I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16249

The Fast and the Studious? Ramadan Observance and Student Performance

Kyra Hanemaaijer
Olivier Marie
Marco Musumeci

JUNE 2023

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 16249

The Fast and the Studious? Ramadan Observance and Student Performance

Kyra Hanemaaijer
EUR and TI

Olivier Marie
EUR & ROA, TI, CEPR, IZA and CESIfo

Marco Musumeci
EUR and TI

JUNE 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Fast and the Studious? Ramadan Observance and Student Performance*

What are the consequences of religious obligations conflicting with civic duties? We investigate this question by evaluating changes in the performance of practicing Muslim students when end-of-secondary-school exams and Ramadan overlapped in the Netherlands. Using administrative data on exam takers and a machine learning model to individually predict fasting probability, we estimate that the grades and pass rate of compliers dropped significantly. This negative impact was especially strong for low achievers and those from religiously segregated schools. Investigating mechanisms, we find evidence that not being able to sleep in the morning before an afternoon exam was particularly detrimental to performance.

JEL Classification: I2, I24, Z12, J15

Keywords: religion, productivity, Ramadan, education, The Netherlands

Corresponding author:
Olivier Marie
Erasmus University Rotterdam
Burgemeester Oudlaan 50
3062 PA Rotterdam
The Netherlands
E-mail: marie@ese.eur.nl

* The authors are grateful for financial support from the Netherlands Organization for Scientific Research (NWO 016.Vidi.185.049). We thank Robin Verbeek and Elena Castellaro for providing excellent research assistance on this project. We are grateful to Hessel Oosterbeek, Gabriela Smarrelli, Renske Stans, Zihan Ye and Dinand Webbink for taking the time to give detailed comments on an earlier version of the paper. We also wish to thank participants at multiple seminars, workshops, and conferences for their valuable suggestions.

Do you think there are possible effects of taking tests and exams during periods of religious fasting? If so, do these effects form a possible obstacle to the progress of pupils and/or students in education?
Tunahan Kuzu, Member of the Dutch Parliament

It is the individual responsibility of a student to prepare as well as possible for tests and exams. It may be that students find religious obligations important, even during periods when tests are held. I leave that decision to the individual.
Sander Dekker, Secretary of State at the Dutch Ministry of Education.

Question time at the Dutch Parliament on the 6^{th} of June 2016

1 Introduction

Individuals' religious obligations and civic duties are sometimes misaligned. This tension may be particularly pertinent for religious minorities, including the quarter of the worldwide population of Muslims who reside outside of Muslim-majority nations (Lugo et al., 2011). In the Muslim calendar, the ninth month, Ramadan, is a month of fasting, prayer and faithful intention. During Ramadan, all healthy post-pubertal Muslims must refrain from eating and drinking between sunrise and sunset for up to 30 days.[1] In countries where Muslims are a minority group, there will be less accommodation in public life for the observance of this holy month. While some may be able to adapt their schedule to minimize disruptions (e.g. by changing working hours or taking holidays), for many this is not an option. This may be especially problematic, and potentially have long term consequences, when Ramadan falls during a period when an individual's relative productivity is tested: i.e. during high-stakes exams.

The religion-productivity relationship our paper explores closely relates to two recent studies which have provided compelling macro-level evidence of negative impacts of religious practices on economic growth. Campante and Yanagizawa-Drott (2015) exploit differences in the length of days during Ramadan in Muslim countries and Montero and Yang (2022) exploit the timing of Catholic saints festivals in Mexico. In this paper, we focus on the effect of religious practices on productivity at the micro-level, in particular when the macro-context of our sam-

[1] Some exceptions to this obligation exist and are mainly related to health conditions, pregnancy, menstruation, old-age, or traveling.

ple is not affected in the same way. This matters because it could make a religious minority group appear relatively less productive than the majority group. Exacerbating already existing (large) differences in outcomes as a consequence, the impact of religious practices on the relative productivity of minorities is an important angle to consider when investigating the causes of educational inequalities across groups (Blanden et al., 2022).

The Islamic calendar is shorter than the Gregorian calendar, so Ramadan starts 10-12 days earlier every year. In recent years the Muslim holy month has fallen in May and June, the period when high school graduation examinations take place in several Western countries. In these cases, there has been no accommodation for Muslim students who are fasting.[2] As a consequence, Muslim students - who are already part of an under-privileged group - are required to take high-stakes exams while observing Ramadan. The potential negative repercussions on their academic performance have been highlighted in the media of most Western countries.[3] Despite the media attention this generated and the growing share of Muslims living in affected countries,[4] the causal evidence of the impact of Ramadan exposure on such academic tests is very limited. We address this knowledge gap with a close examination of the impact of Ramadan observance on students' performance during high-stakes exams in the Netherlands.

In the medical literature, short-term fasting and dehydration are associated with cognitive impairment (Adolphus et al., 2016; Wittbrodt and Millard-Stafford, 2018). However, fasting is not the only factor that could impact Muslim students' performance during Ramadan. When Ramadan falls at times of the year when there are more hours of daylight, and the morning meal (*Suhoor*) and late night meal (*Iftar*) are further apart, this lengthens the fast and reduces hours of sleep. Consequently, Ramadan observance has been associated with important disruptions in sleep patterns (Bahammam, 2003, 2006, 2013; Margolis and Reed, 2004) potentially leading to sleep deprivation, which is itself strongly linked to loss of cognitive functions (Curcio et al., 2006; Csipo et al., 2021). Prima facie one would expect the observance of Ramadan to negatively impact performance. Surprisingly, the existing small scale medical experiments, looking

[2]In some Muslim-majority countries, schools change their schedule to take Ramadan into account. For example, in Morocco, classes are often made shorter, while in Saudi Arabia, schools are closed for (most of) the duration of Ramadan (this was not the case in 2022).

[3]There has been attention on this topic in several countries, such as the UK ("Popular exams in UK to be rescheduled to avoid Ramadan", The Guardian), the USA ("Testing and Fasting", Inside Higher ED), France ("*Jeûner pendant le bac : le défi des lycéens qui font le Ramadan*", France24), Germany ("Hungry Students? Postponed Exams? Ramadan In German Schools", Worldcrunch), Italy ("*La Maturità incrocia il Ramadan*", Corriere della Sera), and the Netherlands ("*Vasten en eindexamens, gaat dat wel samen?*", NOS).

[4]Muslims represent today around 5% of all the Europeans and by 2050 this number is expected to grow to 11% (Lugo et al., 2011)

at the effect of Ramadan on cognitive functions in a lab setting, provide mixed evidence. Tian et al. (2011) find that psycho-motor function, vigilance, verbal learning and memory are impaired in male athletes complying with Ramadan obligations, while Ghayour Najafabadi et al. (2015) and Yasin et al. (2013) find no adverse effect on cognitive functions.

The few studies which, as in our setting, have used changes in student performance in real-world exams to explore the impact of Ramadan observance also come to inconsistent conclusions. Oosterbeek and van der Klaauw (2013) are the first to investigate how Ramadan exposure affects academic performance. They find that each week of overlap between a university microeconomics class and Ramadan is associated with a decrease in the exam grade of students with Muslim sounding names of almost 10% of a standard deviation. Hornung et al. (2023) and Kökkizil (2022) use cross-country data on TIMMS and PISA tests to compare performance at tests taken up to one year after the end of Ramadan of students with parents from Muslim majority countries. Hornung et al. (2023) find positive effects, which they interpret as Ramadan observance having a positive impact on social capital formation of young Muslim students. On the other hand, Kökkizil (2022) finds that Ramadan makes the existing gender norms and stereotypes more salient: in the three months after the end of Ramadan, the academic performance of female Muslim students in STEM subjects significantly worsens compared to their male peers. Finally, Nuryakin et al. (2022) use a difference in differences model to investigate the effect of Ramadan fasting on students' performances in the final exams at the University of Indonesia. They find that the overlap between Ramadan and the exam period has no effect on the students' test scores, but they cannot test for parallel trends in exam results between Muslim and non-Muslim students. Our study builds on this small literature, and improves on it in a number of ways.[5] We are the first to measure students' performance *before* and *during* Ramadan for a large group of students facing a high-stakes exam. This allows us to estimate the causal effect of Ramadan exposure on students' test scores. Second, we explore which students are likely to comply with Ramadan obligations, by developing a machine learning algorithm that assigns compliance probabilities to students, rather than relying on narrow identifiers of nationality, country of origin or name. Third, we use rich administrative data to obtain precise

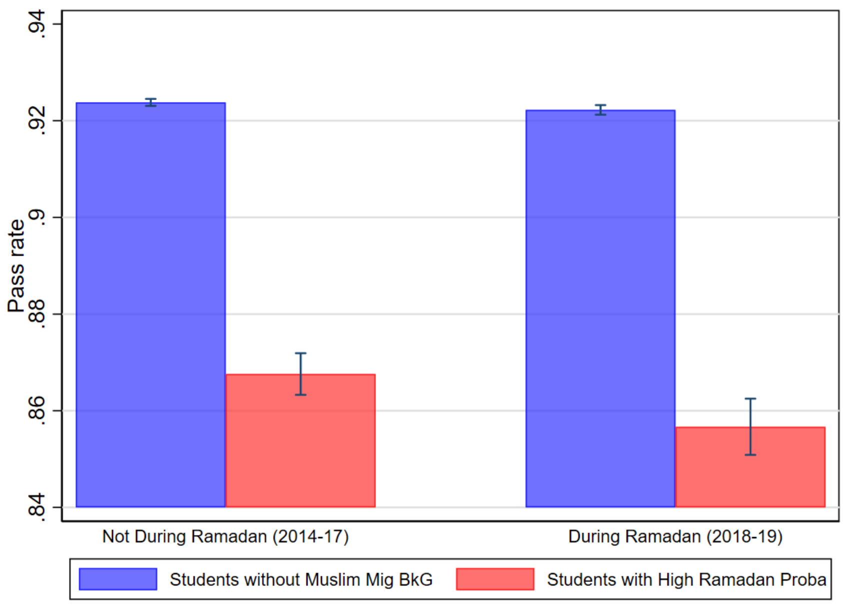
[5]There is a much more developed literature that has exploited the timing of Ramadan on other outcomes. Almond and Mazumder (2011), Van Ewijk (2011), Almond et al. (2015) have shown that prenatal exposure to Ramadan results in worse health outcomes at birth and later in life in various contexts. Schofield (2020) considers the impact of Ramadan on agricultural output in India. Colussi et al. (2021) look at voting behaviour and violent attacks against Muslim communities in Germany, and Bertoli et al. (2022) find that workplace accidents increase during Ramadan in Spain.

average estimates, and to look into heterogeneities and peer effects. Fourth, we attempt to distinguish between the impact of the eating and sleeping disruptions of Ramadan by considering the timing of exams. This allows us to make more specific policy recommendations on how to attenuate the negative effect of Ramadan on Muslim students that only require a change to the exam schedule rather than avoidance of the entire Ramadan period. Finally, we use data on a standardized end-of-primary-school test at the national level to control for changes in ability. We show that this is especially relevant for students with a migration background, as their performance improves relative to native students over time.

