allow me to compare students against themselves on hotter days, netting out variation in ability, geography, or socioeconomic status. Naturally, I assume that any relevant individual-specific effects are non-varying over time. The general specification takes the following form:

$$Y_{ijsty} = \gamma_i + \eta_s + \beta_1 T_{jsty} + B_2 X_{jsty} + B_3 Time_{sty} + \mathrm{B}_4 \mathrm{DOW}_{\mathrm{sty}} + \epsilon_{\mathrm{ijsty}}$$

Employing the empirical strategy of Park (2022), the above specification uses $Y_{ijsty}$ to denote the standardized exam performance for student *i* taking an exam in subject s in school $j$ on date $t$ in year $y$. Next, $\gamma_i$ and $\eta_s$ represent student and subject fixed effects, respectively. The subject fixed effects take into account that different subjects may inherently yield different score outcomes, either due to their difficulty for students or the fact that they are consistently scheduled for a certain time. Ultimately, both fixed effects represent the time-invariant unobservable characteristics, counteracting potential omitted variable bias. Next, $T_{jsty}$ represents the outdoor temperature at the weather point nearest to school *j* during the exam for subject *s* on the appropriate date *t* in year *y*. $X_{jsty}$ is a vector of weather and air quality controls, which include precipitation, dew point temperature, relative humidity, air pressure, wind, and ozone levels. As previously mentioned, these particular controls were selected following their use in Park (2022) and Zivin et al. (2020). Next, $Time_{sty}$ is a dummy variable indicating the time of day in which an exam occurred (Time=0 represents an evening exam, while Time=1 is in the afternoon), and $\mathrm{DOW}_{\mathrm{sty}}$ denotes a vector of fixed effects for weekday exams were taken on.

# 4 Results

## 4.1 Primary Results

The result of running the previously discussed specification is presented in column 1 of Table 1, where the coefficient represents the change in z-score when temperature during the exam increases by 1˚C. To account for potential serial correlation within a school over time, robust standard errors are clustered by school and shown in parentheses. The results of the original specification indicate that heat negatively impacts student performance, as all of the estimates are negative and statistically significant at the 1% significance level. More specifically, temperature reduces exam performance by -0.016 standard deviations per degree Celsius. This represents a decline of 5.8 percent of a standard deviation in performance per standard deviation increase in the temperature (+3.6°C) at the time of the exam, indicating a 14.4 percent standard deviation decline if a student takes an exam when it is 33°C compared to 24°C. These values are similar to the findings of Park (2022), where a one standard deviation increase in exam-time temperature amounts to a decrease of 5.5 percent of a standard deviation in performance, or -13 percent of a standard deviation if a student takes an exam when it is 90°F (32.2 ºC) outside as opposed to a more optimal 75°F (23.9ºC) (Park 2022). Furthermore, the results also align with the finding of Zivin et al. (2020) in China, where a one-standard-deviation increase in temperature during the exam period within counties decreases the total test score by 5.83% of a standard deviation.

Columns 2 and 3 of Table 1 show the results of two robustness checks, where student fixed effects are replaced by school and year and school by year fixed effects respectively, following their use in Park (2022). Across the different specifications, the significantly negative point estimate remains. Further, the effect is not sensitive to alternative clustering of standard errors, including by school and date, weather point and date, weather point and year, or by ward, all of which can be found in Appendix Table A.1.

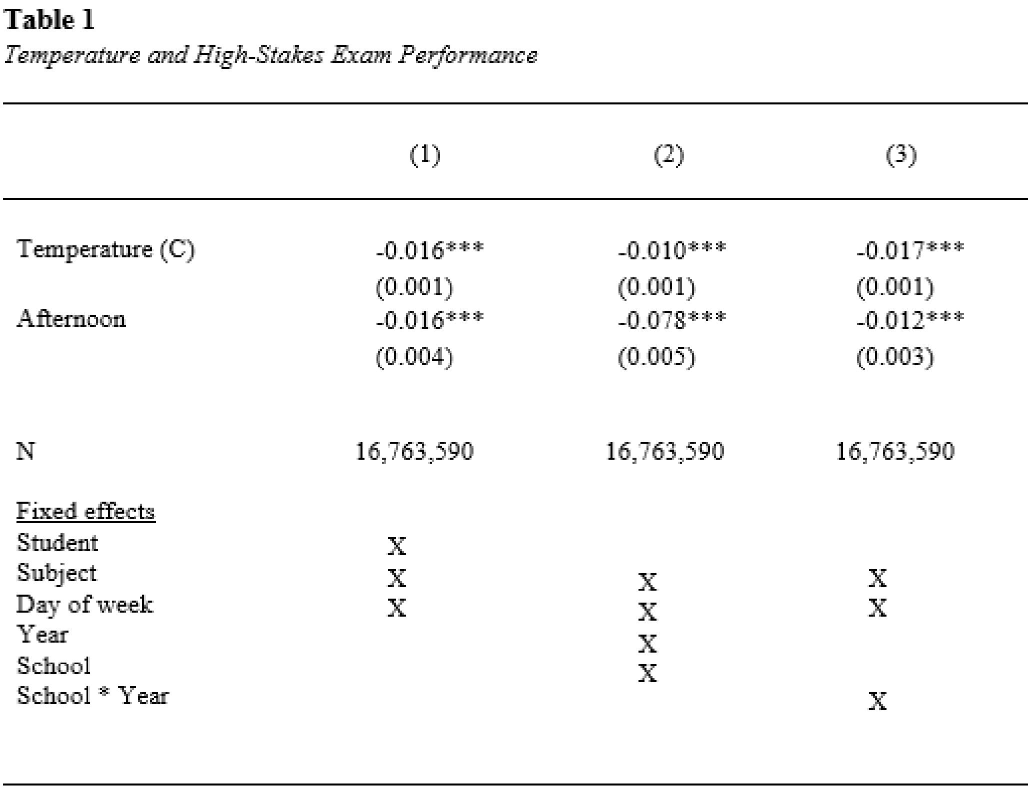


**Table 1**

*Temperature and High-Stakes Exam Performance*

| | (1) | (2) | (3) |
|---|---|---|---|
| Temperature (C) | -0.016***<br>(0.001) | -0.010***<br>(0.001) | -0.017***<br>(0.001) |
| Afternoon | -0.016***<br>(0.004) | -0.078***<br>(0.005) | -0.012***<br>(0.003) |
| N | 16,763,590 | 16,763,590 | 16,763,590 |
| Fixed effects | | | |
| Student | X | | |
| Subject | X | X | X |
| Day of week | X | X | X |
| Year | | X | |
| School | | X | |
| School * Year | | | X |