Our focus is on the impact of Ramadan exposure on (compliant) students' performance at the secondary graduation exams in the Netherlands. These exams comprise several written tests covering different subjects, that are the same at the national level. High school students must pass these exams to graduate and to be eligible for tertiary education. We focus on the final exams that took place from 2014 to 2019, with the Muslim holy month overlapping with the exams in both 2018 and 2019. In all previous years included in our analysis (2014 - 2017), Ramadan fell after the final exams and thus barely affected the instruction of previous cohorts, as this was mostly during the summer holidays. This Ramadan-exams overlap is the first difference we exploit in our difference-in-differences identification approach. The second difference stems from the variation in Ramadan observance of students. This strongly depends on correctly identifying the students that are most likely to comply with Ramadan obligations. Based on Dutch survey data on religiosity and fasting behavior, we show that simply considering as treated all students with a migration background from a Muslim majority country would missclassify many students. We instead assign treatment based on a machine learning (ML) model that predicts Ramadan compliance probability based on a large set of observable characteristics for all individuals in our sample. This reveals a large difference in compliance between and within the two main Muslim minorities in the Netherlands (individuals with a Turkish and Moroccan migration background). While most students with a Moroccan background are very likely to comply, those with a Turkish background have a much wider distribution in their observance of the Ramadan fast. This novel and more precise classification into treatment is the second difference our identification strategy relies on.

The most basic presentation of this difference-in-differences approach is shown in Figure 1 which reports average pass rates at final exams that did and did not overlap with Ramadan for treated and non-treated students respectively. Since most minority students in the Netherlands are from a socioeconomically disadvantaged population group, their pass rates are always much lower on average. Still, the existing large achievement gap between Muslim and non-Muslim students was strongly and significantly exacerbated - widening by almost 17% - when the high school graduation exams took place during Ramadan.

Figure 1: Average pass rate at final exams by Ramadan compliance probability



Notes: The bar graph shows the average pass rate and 95% confidence interval of students with a high Ramadan-compliance probability (in red) and students with no Muslim-majority country migration background (in blue). The graph displays the averages for the years in which Ramadan did not overlap with the final exams (2014-2017) on the left, and when it did overlap with the final exams (2018-2019) on the right.

These findings are confirmed using statistical models which interact treatment status of students with Ramadan exam years, while controlling for a large set of individual, family, school, neighbourhood, and time characteristics. We estimate that taking a high-stakes exam during Ramadan decreases the grade of students most likely to fast by 5.8% of a standard deviation (SD). This in turn leads to an increase of 8.3 percent in the probability of these students failing to graduate from high school.

We corroborate the validity of these findings with robustness checks to confirm that it is Ramadan observance that drives our results. First we check that these results are not driven by a change in students' academic ability. We treat the average grade obtained in all tests and homework performed in the last years of high school before the final exam as the dependent variable.[6] These 'continuous' exams should not be affected by Ramadan, as in the years that we consider in our analysis they do not overlap with it, but are a good proxy for the academic

[6]The change in academic ability of students is also controlled for by including the standardized end of primary school test score in all regressions.

ability of students. We find no effect in 2018 and 2019, confirming that our findings are not driven by lower skilled students in the treatment group taking the exams in those years. Another revealing check is a placebo test which considers as treated those students that are very unlikely to observe Ramadan, because they do not have a migration background from a Muslim-majority country. We select this group based on individual characteristics that our ML model identifies as the most relevant to predict Ramadan compliance. Hence, these students are similar in almost all aspects to our 'real' treated group, but should not be affected by Ramadan. Here again we do not find any effect for placebo students, which shows that the results are unlikely to be driven by some shock affecting students with specific characteristics (e.g. disadvantaged background) in the years of the Ramadan-exams overlap. Finally, we check the robustness of our results by including: i) sibling fixed effects, and ii) family fixed effects, which includes cousins as well. In this way we can control for unobserved characteristics that are shared among siblings and cousins. These alternative specifications produce results that are quantitatively similar to our main results. We take this as confirmation that there is a causal relationship between the observed drop in performance and final exams coinciding with the Muslim holy month.

We perform a large number of investigative heterogeneity analyses to uncover who is most affected by Ramadan. There are no significant gender differences and students of all ability and income levels show a relatively similar drop in grades. However, the drop in performance is especially detrimental for the probability that students in the lowest quartile of pre-final exam performance will graduate. School-level heterogeneity analyses reveal that students in the schools with the most segregation, in terms of parental income or the proportion of students with a high Ramadan compliance probability, are driving our results. Those in schools in the top quartile of the 'complier segregation' saw their grades drop by 7% of a standard deviation and were 12.1% more likely not to graduate. Because of this strong segregation effect we explore the possibility of peer-effects in the likelihood of Ramadan observance. We exploit within-school across-cohort differences in the proportion of students with high Ramadan compliance probabilities. We find that there is a strong additional impact on treated students' test scores, when students have more treated peers in their cohort.

Considering the potential impact of the outcome of final high school exams, we also explore whether sitting this high-stakes exam during Ramadan has long term consequences. We find that students with a high probability of complying with Ramadan obligations are 12.4% more likely to repeat the final year of high school. Moreover, we observe that students older than 18, those allowed to leave school, are 22.5% more likely to drop out from school without a diploma if the final exams occur during Ramadan.

Finally, we attempt to disentangle the impacts of fasting and disruptions to sleep patterns that accompany Ramadan observance. For this, we first consider exam results depending on the type of exam taken, as lack of sleep does not affect all types of cognitive performance similarly (Lim and Dinges, 2010). Short-term memory based tests are somewhat more affected, suggesting sleep deprivation matters. We then consider how far into Ramadan each exam is taken. We hypothesize that fasting would mostly affect exams in the first week of Ramadan, while the effect of sleep deprivation would be cumulative and have the greatest impact in later weeks. However, we do not find evidence of a differential effect of when in Ramadan the exam is taken. We then investigate whether there is a difference between morning and afternoon exams. We find that afternoon results are especially negatively affected, but only when the student already took an exam in the morning. We believe this is because students counter the effects of sleep deprivation by taking (long) naps after the morning meal when they have no exams in the morning. These effects could therefore be mitigated by only scheduling tests in the afternoon, and not having exams that morning, in cases where it is not feasible to move performance evaluations away from the Ramadan period altogether. This will at least reduce the large negative impact we uncover on Muslim-minority students whose religious observances conflict with their civic duties.

The rest of the paper is structured as follows. In the next section we describe the background of Ramadan and Muslims in the Netherlands, the Dutch secondary school system and the overlap between the central exams and Ramadan. In section 3 we describe the data we use. In section 4 we discuss the identification of the ‘treated’ group, our difference-in-differences identification strategy and the machine learning model we use to predict Ramadan compliance. In section 5 we describe our main results and in Section 6 we explore potential heterogeneity by student and school characteristics and by subject. In section 7 we look into underlying mechanisms and propose policy implications. In section 8 we conclude.

2 Background

2.1 Ramadan

Ramadan, which is the ninth month of the Islamic calendar, imposes a fast that is considered one of the five pillars of Islam. It requires all healthy Muslim individuals who have reached puberty to abstain from eating and drinking between sunrise and sunset. There are some exceptions, including health conditions, pregnancy, menstruation, age, or travel. Ramadan lasts for 29 or 30 days and its starting date is determined by the Islamic calendar. This calendar is shorter than the Gregorian one by approximately 11 days, so that Ramadan moves earlier relative to the Gregorian calendar every year by 11 days. As Ramadan moves over time, the daily duration of fasting, which depends on sunrise and sunset, varies by year. In the Netherlands, the sun rises at 5.20 AM and sets at 10 PM on the longest day of the year (21 June), while it rises at 8.50 AM and sets at 4.30 PM on the shortest day (21 December). As a result, Ramadan has a very different impact on daily life depending on when during the year it occurs. On the shortest day of the year, Muslims only have to fast for 7 hours and 40 minutes, and can take their meals at their normal times. If Ramadan takes place during the summer, observers need to have breakfast before 5.20 AM and dinner after 10 PM, and hence will have to abstain from eating and drinking for 16 hours and 40 minutes.

In some Muslim-majority countries, working hours are reduced during Ramadan and schools change their schedule to accommodate students' needs. For example, classes can be shorter, as happens in Morocco, or schools can close completely during Ramadan, which is the case in Saudi Arabia. In most European countries, employers and schools do not make accommodations for Muslims who observe the fast. However, Muslims represent today around 5% of all Europeans and by 2050 this number is expected to grow to 11% (Lugo et al., 2011), which will lead to a rising demand for accommodations for Ramadan in non-Muslim countries.

2.2 Muslims in the Netherlands

According to estimates from Statistics Netherlands, in 2021 around 5% of the population aged 15 or above in the Netherlands was Muslim.[7] There are more than 450 mosques throughout the country.[8] This is not surprising considering that the two largest groups of migrants in the

[7]Data come from Statistics Netherlands.

[8]Data was collected from Moskeewijzer.nl in 2021.

Netherlands come from Muslim majority countries: Turkey and Morocco. First, second and third-generation immigrants from these two groups represent respectively 2.6% and 2.5% of the population.[9] Individuals with Moroccan or Turkish migration backgrounds tend to concentrate in large cities. As of 1 January 2022, around 45% of the population with a Moroccan migration background and 35% of the population with a Turkish migration background was living in one of the four large cities in the Netherlands (Amsterdam, Rotterdam, Utrecht, The Hague).[10] This is shown in the map in Figure 2 that illustrates the percentage of individuals with a Moroccan or Turkish migration background at municipality level.[11] The darker areas correspond to large cities in the Netherlands. Each dot indicates a mosque where the language of the sermon is Moroccan Arabic or Turkish.

The age profile of the population groups with a Moroccan or Turkish migration background is quite different from the population group without a migration background. Individuals with a Moroccan or Turkish migration background are relatively younger (on average in their mid-thirties) than those without a migration background (on average around fifty years old).[12] Among the high school population in 2017, these two ethnic groups represent a larger share of students than their share in the general population. They jointly represent 7.1% of students, with the students with a Moroccan migration background accounting for 3.6% and those with a Turkish migration background accounting for 3.5%. The majority of the students with a Moroccan or Turkish migration background are second generation migrants, meaning that respectively 95.1% and 92.7% have at least one parent who was born in Morocco or Turkey. Most students have parents that migrated to the Netherlands in the end of the 1980s, when a large wave of labour migrants was followed by subsequent migrations, as the wives and children of these labour migrants also moved to the Netherlands. Approximately 3.2% (3.6%) of the students with a Moroccan (Turkish) background migrated themselves (first generation), and 1.7% (3.7%) has at least one grandparent born in Morocco (Turkey), but both parents born in the Netherlands (third generation).

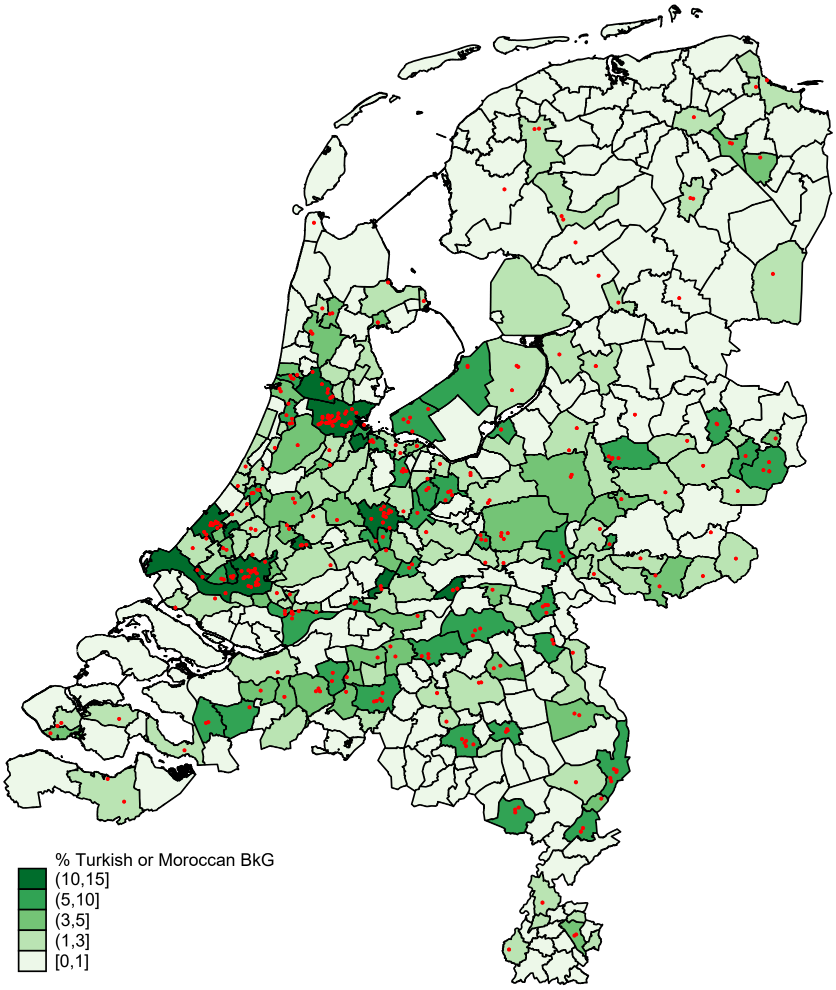
[9] See data section for an explanation on how individuals are classified as 1st, 2nd or 3rd generation immigrants. Numbers are author's calculations based on data from Statistics Netherlands on 1st and 2nd generation migrants and 3rd generation migrants.

[10] Within these cities, there is a lot of variation in the level of segregation across neighbourhoods. In certain neighbourhoods, up to 65% of the residents have a Moroccan or Turkish background.

[11] See Figure A1 for a map of individuals with a Moroccan migration background and Moroccan mosques and see Figure A2 for a map of individuals with a Turkish migration background and Turkish mosques.

[12] Data come from Statistics Netherlands.

Figure 2: The share of municipalities with a Moroccan or Turkish migration background and the location of mosques in the Netherlands



Notes: Municipality level proportion of people with a Moroccan or Turkish migration background in 2017 in the Netherlands. The locations of mosques where the language of the sermon is either Turkish or Moroccan Arabic are indicated with red dots. Source: CBS Statline, statline.cbs.nl; Moskeewijzer, https://moskeewijzer.nl/.

2.3 Secondary education in the Netherlands

Pupils in the Netherlands usually start secondary school at the age of 12.[13] They can enroll (advised by the results of a standardized national test called CITO test and their teacher) in one of the following three types (tracks) of secondary education: pre-vocational secondary education (*VMBO*), senior general secondary education (*HAVO*) or pre-university education (*VWO*). The duration of these programs is respectively four, five and six years. At the end of their third year (or second, in pre-vocational secondary education) students have to choose a certain sub-

[13]In this section we focus only on secondary education in Netherlands but in Appendix Table A1 we provide a detailed graphical overview of the entire Dutch educational system.

track (*profiel*) that corresponds to a selection of topics. In the years after the subtrack choice, students will mainly focus on the topics that correspond to this subtrack. However, certain subjects are compulsory in every subtrack, such as Dutch, English language and literature and physical education. Students are required to stay in school up to the age of 16. After the age of 16, compulsory education is replaced with a compulsory "start qualification" up to the age of 18. A start qualification is obtained when students obtain a senior general secondary education or pre-university education diploma, or a diploma in secondary vocational education (which students can attend after graduating from pre-vocational secondary education).

To qualify for the secondary school diploma, students have to pass the secondary school graduation exam. They usually take this exam aged 16 to 18, depending on the track they are in. Fifty per cent of the final grade is determined by the average of decentralized continuous school exams and 50% by the average of the centralized final exams in May of the final year.[14] Continuous school exam results are based on assignments and tests from the last two or three years of high school. These tests and assignments are different for each school because they are formulated and graded by the student's teachers. By contrast, the centralized exams are the same for everybody and are assessed by the student's teacher and by a teacher that is randomly selected from another school in the Netherlands. The number of subjects that are tested during the central exam depends on the track and on the student's choices, but it usually ranges from 5 to 8. The grades of both the central and continuous exams are on a scale from 1 to 10, with the latter being the maximum grade that a student can achieve and 5.5 being the minimum passing grade. There are some other conditions on graduation that depend on the student's track.[15]

2.4 Central Exam and Ramadan

Central exams are usually taken in May over a period of three weeks, according to a schedule determined by the government.[16] In June there is a retake session for the central exams. During

[14]Some subjects do not have a central final exam (for example physical education) and therefore the final grade is completely determined by the school interim examination.

[15]In addition to the requirement of a 5.5 as overall average grade of the central exam, students in senior general secondary education and pre-university education can only have one 5 as a final grade (which consists for 50% of the continuous exam grade and for 50% of the central exam grade) for Dutch, English and Math, and they need to have an average final grade overall of 6 or higher, no final grade lower than a 4 and a pass for physical education. In addition to the requirement of a 5.5 as overall average grade of the central exam, students in pre-vocational secondary education need a 5 or higher as a final grade for Dutch, an average final grade overall of 6 or higher, no final grade lower than a 4, a pass for physical education, art education and research project (this latter requirement only applies to students in the theoretical type of pre-vocational secondary education (*VMBO-TL*)).

[16]See Examenblad.nl for the detailed schedule of (and more information on) all exams.

this session, students can retake exams if they were severely ill during the main session in May or they had another serious personal reason not to take the exam. Moreover, in the retake session a maximum of one exam can be taken for a resit, if a student wants to improve their grade. In this case, the final grade is the student's highest grade (so either the original or resit grade). In July there is a second retake session which is only available when students could not take the exam in the original exam period and the first retake session.