It should be noted that the meaning of the afternoon dummy differs compared to Park (2022), where students either take exams in the morning or afternoon. For students in New York City, taking an exam in the afternoon a has multiple disadvantages in that the temperatures at this time are higher and students are relatively more fatigued from their morning exams. However, in the case of the PSLE, students either take exams in the afternoon (from 14:00-15:30 or 16:30-18:30) or in the evening (20:30-22:00). On the one hand, afternoon tests are hotter, with an average exam-time temperature of 25.8ºC compared to an average of 21.6ºC for evening exams. However, students taking the PSLE are more fatigued in the *evening* after having already taken other exams. Nevertheless, I find that taking a test in the afternoon has a significantly negative point estimate, similar to Park (2022).

The work of Graff Zivin et al. (2018) found differences in the effect of temperature on math versus reading outcomes, which inspired Park (2022) to run models which separated quantitative versus verbal subjects. As a potentially interesting aside, I run the models again, separating math and science from English, Swahili, and social studies. The results may be found in Appendix Table A.2. Similar to both works, I find that the point estimate is more negative on quantitative subjects. However, unlike them, I still find that the coefficients on the verbal models are distinguishable from zero. One potential explanation for math subjects having a more negative estimate is due to the increased heat sensitivity of the prefrontal cortex, which is a region required more for quantitative rather than verbal tasks (Hocking et al. 2001; Kiyatkin 2007).

### 4.2 Linear versus nonlinear impacts

To access the potential for non-linear effects, I employ two approaches from the literature. The first is known as degree days (DD), which uses a piece-wise linear function that measures an observation's number of degrees above or below a certain threshold. In this case, the threshold is 26°C, meaning that the variable "DD >=26" is equal to an observation's temperature subtracted by 26°C, conditional on it being over the threshold. As such, it is equal to zero if it is below the threshold. Similarly, the "DD <26" variable is equal to an observation's temperature subtracted by 26°C if it is under the threshold, or zero if above the threshold. According to the literature, the degree days measure is useful when studying the impacts of temperature if the response is roughly constant but changes nonlinearly in temperature, or if the response can be accurately estimated by a piece-wise linear function, with kinks at the threshold temperature (Graff Zivin et al. 2018). In this case, the threshold of 26°C was chosen for being the local maximum in the nonparametric approach, as can be seen in Figure 4.

The nonparametric specification employs a full set of indicator variables for every 2°C bin, as is the standard across the literature (Deschenes and Greenstone 2011 ; Barreca et al. 2016; Graff Zivin and Neidell 2014; Graff Zivin and Neidell 2018). Due to data sparseness at the extremities of the distribution, the lowest bin includes all observations with temperatures below 16˚C and the highest bin includes all those with temperatures above 32˚C, as was done in Zivin et al. (2020). For the non-linear binned approach, the coefficient represents change in z-sore when the exam-time temperature is within that 2°C bin relative to the reference bin.