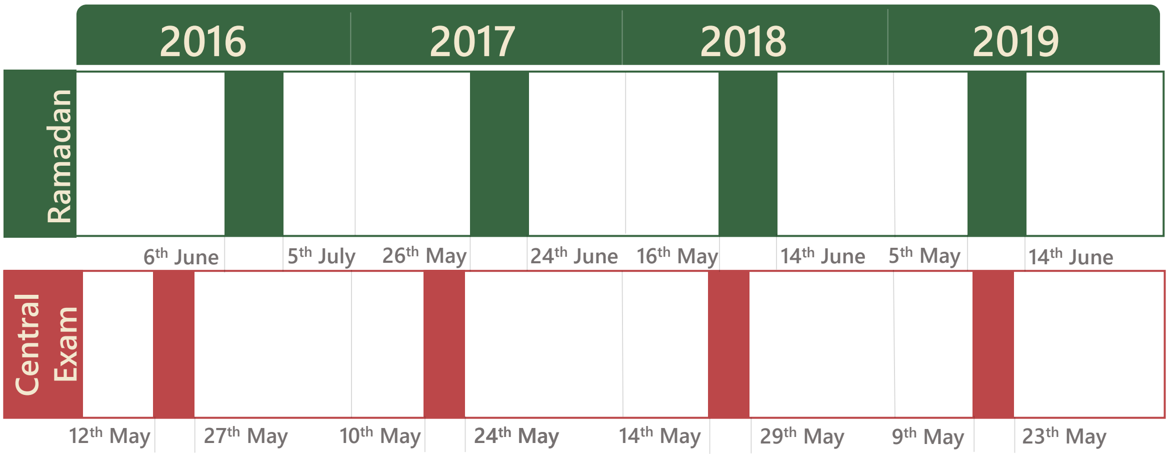
In recent years, there has been an overlap between the central exam, which is always held in May, and Ramadan, as illustrated in Figure 3. Prior to and including the year 2017, the month of Ramadan occurred after the central exams (so Ramadan did not overlap with the central exam, nor with the lead up to the central exam). However, in the years 2018 and 2019, the central exam overlapped with the ninth month of the Islamic calendar.[17] This overlap was complete in 2019 while in 2018 only partial, as the first three days of the central exam took place before the beginning of Ramadan. When Ramadan overlapped with the final exams in May, sunrise was at approximately 5.50 AM and sunset at approximately 21.30 PM, meaning that students had breakfast before 5.50 AM and dinner after 21.30 PM, and had to fast for about 16 hours per day. Morning exams started at 9 AM and afternoon exams at 13.30 PM. Hence, students did not have (much) time to sleep in after breakfast when they had a morning exam, and had abstained from drinks and food for more than 7.5 hours when they started an afternoon exam.

A member of the parliament of left-wing political party "DENK" had already requested the rescheduling of the central final exams in 2016 to avoid an overlap with Ramadan. The State Secretary of Education responded by saying that "it is practically impossible to reschedule the final exams to accommodate Ramadan. [...] It could be that students find religious obligations important, even in periods when they take tests. I leave that decision to the individual.".[18]

[17] School assessments take place at multiple moments (continuously) over the year and are scheduled by schools themselves. Continuous exams have to end before the central exams start for students in their final year who take the central exams. Students in the years before may also take continuous exams in June or July, and hence Ramadan might have partly overlapped in the years before 2019 (as it took place in May-July in 2014-2018) with one of their continuous exams. However, as the continuous exam grade that we use is constructed out of many tests and assignments over 2 to 3 years, this would mean that only a very small fraction of the continuous exam grade may have been affected.

[18] The original and complete answer in Dutch in the parliamentary document was: "Jaarlijks doen ongeveer 200.000 leerlingen het centrale examen in het vastgestelde tijdvak. Het is de individuele verantwoordelijkheid van een leerling om zich zo goed mogelijk voor te bereiden op toetsen en examens. Het kan zijn dat leerlingen religieuze verplichtingen belangrijk vinden, ook in periodes waarin toetsen worden afgenomen. Die afweging laat ik aan het individu. Islamitische leerlingen die graag deelnemen aan het vasten kunnen er overigens ook voor kiezen om - eventueel in overleg met hun geestelijke - de vastenperiode te verplaatsen. Er zijn leerlingen die dat ook doen.'

Figure 3: Graphical representation of overlap between final exams in secondary school and Ramadan.



Notes: Graphical representation of the overlap between the central exams and Ramadan. In the years before and including 2017, there is no overlap between the two, even if in 2017 they are very close. From 2018, the central exam and Ramadan occur in the same period (only partially in 2018). Source: Examenblad.nl, https://www.examenblad.nl/.

3 Data

3.1 Administrative Data

We use administrative data from different registries compiled by Statistics Netherlands.[19] We focus on secondary school students who took the final exam between 2014 and 2019 and for whom we observe all relevant characteristics[20] (841,376 observations, or 789,259 students)[21] in 549 schools. For those students, we observe their central exam grade (both including and excluding retakes) and continuous school exam grade both as an average of all courses together and at the subject level, whether they passed the exams/graduated, their track and sub-track (and the tracks they were in before, if they moved up or down, and whether they repeated a year), their result on the national end-of-primary-school (*CITO toets*) test (as a standardized measure of ability), and demographic characteristics such as birth date and migration background. We also observe neighbourhood characteristics (share with a Turkish or Moroccan migration background, share that receives welfare benefits, average household size, share under 45 years old,

[19]See Appendix Section B for a detailed description of the data set-up, sample selection process and variable definitions.

[20]The characteristic that causes the largest selection in sample is the national end of primary school test that we use as a standardized measure for ability. We can only link about 75% of the exam takers to their result on this test.

[21]We have for approximately 6.5% students multiple observations, as they do not graduate from high school in one attempt and have to repeat the final year including the final exams.

fertility), school characteristics (share with a Turkish or Moroccan migration background) and characteristics of their parents (income, unemployment, illness).

We classify individuals as having a foreign background if they are first, second or third-generation immigrants. An individual is considered a first-generation immigrant if they are born outside the Netherlands and one of their parents is born abroad as well. A second-generation immigrant is an individual who is born in the Netherlands but has at least one parent who was born abroad. A third-generation immigrant is an individual born in the Netherlands, with parents who are also born in the Netherlands, but at least one grandparent is born abroad. If a person does not belong to any of the above categories, then they are considered to have no migration background.[22]

Figure 4 shows the distribution of secondary school central exam grades (standardized at track level) in 2017 for students with the two most common migration backgrounds (Moroccan migration background in red and Turkish migration background in green) and students with no migration background or a migration background from a non-Muslim majority country (in blue). This figure shows an existing gap in the standardized average exam grade between Moroccan (Turkish) background students and the students without a Muslim-majority migration background of approximately 48% (50%) of a standard deviation. Notice that this gap is measured in 2017, when there was no overlap between Ramadan and the final exams.

Figure 4 also provides additional descriptive statistics for these three groups. Students with a Moroccan or Turkish background have on average poorer parents, live in more segregated neighborhoods and performed worse in the standardized end-of-primary-school test. Put together, the descriptive statistics presented underline how students with a Moroccan or Turkish background come from a strongly under-privileged population group.

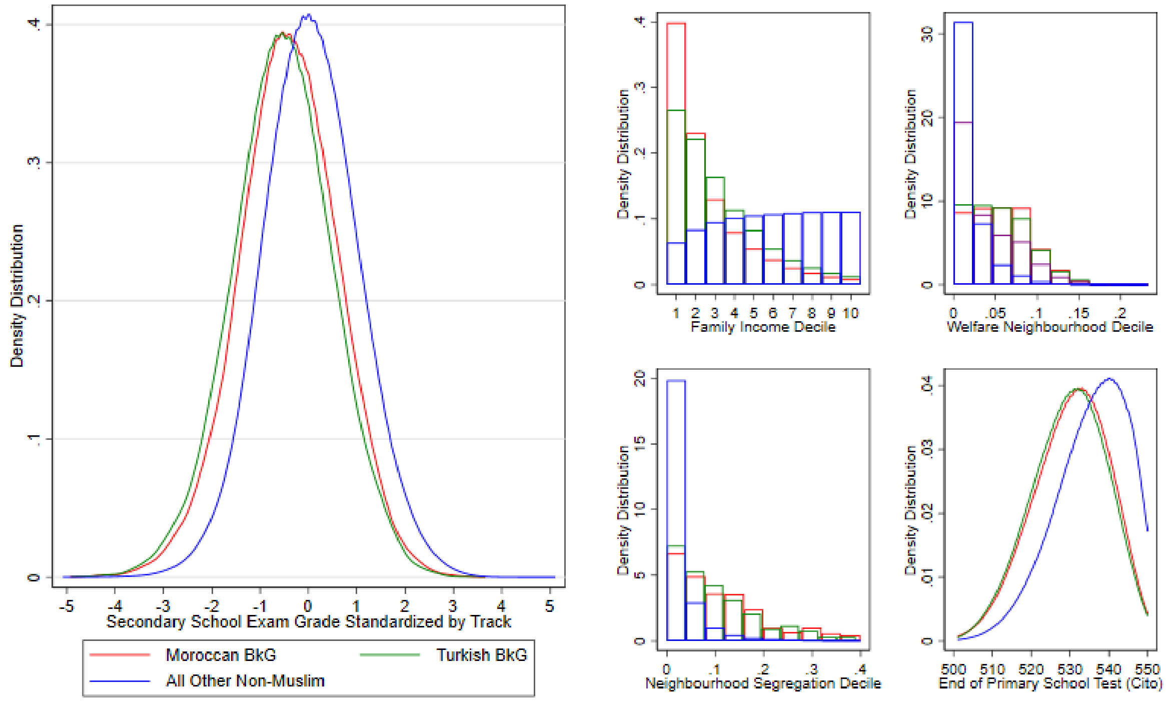
3.2 Survey Data

To determine who is observing Ramadan in the Netherlands, we use the SIM Survey (*Survey Integratie Minderheden*, meaning survey on minorities' integration). This survey was conducted in 2006, 2011, 2015 and 2020. It includes questions on religiosity, and specifically on how strictly individuals comply with fasting during Ramadan, answered by 19,002 respondents in total. We are able to link the respondents to our administrative data, and therefore have information on a rich set of characteristics of the respondents. There are 5,587 respondents with a

[22]See Appendix Section B for a detailed description of the classification of migration background.

Figure 4: Distribution of characteristics of students by migration background

A: Standardized final exam grade **B:** Other characteristics



Notes: The left graph shows the density distribution of the average central exam grade standardized by track and exam year in baseline year 2017. The right graph plots the density distribution of family income divided in deciles, the share of a neighbourhood using welfare (*bijstand*), the share of a neighbourhood with a Moroccan migration background and the grade of the end of primary school test (standardized by year). The distributions are given for students with a Moroccan migration background (red), Turkish migration background (green) and a non-Muslim migration background or no migration background (blue).