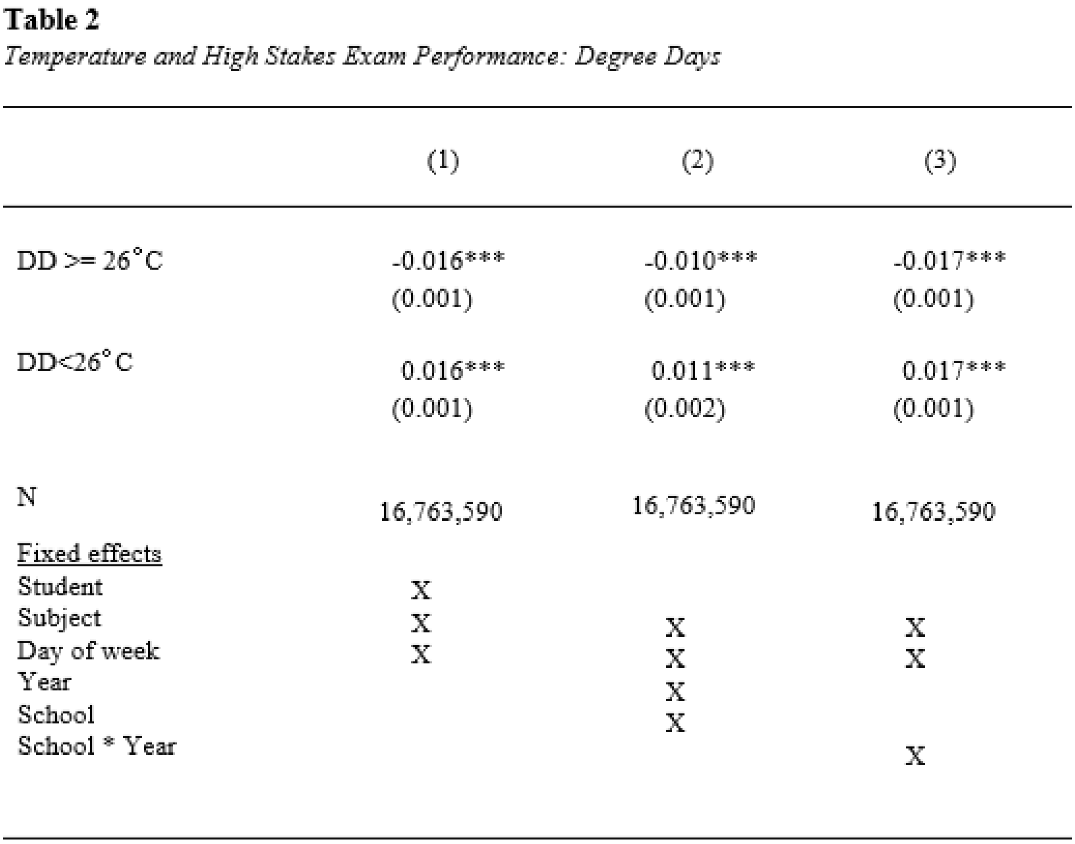


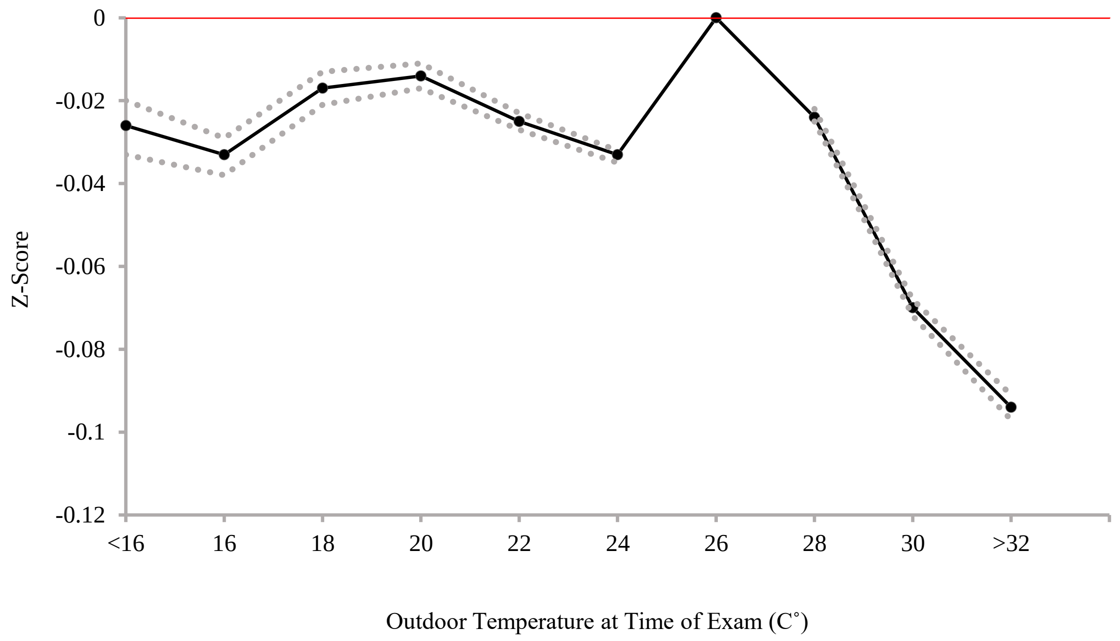
**Table 2**

*Temperature and High Stakes Exam Performance: Degree Days*

| | (1) | (2) | (3) |
|---|---|---|---|
| DD >= 26°C | -0.016*** | -0.010*** | -0.017*** |
| | (0.001) | (0.001) | (0.001) |
| DD<26°C | 0.016*** | 0.011*** | 0.017*** |
| | (0.001) | (0.002) | (0.001) |
| N | 16,763,590 | 16,763,590 | 16,763,590 |
| Fixed effects | | | |
| Student | X | | |
| Subject | X | X | X |
| Day of week | X | X | X |
| Year | | X | |
| School | | X | |
| School * Year | | | X |

Relaxing the linearity assumption, Table 2 shows the results of specifying temperature in terms of degree days. The coefficient of DD >= 26 measures the change in z-score if temperature increases by 1˚C conditional on the temperature being above 26˚C. Likewise, the coefficient of DD < 26 represents the change if temperature decreases by 1˚C conditional on the temperature being below 26˚C. Once again, robust standard errors are clustered by school and displayed in parentheses. Here, the effect of DD >= 26 is significantly negative, with a magnitude nearly identical to the linear effects of the fixed effects in Table 1.

**Figure 4:** *Coefficients of Non-Parametric Binned Approach*



For the non-linear binned approach, the coefficient decreases monotonically for all bins hotter than 26˚C. This can be seen in Figure 4, which plots the coefficients and 95% confidence intervals when the dependent variable is the exam z-score and the students fixed effects model is employed. Similar to the degree days method, the magnitude here is also comparable to the effects seen in column 1 of Table 1. For example, the estimated coefficient for the above 30˚C bin is -0.071. Since the difference between this bin and the reference bin (26˚C) is approximately 4˚C, each 1˚C increase in temperature decreases the z-score by 0.018 (-0.07/4), as compared to 0.016 in column 1 of Table 1. This aligns with the results seen in Zivin et al (2020). Furthermore, the relationship between temperature and exam performance plotted in Figure 4 follows roughly the same pattern as seen with math performance in Graff Zivin et al. (2018). Interestingly, the effect of temperature on math scores in Graff Zivin et al. (2018) was only significant beyond 26ºC, despite having a different threshold.

## 4.3 Temperature and Pass/Proficiency Status

In this specification, the outcome variable of standardized exam scores is replaced with a dummy variable for whether or not students passed a given subject, defined as receiving a C or better, as this is the average required to pass the PSLE overall. The results of this are found in Table 3, once again using robust standard errors clustered by school. The results indicate that heat stress reduces the likelihood of passing a given subject exam. As before, all of the estimates are statistically significant at the 1% significance level.

Here, a one standard deviation (3.6°C) increase in temperature results in a 3.2 percent lower probability of receiving a passing score on a given exam. This translates to a 1.4 percentage point decline per degree Celsius, relative to a mean likelihood of 66 percent. Therefore, taking an exam when outdoor temperature is 33°C results in a 12.6 percent lower chance of receiving a passing score on a given exam relative to a day with a temperature of 24°C. Again, columns 2 and 3 test the robustness of the finding to alternative specifications by replacing student fixed effects with school and year along with school-by-year fixed effects. Across models, the point estimates remain roughly similar and equally significant.

These results are in line with, although slightly higher than, the findings of Park (2022), where a one standard deviation increase in temperature resulted in a 2.4 percent lower probability of passing (relative to the mean), and taking an exam when it is 90°F outside results in a 10 percent lower chance of passing a given exam relative to it being 75°F (Park 2022). Interestingly, the afternoon dummy now has a positive effective, indicating that students may be more likely to pass when less fatigued from other exams. Park (2022) also found a different effect for the afternoon variable when the outcome is a passing dummy.

However, the coefficients were still negative, although only significant at the 10% for the student fixed effects models, and not significant at all for the school effects models.

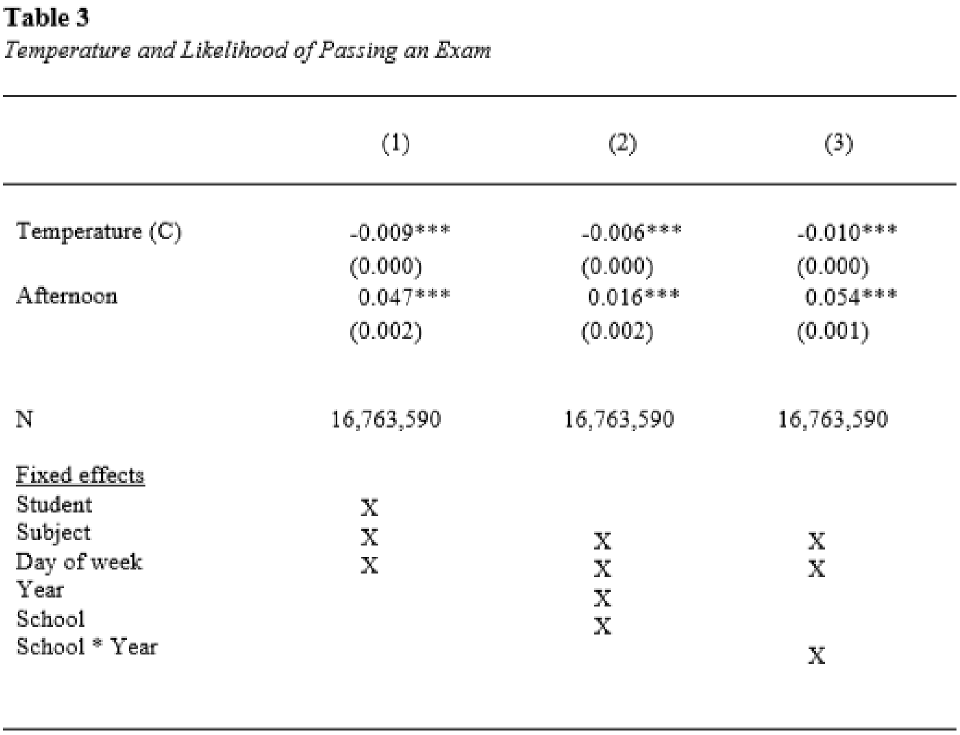


**Table 3**

*Temperature and Likelihood of Passing an Exam*

| | (1) | (2) | (3) |
|---|---|---|---|
| Temperature (C) | -0.009*** | -0.006*** | -0.010*** |
| | (0.000) | (0.000) | (0.000) |
| Afternoon | 0.047*** | 0.016*** | 0.054*** |
| | (0.002) | (0.002) | (0.001) |
| N | 16,763,590 | 16,763,590 | 16,763,590 |
| Fixed effects | | | |
| Student | X | | |
| Subject | X | X | X |
| Day of week | X | X | X |
| Year | | X | |
| School | | X | |
| School * Year | | | X |