Turkish or Moroccan migration background (after cleaning and linking the data).[23] Descriptive statistics on the SIM survey data are in Table A3.

3.3 Mosque Data

Finally, we collect data on the language used and location of mosques in the Netherlands. We do so by scraping the geolocation of mosques, together with other information, from the website moskeewijzer.nl.[24] We then link the dataset with the administrative data from Statistics Netherlands. This allows us to compute the distance between the neighbourhood where a student is living and the neighbourhood of the closest mosque.

[23]See Appendix Section B for a detailed description of the SIM survey sample (selection) and cleaning of the SIM data.

[24]The data were collected in 2021.

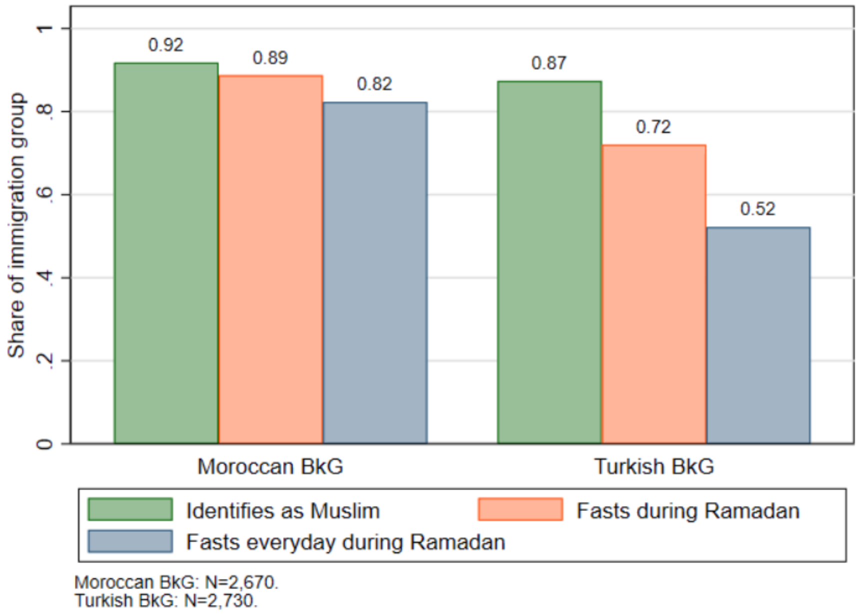
4 Identification strategy

4.1 Identification of the 'Treated' Group

To causally identify the effect of Ramadan on educational outcomes, we compare students who observe Ramadan during their high school final exams with students who don't. While we do not have direct insight into students who definitely observed Ramadan, we do have information on self-reported religiosity of 2,670 people in the Netherlands who have a Moroccan migration background and 2,730 who have a Turkish migration background. We focus on these groups for two reasons. First, they are the two largest migration groups in the Netherlands and they jointly represent more than 7% of the high school graduation exam takers (in 2017). Secondly, we don't have detailed information about Ramadan compliance for other migration groups from Muslim-majority countries, and they have unknown, and potentially very low, Ramadan compliance. The individuals who migrated to the Netherlands might have different religious attitudes from individuals in the country of origin, and using the major religion in the country of origin might not be a good proxy. For example, a relatively large proportion of the Iranian Dutch are not religious and around 20% of them are Christian, while the majority of the people living in Iran are Muslim (Huijink, 2018).

Figure 5 shows the share of individuals with a Turkish or Moroccan migration background who identify as Muslim and report that they fast every day during Ramadan. These two groups are almost equally likely to identify as Muslim. However, their behaviours diverge significantly on the question of how strictly they observe the fast. 84% of the respondents with a Moroccan background say they fast everyday during Ramadan, compared to only 54% of the respondents with a Turkish background. When we take this as a proxy for how strictly people comply with Ramadan, we would expect students with a Moroccan background to be more impacted during the overlap of Ramadan with the high-stakes secondary school exams. Other measures of compliance with the rules of Islam are shown in Table A2, confirming that on average, individuals with a Moroccan migration background in the Netherlands comply more with Islamic rules than individuals with a Turkish migration background. The sample of respondents in the SIM data is very similar to our sample of exam takers (see Table A3 in the Appendix), except for age. The sample of exam takers is younger than the sample of respondents in the SIM data. However, Figure A3 in the Appendix shows that everyday Ramadan compliance is very stable over age, which reassures us that our sample of 16 to 18 year old exam takers is not very different in terms of compliance with Ramadan.

Figure 5: Self-reported religiosity of SIM respondents with a second generation Moroccan or Turkish background



Notes: Graphical representation of the average self-reported religiosity (being Muslim, fasting during Ramadan and fasting every day during Ramadan) of respondents with a Moroccan migration background (on the left) and a Turkish migration background (on the right). Source: Author's calculations based on integration survey (*Survey Integratie Minderheden*) data on 2006, 2011, 2015, and 2020.

Individuals with a Moroccan migration background show high compliance with Ramadan fasting, making them an obvious sample for the treated group. However, it is not clear how to classify students with a Turkish migration background in this case. If we include them in the treatment or control group, we would incorrectly classify half of them. For this reason, rather than relying solely on migration background to classify individuals, we develop a classification algorithm based on machine learning. The algorithm exploits individual and neighbourhood characteristics, and assigns to each individual with a Moroccan or Turkish migration background a probability of complying with Ramadan obligations. The model is trained on the SIM survey respondents and used to populate the sample of high school exam takers with individual compliance probabilities. More details about this model are presented in section 4.3.

Our main specification uses treated individuals with a predicted probability above a certain threshold. In an alternative specification we use the continuous probability itself as a measure of treatment intensity. Finally, we use a third specification in which students with a Moroccan migration background are the treated group and students with a Turkish background are dropped from the sample. In all of our specifications the control group is all of the students who do not have a migration background from a Muslim-majority country (non-Muslim background) (see

Data Appendix).

4.2 Main Econometric Specification

To estimate how Ramadan compliance affects students' performance in the high school graduation exams, we exploit a natural experiment in which the timing of Ramadan and the exam period varies over the years, as described in Section 2.4. We apply a difference-in-differences approach comparing 'treated' students, those with a high probability of complying with Ramadan obligations, with the control group, composed of students who do not have a migration background from a Muslim-majority country. Equation (1) is our difference-in-differences main specification:

$$y_{i,t,s,p} = \alpha_{s,p} + \sum_{T=2014}^{2019} \beta_T Treated_i * T_t + \zeta Treated_i + \delta X_{i,t} + \gamma_t + \varepsilon_{i,t,s,p} \tag{1}$$

The outcome variable $y_{i,t,s,p}$ indicates the final exam grade (standardized by track and exam year) obtained in the central exam by student i, in year t, school s and subtrack p, or a pass dummy taking value one if the student graduated from high school. In our preferred specification, the variable $Treated_i$ is a dummy equal to one for individuals with a high probability of Ramadan compliance. In the alternative specifications, $Treated_i$ is the continuous probability predicted by the machine learning model or a dummy that takes value one if the individual has a Moroccan migration background. T_t represents a dummy equal to one if the exam is taken in year T. This will generate five estimates of β which we can plot to check for consistent differences in outcomes and pre-trends, which will help us to validate the common trends hypothesis. This specification also serves as a 'placebo in time' as it will reject or not the possibility that we would detect an impact of Ramadan on treated students before Ramadan actually coincided with the final exams. The coefficients that measure the average effect of Ramadan on exam outcomes are β_{2018} and β_{2019}. $\alpha_{p,s}$ are school-track-subtrack fixed effects, $X_{i,t}$ is a vector of controls and γ_t are exam year fixed effects. Controls included in vector $X_{i,t}$ are at the individual level: gender, age at exam, ability as measured by the end of primary school test (standardized by test year), continuous exam grade (standardized by track and exam year, which is dropped when it is used as a placebo outcome instead), moved up from lower track or failed the final exams last year, number of siblings, a dummy for having a Moroccan migration background (which is dropped when treatment is based on this), a dummy for having a Turkish migration background (which is dropped when treatment is based on Moroccan migration background), a dummy for a different migration background from Muslim-majority countries, time spent abroad, parental income, unemployment or illness benefits, and on the neighbourhood level: the share receiving assistance benefits, the average household size, the ratio of residents under 45 years old, fertility rate, share with Turkish migration background, and share with Moroccan migration background. Standard errors are clustered at school-track-subtrack exam year level.

We know from Figure 4 that there are important level differences in characteristics between the students in our treatment and control group, so we validate that students in our treatment group are not changing in other respects - compared to the students in the control group - in those years that Ramadan overlapped with the final exams. We do not find evidence of a compositional change in parental and neighbourhood characteristics over time (family income, parental unemployment, and neighbourhood segregation) (see Figure A4). A particular concern is selection in ability over time of the treated students. Figure A4 suggests positive selection in ability over time, as measured by the end-of-primary-school test (standardized by year), of the students in the treatment group. If we did not control for ability we would capture this trend, which might be explained by the integration of migrants over time and could bias our estimates. Therefore we control for ability in every analysis.