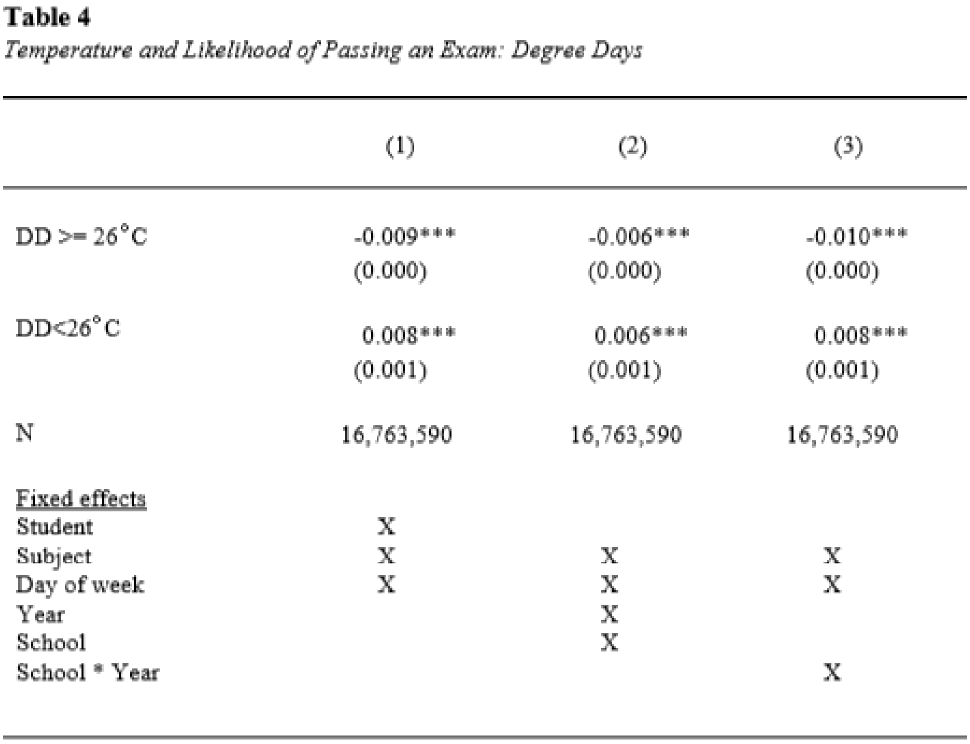


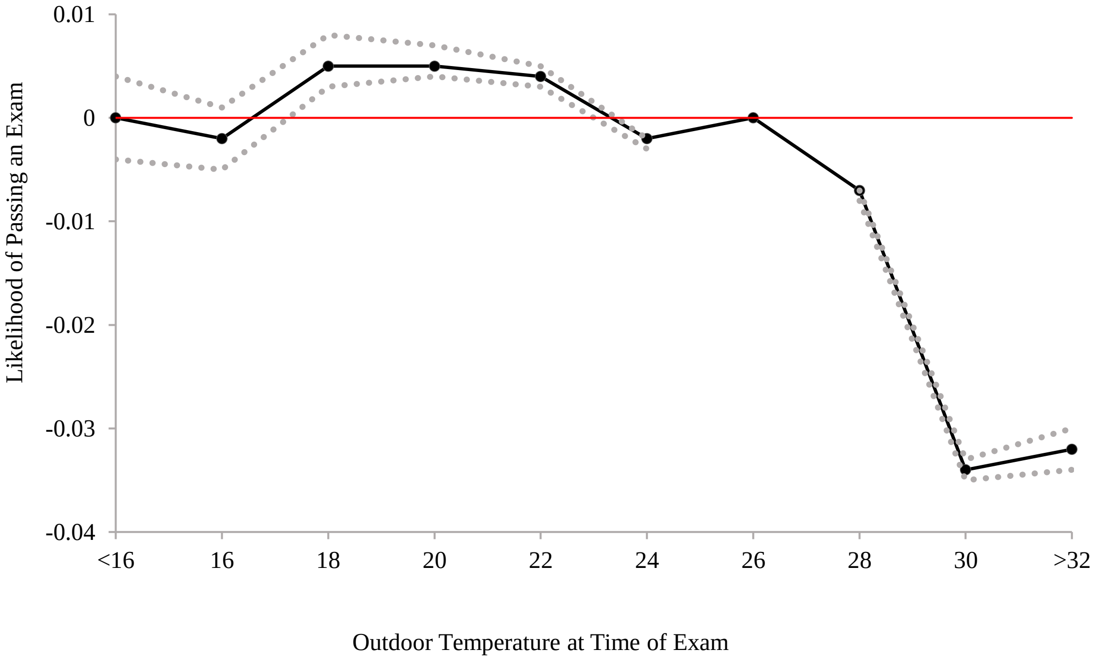
**Table 4**

*Temperature and Likelihood of Passing an Exam: Degree Days*

| | (1) | (2) | (3) |
|---|---|---|---|
| DD >= 26°C | -0.009*** | -0.006*** | -0.010*** |
| | (0.000) | (0.000) | (0.000) |
| DD<26°C | 0.008*** | 0.006*** | 0.008*** |
| | (0.001) | (0.001) | (0.001) |
| N | 16,763,590 | 16,763,590 | 16,763,590 |
| Fixed effects | | | |
| Student | X | | |
| Subject | X | X | X |
| Day of week | X | X | X |
| Year | | X | |
| School | | X | |
| School * Year | | | X |

Once again, I also access the potential for nonlinear effects through both the degree days and nonparametric binned approach. The results of the former are shown in Table 4. I again find temperature to have a significantly negative point estimate of a similar magnitude to the linear approach. For the nonparametric specification, Figure 5 reveals that the coefficients are decreasing for bins higher than the threshold. Again, the coefficients are roughly similar to the other approaches. For example, the coefficient on the 30ºC bin is -.034. Since the difference from reference bin is around 4 degrees, the magnitude is the same as the linear approach (-.034/4=-.009). Additionally, the shape of this graph follows a similar pattern as the relationship between temperature and the probability of getting into first-tier universities in Zivin et al. (2020), which also uses a dummy outcome variable and the nonparametric binned approach.

**Figure 5:** *Coefficients of Non-Parametric Binned Approach*



# 5 Verifying Assumptions

In this section, I will exploit unique features of the PSLE data in order to test some of the key assumptions of the conceptual framework. Firstly, the high-stakes empirical setting should eliminate any extensive margin response (absenteeism). This is a key assumption for limiting potential selection bias and ensuring that the effect of temperature is not simply due to students reducing effort or concentration. I utilize the PSLE data on exams missed by students to test this. Overall, only 1.2% of exams are missed by students, supporting the presence of a high-stakes environment. To test the impact of temperature of absenteeism, I run regressions similar to the main specification, simply replacing the outcome variable with a dummy for whether a student appeared for their exam. Here, temperature had a coefficient of 0.000, which was statistically insignificant. In the model with student fixed effects the p-value was equal to 0.761 while in the model with school by year fixed effects the p-value was 0.643.

Next, I attempt to further confirm the impact of exam stakes on the effect of temperature. The PSLE exam is unique compared to the high-stakes exams seen in Park (2022) and Zivin et al. (2020) in that students are only required to receive a C average overall, without receiving any added benefit for higher scores. In both Park (2022) and Zivin et al. (2020), the scores of the exams are used for university admissions, meaning that students are likely aiming to receive the highest score possible, and therefore exert the proper effort and concentration to do so. However, this is not the case with the PSLE, meaning that the disutility of additional effort past the passing threshold is higher. According to the theory, this means temperature should have a greater negative impact on the probability of receiving a B on a given exam, and an even greater negative effect on the probability of receiving an A. To test this, I run two sets of regressions with dummy outcome variables for whether a student received a B or an A on a given exam, relative to the mean. The results are presented in Table 5.