We estimate a second model where the year dummies ($Year_t$) are replaced by a single dummy ($Ramadan\ Year_t$) that takes value one if the central exam is taken in one of the years when it overlaps with Ramadan. The model is the following:

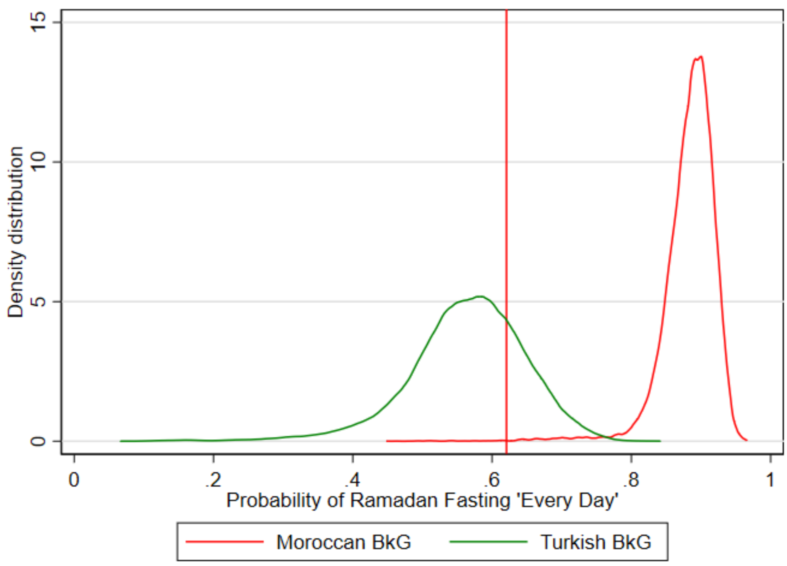
$$y_{i,p,s,t} = \alpha_{p,s} + \beta Treated_i * Ramadan\ Year_t + \zeta Treated_i + \delta X_{i,t} + \gamma_t + \varepsilon_{p,s,t} \qquad (2)$$

Standard error clustering, fixed effects and the included controls in equation (2) are the same as in equation (1).

4.3 Machine Learning (ML) Prediction of Ramadan Compliance

Before moving to the results, we present more details on how the treated group is constructed. Our aim is to identify a group with high compliance with Ramadan fasting. We build two machine learning classification algorithms: one for individuals with a Moroccan migration background and one for individuals with a Turkish migration background. These models are trained to predict the probability of fasting every day. We expect this measure to represent a proxy for the individual involvement in Ramadan, not only in terms of fasting compliance, but also in terms of participation in activities of the Muslim community during Ramadan. The model is based on individual and neighbourhood level characteristics and the training is performed on the respondents of the SIM survey. We test the performance of two different ML techniques: Random Forest and penalised logistic regression. As the latter outperforms the former in terms of out-of-sample prediction accuracy, we only use the penalized logistic regression to predict the exam takers' compliance probability. We produce these probabilities as follows: first, we train the classification algorithm on the SIM survey respondents. Then, we use this model to predict the probability of everyday compliance with Ramadan fasting for individuals with a Moroccan or Turkish migration background. Finally, we assign individuals in the control group (students with no or other migration background from non-Muslim majority countries) a probability of zero.

Figure 6: Predicted Ramadan compliance probabilities



Notes: Density distribution of the Ramadan compliance probability as predicted by the ML model by migration background (Moroccan BkG in red and Turkish BkG in green). The red vertical line defines the threshold for a student being classified as having a high Ramadan compliance probability.

Figures A6 and A7 in the appendix show the coefficients of the models. The distance to the closest mosque is the strongest predictor in both models. This is consistent with the idea that the distance to a mosque represents a strong proxy for the level of religiosity of individuals. Figure 6 shows the probabilities obtained with the two different models. The green line represents the distribution of the predicted probabilities of the Turkish background-specific model and the red one of the Moroccan background-specific model. We observe that the probabilities assigned to individuals with a Moroccan background are much more concentrated than the probabilities of individuals with a Turkish background. The vertical red line indicates the threshold that we use in our main specification for considering a student as treated (having a high probability of Ramadan compliance). This threshold is defined at 0.62074, which corresponds to the top quartile of the predicted probability among students with a Turkish migration background (who have a lower probability than the students with a Moroccan migration background). The sensitivity of the results to the use of alternative thresholds, as well as the use of the probability itself as a treatment intensity variable, is extensively investigated in our robustness analysis.

Table A4 compares the different compositions of the treatment group between our main specification, students with a high probability of Ramadan compliance, and the alternative specifications: i) individuals with Moroccan background considered as treated, and ii) the continuous probability of Ramadan compliance used as treatment intensity variable. If we use our main specification, students with a high Ramadan compliance probability, our treatment group includes practically all students with a Moroccan migration background, and 25% of the students with a Turkish migration background. Under the alternative specification of Moroccan migration background as treatment group, we drop students with a Turkish migration background from the sample. Using continuous treatment intensity, the average Ramadan probability of students with a Moroccan migration background is 0.88 and for students with a Turkish migration background it is 0.57. In all treatment specifications we exclude from the sample those students that do not have a Moroccan or Turkish migration background, but come from a Muslim-majority country (see Table A14). We do not have information on their exact Ramadan compliance probability, meaning that we cannot consider them as treated, even if we know that a share of the students with these migration backgrounds are likely to observe Ramadan. For the same reason, they would not represent a good control group. Therefore, we drop them from our sample rather than include them in the control group.

5 The Impact of Ramadan on Student Performance

5.1 Main Results

In this section, we present the results for the effect of Ramadan on secondary school outcomes, based on equation (1). Panel A of Figure 7 shows for students with a high probability of compliance with Ramadan, their central exam grade drops significantly in the years when the exams fall during Ramadan.[25] The size of the effect is 5.9 (in 2018) to 6.6% (in 2019) of a standard deviation, while the lack of evidence for pre-trends in the years before 2018 confirms the common trends hypothesis. As students with a high probability of observing Ramadan achieve lower grades in the high stakes exams, when they take those exams during Ramadan, they might be more likely to fail those exams. Panel B of Figure 7 shows a small drop in the pass rate for students with a high Ramadan probability in 2018 and 2019, which suggests that students with a high Ramadan probability might be less likely to pass their final exams.

Table 1 shows the simple DiD estimates based on equation (2) with students with a high Ramadan probability in the treatment group. The final grade of those students decreased by 5.8% of a standard deviation during the years that Ramadan coincided with the final exams (Column 1). The probability that students with a high Ramadan probability graduated from high school also dropped during Ramadan, as they were 1.1 percentage points less likely to pass the final exam (Column 2), corresponding to an increase of 8.3% in their probability of failing the final exam. The probability of passing the final exam depends for 50% on the continuous exam, which is not affected by Ramadan. This implies that the effect size would be larger (doubled) if passing the exam depended uniquely on the final exam. Table A5 shows that the results are robust to the use of alternative definitions of the treatment group.

5.2 Robustness and Placebos

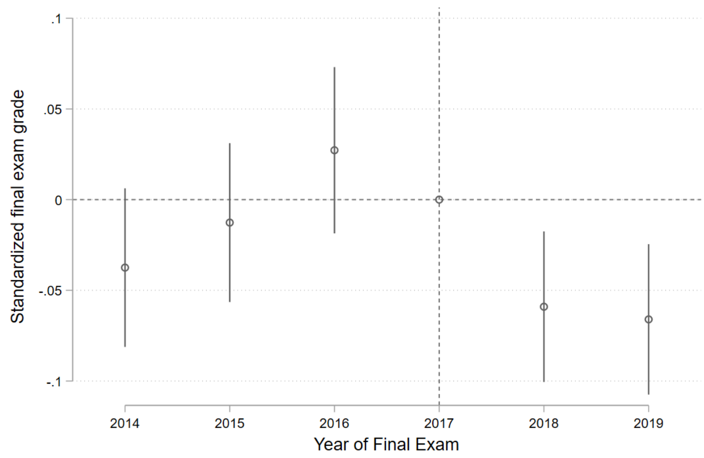
We validate that students in our treatment group are not changing in other respects. One concern is that the negative impact can be explained by negative selection in ability over time of the treated students. Therefore, we look at whether the standardized ‘continuous’ exam grade, as a measure of ability, changes for students in the treatment group compared to the control group in those years that Ramadan overlapped with the final exams. This grade is an average of several decentralized exams that students take in the last two or three years (depending on the track) of high school. Figure A8 shows that, once we control for standardized end-of-primary-school test scores (our other proxy for ability), there is no evidence of a compositional change in ability of exam takers with a high Ramadan probability (apart from a small jump in 2014, which is far from our baseline and treated years).[26] Column 3 in Table 1 confirms this.

[25]Results are robust to using Moroccan migration background or continuous Ramadan probability as treatment instead of high Ramadan probability. We have also checked whether Ramadan observing students avoid the final exams by not showing up, but find no evidence for this.

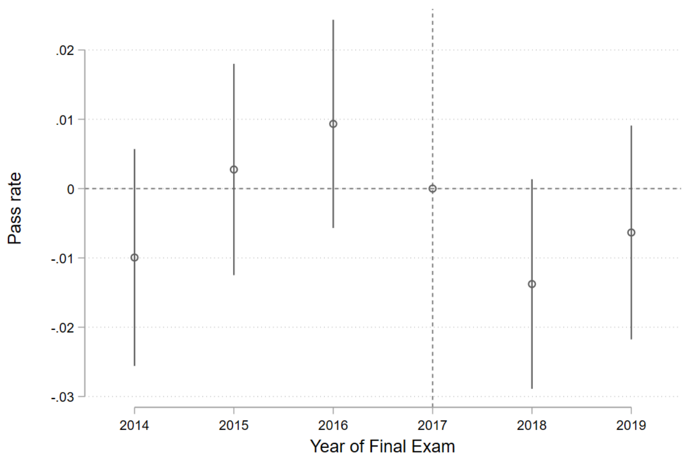
[26]Figure A4 does show evidence for a compositional change in ability of exam takers with a high Ramadan probability when we take the standardized end-of-primary-school test score as an outcome variable. Hence, it is

Figure 7: Impact of Ramadan on secondary education outcomes of students with a high Ramadan compliance probability

A: Standardized final exam grade



B: Pass rate



Notes: The figures plot the point estimates and 99% confidence intervals for the coefficients (β_t) that are estimated using our main specification (equation (1)). Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan compliance probability, as predicted by our machine learning model. The exams in years 2014 to 2017 did not overlap with Ramadan (and hence we would not expect an effect here), while exams in 2018 and 2019 did overlap with Ramadan. Final exam grade is measured as the average of the central exam grades standardized with mean zero and standard deviation one by track and exam year. Probability of overall pass is measured by a dummy that indicates whether the student passed the final exams (and graduated from secondary education). All controls are included, as in Table 1. School-subtrack and exam year fixed effects are included. Robust standard errors are clustered at the school-track-subtrack-year level. The horizontal dashed line marks the zero or no effect.

very important to control for ability. This test, which takes the continuous exam grade as an outcome variable while controlling for the standardized end-of-primary-school test score, shows that conditional on this earlier test score, there is no evidence of a compositional change.