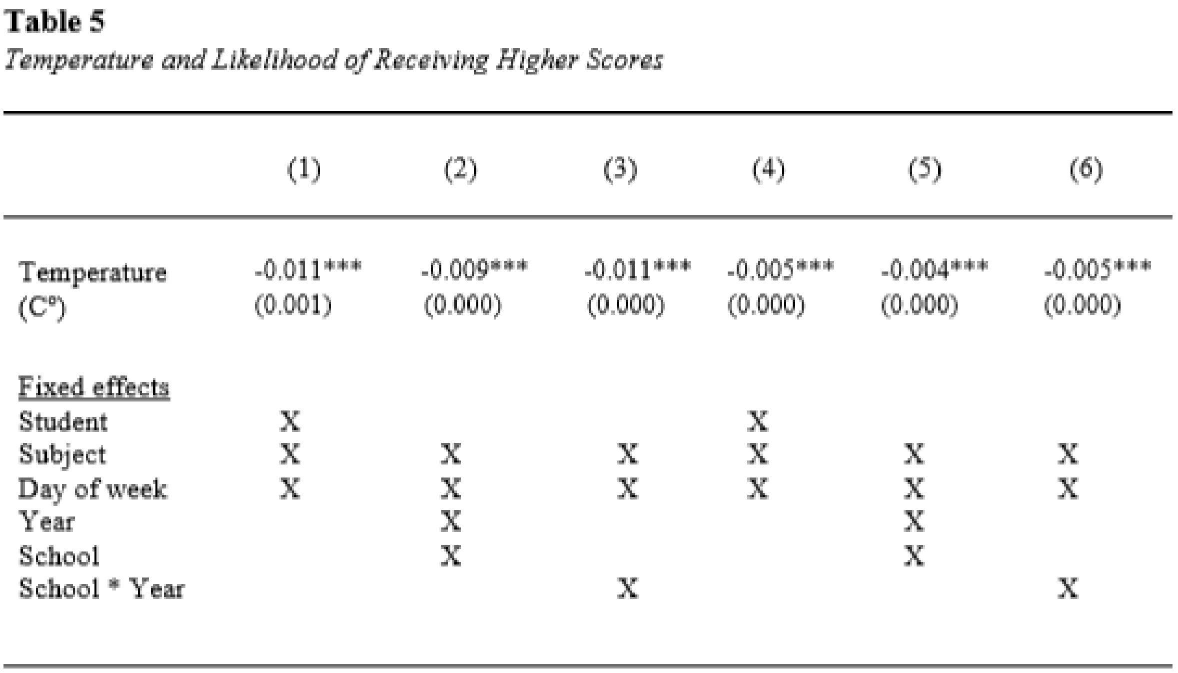


**Table 5**

*Temperature and Likelihood of Receiving Higher Scores*

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Temperature<br>(C°) | -0.011***<br>(0.001) | -0.009***<br>(0.000) | -0.011***<br>(0.000) | -0.005***<br>(0.000) | -0.004***<br>(0.000) | -0.005***<br>(0.000) |
| <u>Fixed effects</u> | | | | | | |
| Student | X | | | X | | |
| Subject | X | X | X | X | X | X |
| Day of week | X | X | X | X | X | X |
| Year | | X | | | X | |
| School | | X | | | X | |
| School * Year | | | X | | | X |

Columns 1-3 show the effect temperature on the likelihood of receiving a B on a given exam, while columns 4-6 represent the effect temperature on the likelihood of receiving an A. Given the coefficient in column 1, a one-degree Celsius increase in temperature would lead to a 4.4 percentage point decline, relative to a mean likelihood of 25.2 percent. Therefore, taking an exam when outdoor temperature is 33°C results in a 39.6 percent lower chance of getting an B on an exam relative to a 24°C day. This is much higher than the effect of 12.6 percent seen for the likelihood of simply receiving a passing score.

Next, I evaluate the effect of temperature on the likelihood of receiving an A. Given the coefficient in column 4, a one-degree Celsius increase in temperature amounts to an 8.6 percentage point decline, relative to a mean likelihood of 5.8 percent. This would result in a 77.6 percent lower chance of getting an A an exam when it is 33°C outside relative to a

24°C. The large impacts seen in these results seem to confirm the established theory that higher stakes mitigate the effect of temperature.

# 6 Discussion and Conclusion

This work examines the impact of heat stress on high-stakes exam performance in Tanzania. Using climate reanalysis data and test results from the important Primary Leaving School Exam, I find that hotter temperatures exert a causal, meaningful impact on student achievement. To do this, I exploited quasi-random variation in test-taking conditions and employed various fixed effects models. Ultimately, I find that temperature reduces exam performance by 5.8 percent of a standard deviation per standard deviation increase in exam-time temperature. This means that taking an exam when outdoor temperature is 33°C results in 14.4 percent of a standard deviation lower exam performance relative to a day at 24°C. For the average Tanzanian student, this results in a 12.6 percent lower chance of passing a given subject.

As pointed out by Park (2022), these results indicate that standardized exams taken across different climates may not be on a level playing field. For example, students in developing countries experiencing increased episodes of heat stress and a lack of cooling infrastructure may be further disadvantaged on key exams which are compared to others around the world in more optimal conditions.  Another issue, brought forth by Graff Zivin et al. (2018) is that cognitive performance forms the notion of rationality which underpins decision making within economic literature. Therefore, the commonality of temperatures impacting performance across the globe could have significant welfare impacts. As such, there may be implications for the optimal timing of key health and financial decisions. Additionally, the examination of timing and temperature may be valuable to incorporate into to future randomized control trials in development economics.

Ultimately, the task of implementing solutions to combat the effect of high temperatures is more complex in the developing world. For instance, in reference to the US, Park (2022) asserts that “providing uniformly climate-controlled test centers represents a relatively low-cost means of further reducing longstanding achievement gaps.” Unfortunately, this is not likely to be as feasible for developing countries such as Tanzania in the short-run, meaning that achievement gaps across nations is likely to persist. Furthermore, there are a number of environmental concerns regarding the mass take-up of air conditioning across the developing world. For example, the electricity demand for air conditioning is set to increase 10-fold by 2040 across Africa (Dunn 2020). Given the energy-intensive nature of AC, increased usage could worsen global greenhouse gas emissions. At the same time, its use has proven effective in mitigating the devasting impacts of heat, not only in the context of education. According to projections, effective policy and technology is actually capable of doubling the efficiency of air conditioning units and reducing their energy demand by 45 percent by 2050 (International Energy Agency 2018). Therefore, if leveraged properly, air conditioning in schools could potentially facilitate equity, economic growth, and the transition to clean energy within developing countries (Energy for Growth 2022).

In the meantime, another option would be for administrators of exams to consider environmental factors and either schedule assessments for cooler months or in the evenings. Park (2022) points out that a solution for international exams may be adjusting the scaling of scores based on geography, although this would not suffice if test-taking conditions vary within a given region. Ultimately, further research will need to explore effective adaptation techniques which are both inexpensive and accessible to schools within different developing countries.

# Bibliography

“Air-Conditioning Should Be a Human Right in the Climate Crisis - Energy For Growth.” n.d. Accessed December 12, 2022. https://www.energyforgrowth.org/op-ed/air-conditioning-should-be-a-human-right-in-the-climate-crisis/.

Alberto, Ivan Carlo, Yang Jiao, and Xiaohan Zhang. "Too hot or too cold to study? The effect of temperature on student time allocation." Economics of Education Review 84 (2021): 102152.

Auffhammer, Maximilian, Solomon M. Hsiang, Wolfram Schlenker, and Adam Sobel. "Using weather data and climate model output in economic analyses of climate change." *Review of Environmental Economics and Policy* (2013).

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S. Shapiro. "Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century." *Journal of Political Economy* 124, no. 1 (2016): 105-159.

Barro, Robert J. "Education and economic growth." *The contribution of human and social capital to sustained economic growth and well-being* 79 (2001): 13-41.

Bowler, K., and R. Tirri. "The temperature characteristics of synaptic membrane ATPases from immature and adult rat brain." *Journal of neurochemistry* 23, no. 3 (1974): 611-613.

Burke, Marshall B., Edward Miguel, Shanker Satyanath, John A. Dykema, and David B. Lobell. "Warming increases the risk of civil war in Africa." *Proceedings of the national Academy of sciences* 106, no. 49 (2009): 20670-20674.

Cho, Hyunkuk. "The effects of summer heat on academic achievement: a cohort analysis." *Journal of Environmental Economics and Management* 83 (2017): 185-196.

Deboer, Tom. "Brain temperature dependent changes in the electroencephalogram power spectrum of humans and animals." *Journal of sleep research* 7, no. 4 (1998): 254-262.

Dee, Dick P., S. Mㅒ Uppala, Adrian J. Simmons, Paul Berrisford, Paul Poli, Shinya Kobayashi, U. Andrae et al. "The ERA-Interim reanalysis: Configuration and performance of the data assimilation system." *Quarterly Journal of the royal meteorological society* 137, no. 656 (2011): 553-597.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. "Temperature shocks and economic growth: Evidence from the last half century." *American Economic Journal: Macroeconomics* 4, no. 3 (2012): 66-95.

Deschênes, Olivier, and Michael Greenstone. "Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US." *American Economic Journal: Applied Economics* 3, no. 4 (2011): 152-85.

Deschênes, Olivier, and Michael Greenstone. "The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather." *American economic review* 97, no. 1 (2007): 354-385.

Dunn, Katherine. “The great African air conditioning boom is about to begin—and it could strain the planet.” Fortune. March 19, 2020. https://fortune.com/2020/03/19/africa-air-conditioning-boom/ (accessed December 5, 2022).

Garg, Teevrat, Maulik Jagnani, and Vis Taraz. "Temperature and human capital in India." *Journal of the Association of Environmental and Resource Economists* 7, no. 6 (2020): 1113-1150.

"GDP per Capita (Current US$) - Tanzania | Data." n.d. Accessed December 12, 2022. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=TZ.

"GIZ Climate Risk Profile Tanzania." n.d. Accessed December 12, 2022. https://www.adaptationcommunity.net/wp-content/uploads/2021/02/GIZ_Climate-Risk-Profile-Tanzania_EN_final.pdf.

Graff Zivin, Joshua, Solomon M. Hsiang, and Matthew Neidell. "Temperature and human capital in the short and long run." Journal of the Association of Environmental and Resource Economists 5, no. 1 (2018): 77-105.

Graff Zivin, Joshua, and Matthew Neidell. "Temperature and the allocation of time: Implications for climate change." *Journal of Labor Economics* 32, no. 1 (2014): 1-26.

Hocking, Chris, Richard B. Silberstein, Wai Man Lau, Con Stough, and Warren Roberts. "Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing." *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* 128, no. 4 (2001): 719-734.

Hsiang, Solomon M. "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America." *Proceedings of the National Academy of sciences* 107, no. 35 (2010): 15367-15372.

Hyman, Joshua. "ACT for all: The effect of mandatory college entrance exams on postsecondary attainment and choice." *Education Finance and Policy* 12, no. 3 (2017): 281-311.

International Energy Agency. *The Future of Cooling: Opportunities for energy-efficient air conditioning*. OECD Publishing, 2018.

Kiyatkin, Eugene A. "Brain temperature fluctuations during physiological and pathological conditions." European journal of applied physiology 101, no. 1 (2007): 3-17.

Melo, Ana Paula, and Mizuhiro Suzuki. "Temperature, Effort, and Achievement." (2022).

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw. "The impact of global warming on agriculture: a Ricardian analysis." The American economic review (1994): 753-771.

Nations, United. n.d. “Country Insights.” Human Development Reports. United Nations. Accessed December 12, 2022. https://hdr.undp.org/data-center/country-insights.

Nelson, Richard R., and Edmund S. Phelps. "Investment in humans, technological diffusion, and economic growth." The American economic review 56, no. 1/2 (1966): 69-75.

Park, R. Jisung, A. Patrick Behrer, and Joshua Goodman. "Learning is inhibited by heat exposure, both internationally and within the United States." Nature human behaviour 5, no. 1 (2021): 19-27.

Park, R. Jisung. "Hot Temperature and High-Stakes Performance." *Journal of Human Resources* 57, no. 2 (2022): 400-434.

Ranson, Matthew. "Crime, weather, and climate change." *Journal of environmental economics and management* 67, no. 3 (2014): 274-302.

“Poverty Headcount Ratio at $2.15 a Day (2017 PPP) (% of Population) - Tanzania | Data.” n.d. Accessed December 12, 2022. https://data.worldbank.org/indicator/SI.POV.DDAY?locations=TZ.

Roach, Travis, and Jacob Whitney. "Heat and learning in elementary and middle school." Education Economics 30, no. 1 (2022): 29-46.

Romer, Paul M. "Increasing returns and long-run growth." Journal of political economy 94, no. 5 (1986): 1002-1037.

Sandholtz, Wayne. " Secondary school access raises primary school achievement in Tanzania." (2022).

Schlenker, Wolfram, and David B. Lobell. "Robust negative impacts of climate change on African agriculture." *Environmental Research Letters* 5, no. 1 (2010): 014010.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher. "The impact of global warming on US agriculture: an econometric analysis of optimal growing conditions." *Review of Economics and statistics* 88, no. 1 (2006): 113-125.

“School Enrollment, Primary (% Net) - Tanzania | Data.” n.d. Accessed December 12, 2022. https://data.worldbank.org/indicator/SE.PRM.NENR?locations=TZ.

“School Enrollment, Secondary (% Net) - Tanzania | Data.” n.d. Accessed December 12, 2022. https://data.worldbank.org/indicator/SE.SEC.NENR?locations=TZ.

Seppanen, Olli, William J. Fisk, and Q. H. Lei. "Effect of temperature on task performance in office environment." (2006).