Table 1: Impact of Ramadan on secondary school outcomes of students with a high Ramadan probability

	Secondary Education Outcomes				
	Final	Pass	Cont-inuous	Improved	Final incl. retakes
	(1)	(2)	(3)	(4)	(5)
Ramadan Exam Years *Treated	- 0.058***	- 0.011***	0.019	0.007	- 0.059***
	(.010)	(.004)	(.012)	(.005)	(.010)
School-Subtrack f.e.	Yes	Yes	Yes	Yes	Yes
Exam year f.e.	Yes	Yes	Yes	Yes	Yes
All controls	Yes	Yes	Yes	Yes	Yes
Mean outcome before	6.10	86.8%	6.27	26.3%	6.18
Share treated	4.7%	4.7%	4.7%	4.7%	4.7%
Observations	776,219	776,284	776,284	776,284	776,284

Notes: The table shows point estimates for the coefficient β estimated using equation (2) with students with a high Ramadan probability being treated and students without a Muslim-majority migration background being the control group. Final is the average of the central exam grades standardized with mean zero and standard deviation one by track and final exam year. Pass is a dummy indicating that a student passed the final exams (including retakes) and graduated from secondary education. Continuous is the average exam grade of the continuous school exams standardized with mean zero and standard deviation one by track and exam year. Improved is a dummy that takes value 1 if the student improved his or her final exam grade in a retake. Final incl. retakes is the average final exam grade of the central exam standardized with mean zero and standard deviation one by track and exam year including the grades of the first and second retake. The difference in number of observations between the analyses with final grade in column 1 and outcomes in other columns can be explained by students not taking the exam during the first period, but during the first or second retake due to e.g. illness. Mean outcome before is the average outcome for treated students in the pre-treatment years 2014 to 2017. Controls included are: gender, age at exam, end of primary school test score (standardized by year), continuous exam grade (standardized by track and exam year), moved up from lower track or failed the final exams last year, number of siblings, a dummy for being Moroccan or Turkish, time spent abroad, parental income, unemployment or illness benefits, and on the neighbourhood level: the share receiving assistance benefits, the average household size, the ratio under 45 years old, fertility rate, share with Turkish migration background, share with Moroccan migration background. Robust standard errors are clustered at the school-track-subtrack-year level in parentheses. $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We also check whether students can compensate for the negative impact of Ramadan by retaking an exam to improve a subject grade. Column 4 of Table 1 shows that there is only a very small (and not significant) increase in the share of students with a high Ramadan probability that improve their final grade with a retake during the years in which the final exam overlapped with Ramadan. The impact of Ramadan on the final grade including retakes is almost identical to the impact on the final grade excluding retakes (Column 5). Hence, students do not compensate for the negative impact of Ramadan with retakes. The preparation period for the retakes overlapped with Ramadan, so the findings of Oosterbeek and van der Klaauw (2013) might explain this result: preparing an exam (a retake in this case) during Ramadan negatively affects the performance of Muslim students.

We further check the robustness of the results by including i) sibling fixed effects and ii) (extended) family fixed effects. We link each exam taker to their siblings and cousins and compare exam takers within families, to take out any family fixed effects. In this way we can control for unobserved characteristics that are shared among siblings and cousins. As not every exam taker has a sibling or cousin that also took an exam in 2014-2019, and as we are not able to identify families for all exam takers (for exam takers with a migration background we can only observe family if their grandparents also migrated to the Netherlands),[27] the number of observations drops by approximately 47.4% when including sibling fixed effects and 36.4% when including (extended) family fixed effects. Table A6 shows that our results are robust to using sibling and family fixed effects (the effect size even becomes slightly larger).

We conduct two placebo in treatment group tests with groups that are similar in socioeconomic characteristics to our treatment group, but that we do not expect to be Muslim. We compute the Ramadan compliance probability, using both the within-Turkish machine learning model and the within-Moroccan machine learning model, for students who do not have a migration background from a Muslim-majority country and hence should not observe Ramadan. We use non-Muslim students with a high placebo Ramadan probability as the treatment group. Using these placebo Ramadan probabilities, we do not find any evidence of a change in secondary education outcomes when Ramadan overlapped with the final exams (see Table A7). The results of these placebo tests provide reassurance that the decrease in secondary education outcomes we find when Ramadan overlaps with final exams for students that are likely to observe Ramadan is due to Ramadan, and not due to another shock that affected students with a migration background or a socioeconomically disadvantaged background.

Lastly, we replicate our main analysis for students with a high Ramadan probability by setting different thresholds for the Ramadan probability being high. Figure A9 shows that our results are robust to using different thresholds and get stronger the higher we set the threshold.

6 Heterogeneity Analysis

6.1 Student and School Characteristics

In this section, we explore who is impacted most by Ramadan by examining heterogeneity of the effect by student and school characteristics. First, we check whether the effect on secondary

[27] See Data Appendix for more details on the creation of family links.

education outcomes differs by gender. We find a very similar impact of Ramadan on the final grade and pass rate of female and male students, with a slightly larger impact on the final grade of male students (see Figure 8). Second, we check whether the effect on the central exam grade and pass rate differs for high and low performers at the individual and school level.[28] At the individual level, we measure ability with the continuous exam grade and find no heterogeneity by student ability in the impact on the final grade, but we do find heterogeneity in the impact on the pass rate (see Figure 8). The effect of Ramadan on the pass rate is (naturally) driven by the 25% worst performing students. Those students, who are already on the margin of graduating, are 11% more likely to fail their exams and not graduate if the final exams overlap with Ramadan.

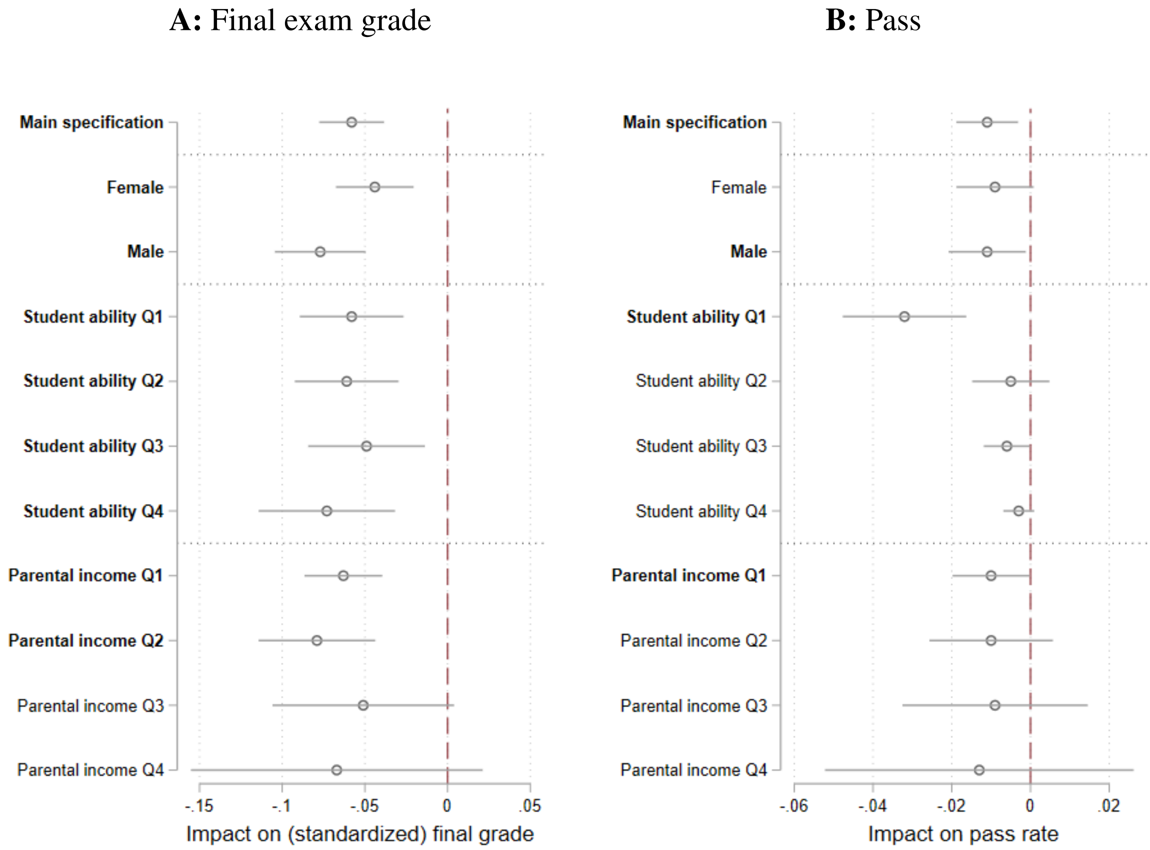
At the school level, we measure mean school performance by the average final exam grade of the years before Ramadan overlapped with the final exams. The effect seems to be slightly larger for the lowest performing schools, and there is no impact at all on Ramadan-compliant students in the top 25% best performing schools (see Figure 9). Similarly, we find the largest effect on the pass rate in the 25% worst performing schools (Ramadan-compliant students in those schools are 11.3% more likely to not graduate when Ramadan overlaps with central exams) and no impact at all in the 50% best performing schools.

Third, we check whether the effect differs by income at the school and individual level. While the income of the family of the student does not seem to interact with the impact of Ramadan (see Figure 8), mean school income (of the student body) clearly does. The impact of Ramadan on final grade and pass rate is fully driven by students with a high Ramadan-compliance probability in the schools in the lowest mean school income quartile (see Figure 9).