“Summary for Policymakers.” In Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

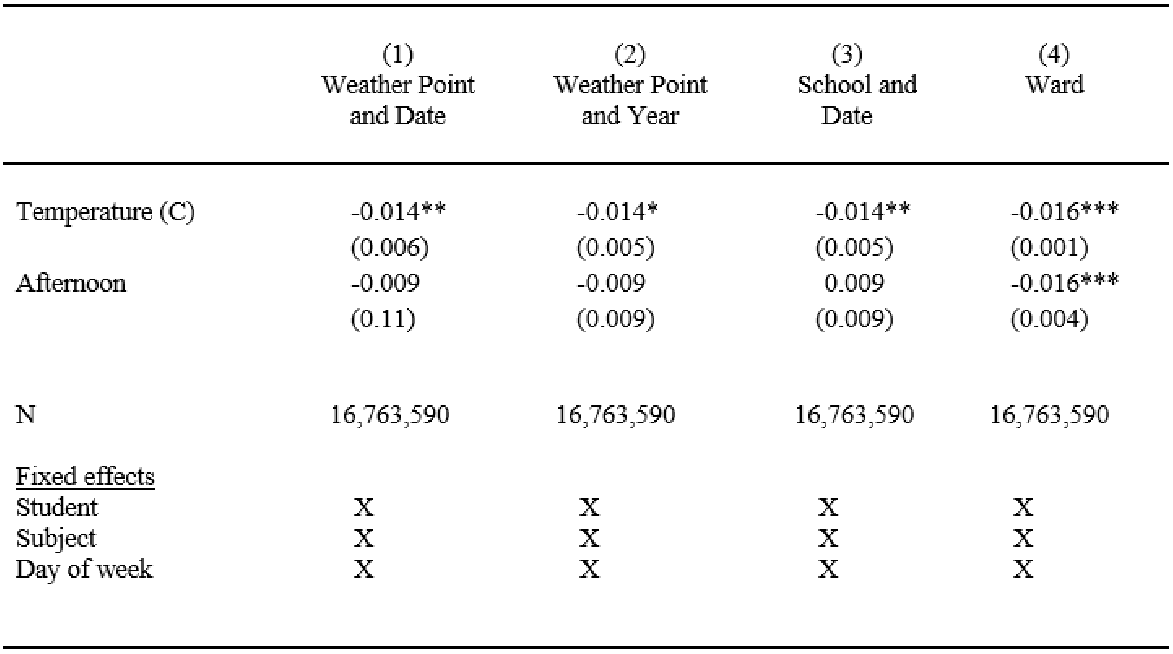
Vasmatzidis, Ioannis, Robert E. Schlegel, and Peter A. Hancock. "An investigation of heat stress effects on time-sharing performance." *Ergonomics* 45, no. 3 (2002): 218-239.

Yablonskiy, Dmitriy A., Joseph JH Ackerman, and Marcus E. Raichle. "Coupling between changes in human brain temperature and oxidative metabolism during prolonged visual stimulation." *Proceedings of the National Academy of Sciences* 97, no. 13 (2000): 7603-7608.

Zivin, Joshua Graff, Yingquan Song, Qu Tang, and Peng Zhang. "Temperature and high-stakes cognitive performance: Evidence from the national college entrance examination in China." *Journal of Environmental Economics and Management* 104 (2020): 102365.

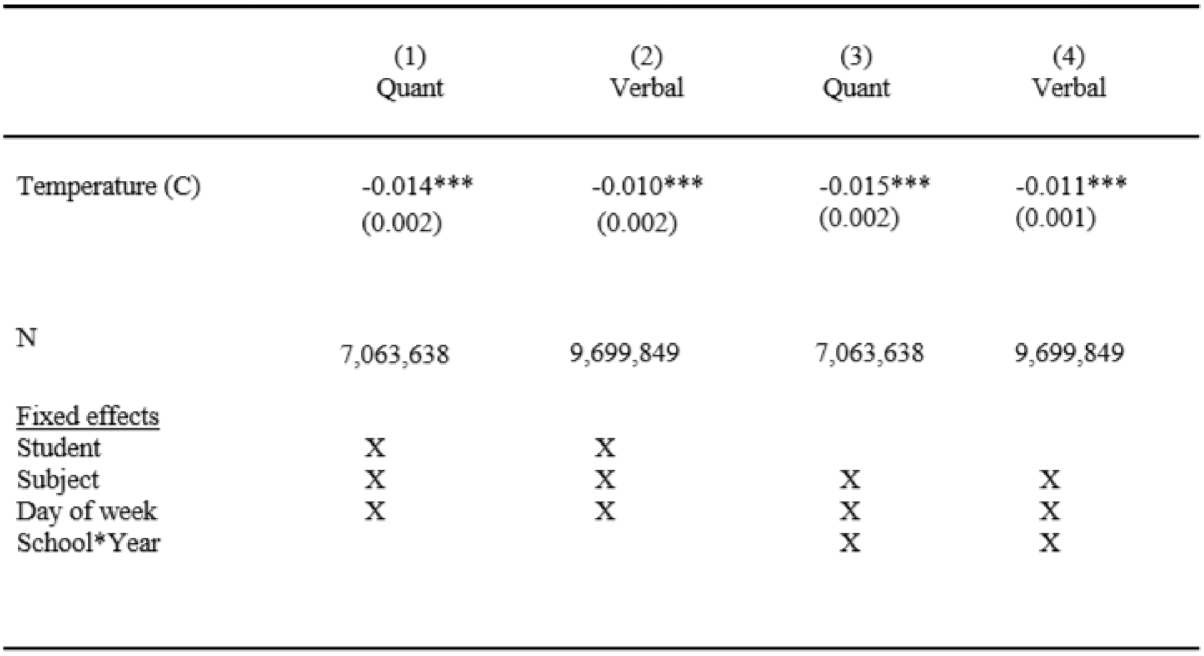
# Appendix

Table A.1: Temperature and High Stakes Exam Performance: Robustness to Alternative Clustering



| | (1) Weather Point and Date | (2) Weather Point and Year | (3) School and Date | (4) Ward |
|---|---|---|---|---|
| Temperature (C) | -0.014** | -0.014* | -0.014** | -0.016*** |
| | (0.006) | (0.005) | (0.005) | (0.001) |
| Afternoon | -0.009 | -0.009 | 0.009 | -0.016*** |
| | (0.11) | (0.009) | (0.009) | (0.004) |
| N | 16,763,590 | 16,763,590 | 16,763,590 | 16,763,590 |
| Fixed effects | | | | |
| Student | X | X | X | X |
| Subject | X | X | X | X |
| Day of week | X | X | X | X |

Table A.2: Temperature and High Stakes Exam Performance: Subject-Specific Impacts



| | (1) Quant | (2) Verbal | (3) Quant | (4) Verbal |
|---|---|---|---|---|
| Temperature (C) | -0.014*** | -0.010*** | -0.015*** | -0.011*** |
| | (0.002) | (0.002) | (0.002) | (0.001) |
| N | 7,063,638 | 9,699,849 | 7,063,638 | 9,699,849 |
| Fixed effects | | | | |
| Student | X | X | | |
| Subject | X | X | X | X |
| Day of week | X | X | X | X |
| School*Year | | | X | X |