Finally, we check whether the effect varies by school segregation level, where we measure segregation by the proportion of Ramadan-compliant students in a school in the years before Ramadan overlapped with the final exams. We only find a negative impact of Ramadan on the final grades and pass rate of students with a high Ramadan-compliance probability in the 25% of schools with the highest share of Ramadan-compliant students. That we only find a Ramadan impact in the most segregated schools may reflect high self-selection, because compliers go to

[28] Results by track show a negative coefficient in each track, but its size and precision varies. In the pre-university education stream we do not find a significant negative impact, in senior general secondary education we do find a large significant negative impact, and in pre-vocational secondary education we also find a (slightly smaller) significant negative impact. The share treated varies a lot across tracks, from 1.6% in pre-university education to more than 6.5% in pre-vocational secondary education.

Figure 8: Heterogeneity in the impact of Ramadan on secondary education outcomes of high probability Ramadan-compliant students, by student characteristics



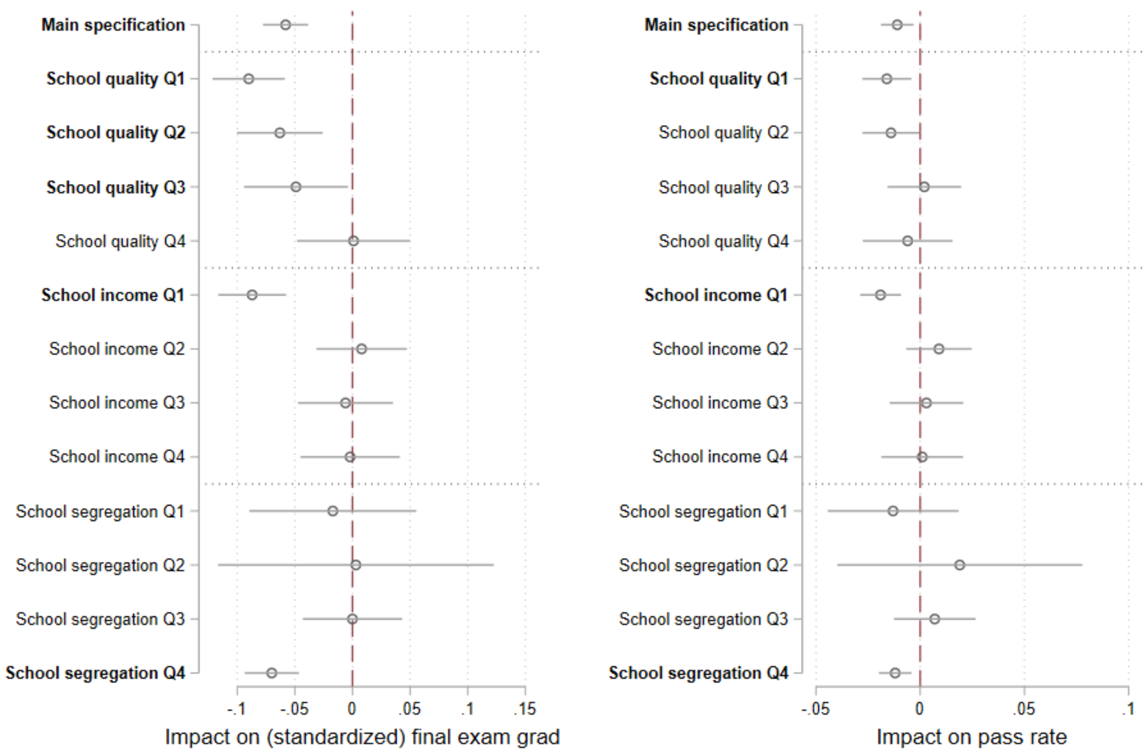
Notes: The figures plot the point estimates and 95% confidence intervals for the coefficient β that is estimated using equation (2), split by student characteristics. Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan-compliance probability, as predicted by our machine learning model. The first row shows our main effect for the full sample of students. The second and third row show the main effect for female and male students separately. The fourth to seventh rows show the effect for students split into four quartiles by ability, which is measured by their standardized continuous exam score. The eight to eleventh rows show the effect for students split into four quartiles by parental income, measured in the year before the exam took place. Final exam grade is measured as the average of the central exam grades standardized with mean zero and standard deviation one by track and exam year. Probability of overall pass is measured by a dummy that indicates whether the student passed the final exams (and graduated from secondary education). Labels on the y-axis are in bold when the coefficient is significant at a 5% significance level. All controls are included, as in Table **??**. School-subtrack and exam year fixed effects are included. Robust standard errors are clustered at the school-track-subtrack-year level. The horizontal dashed line marks the zero or no effect.

the same schools. However, it could also stem from peer effects in compliance, in that Muslim students would follow Ramadan rules more strictly as there are more compliant students around them.

While anecdotal mentions of ’Ramadan peer effects’ are frequent, there is no scientific evidence for it. We test for this mechanism by exploiting our DiD setting and using cohort variations in compliance probability. We define a cohort here as students in the same subtrack, track, school and exam year. Figure A10 shows that there is a lot of cohort variation beyond

Figure 9: Heterogeneity in impact of Ramadan on secondary education outcomes of high probability Ramadan-compliant students, by school characteristics

A: Final exam grade **B:** Pass



Notes: The figures plot the point estimates and 95% confidence intervals for the coefficient β that is estimated using equation (2), split by school characteristics. Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. The first row shows our main effect for the full sample of students. The second to fifth rows show the effect for students split into four quartiles by school quality, which is measured by the average standardized final exam grade in the years before Ramadan overlapped with the final exam. The sixth to ninth rows show the effect for students split into four quartiles by school income, measured as the average of parental income decile in schools. The tenth to thirteenth rows show the effect for students split into four quartiles by school segregation, measured as the average share of students with a high Ramadan-compliance probability in a school. Final exam grade is measured as the average of the central exam grades standardized with mean zero and standard deviation one by track and exam year. Probability of overall pass is measured by a dummy that indicates whether the student passed the final exams (and graduated from secondary education). Labels on the y-axis are in bold when the coefficient is significant at a 5% significance level. All controls are included, as in Table 1. School-subtrack and exam year fixed effects are included. Robust standard errors are clustered at the school-track-subtrack-year level. The horizontal dashed line marks the zero or no effect.

school self selection, both before and after 2018. Cohort variation in the share of students with a high Ramadan compliance probability is more random than school variation, as cohorts suffer less from self selection than schools. Using the share of peers with a high Ramadan probability in a student's cohort might therefore give us a better indication of whether the heterogeneity by segregation is fully driven by self-selection or also by Ramadan peer effects. When we interact our main coefficient of interest with the share of peers with a high Ramadan probability in a student's cohort, we find evidence to suggest an additional impact of Ramadan on final grades and pass rate when a student is likely to have more Ramadan-compliant peers, which suggests there might be Ramadan peer effects (see Figure A11).

6.2 Effect by Type of Exam Taken

In this section we run heterogeneity analyses at subject level (e.g. the exam in math, English, etc.) in the central examination as the outcome variable, rather than the average final central exam grade.[29] We first replicate our main analysis with the new data.[30] Our results are qualitatively the same as in the main specification, but the magnitude of our estimates is smaller (see Table A8).[31]

The availability of data at subject level allows us to show (in Table A10) the results for different groups of exams: languages (reading comprehension tests in Dutch, English, German and French), more memory-based exams (history, geography and biology), and math.[32] Memory-based exams were negatively affected by approximately 4.2% of a standard deviation. The impact on language exams is about 40% smaller (only 2.4% of a standard deviation) than the impact on memory-based exams. We find a very strong effect on math (grades dropped with 8.9% of a standard deviation), which is more than twice as large as the effect on memory-based exams. The first two findings are in line with the literature on sleep deprivation and cognition, that finds that sleep deprivation negatively affects memory, but is much less likely to affect reasoning and crystallized intelligence (Lim and Dinges, 2010). This suggests that the impact of Ramadan on academic performance could be driven by sleep deprivation, more than by fasting. However, math exams are strongly affected, which is not in line with the idea that sleep deprivation is mainly driving our effect because those exams might also be classified as dependent on reasoning and crystallized intelligence. Hence, some results point into the direction of sleep

[29] Figure A12 provides an overview of the schedule of these exams.

[30] Given the different structure of the data, the regression employed to estimate the effects at exam subject level has some differences with equation(2). This regression is presented in equation(3) in the appendix.

[31] The smaller magnitude of the results using the subject level data can be explained by the fact that students in the pre-university education stream take more exams than students in other streams, yielding more observations for those students. As a consequence, by focusing on subject level outcomes, we give more weight to students in the pre-university stream, who are less affected by Ramadan exposure. Moreover, the use of a rich set of fixed effects to account for the timing of the exams might explain the smaller magnitude obtained with this specification. Finally, we standardize exams at subject-track-year level, so that for each subject in a certain track in a specific year the mean grade is zero with standard deviation one. This is different from the standardization adopted in our main specification, where we transform the final (comprehensive of all subjects) grade such that it has mean zero and standard deviation one at track-year level.

[32] Note that in this analysis we can only look at a selected sample that actually takes a final exam for one or more of those subjects.

deprivation as the main mechanism, while others point into the direction of fasting as the main mechanism. In the following section, we explore the timing of exams within days and weeks, to explore whether our effect is driven by sleep or by fasting.

7 Mechanisms & Policy Implications

7.1 Fasting or Sleep Deprivation?

In this section we explore the role of fasting and other features of Ramadan that may impact performance. Muslims often change their sleep patterns during Ramadan, (Bahammam, 2003; Bahammam, 2006; Margolis and Reed, 2004) as the requirement to fast during daylight shifts mealtimes to before sunrise and after sundown. Hence, students might suffer from sleep deprivation during the exams. In addition, the time spent celebrating the holy month of Islam could affect the amount of time that students dedicate to studying, which could also negatively impact academic performance.

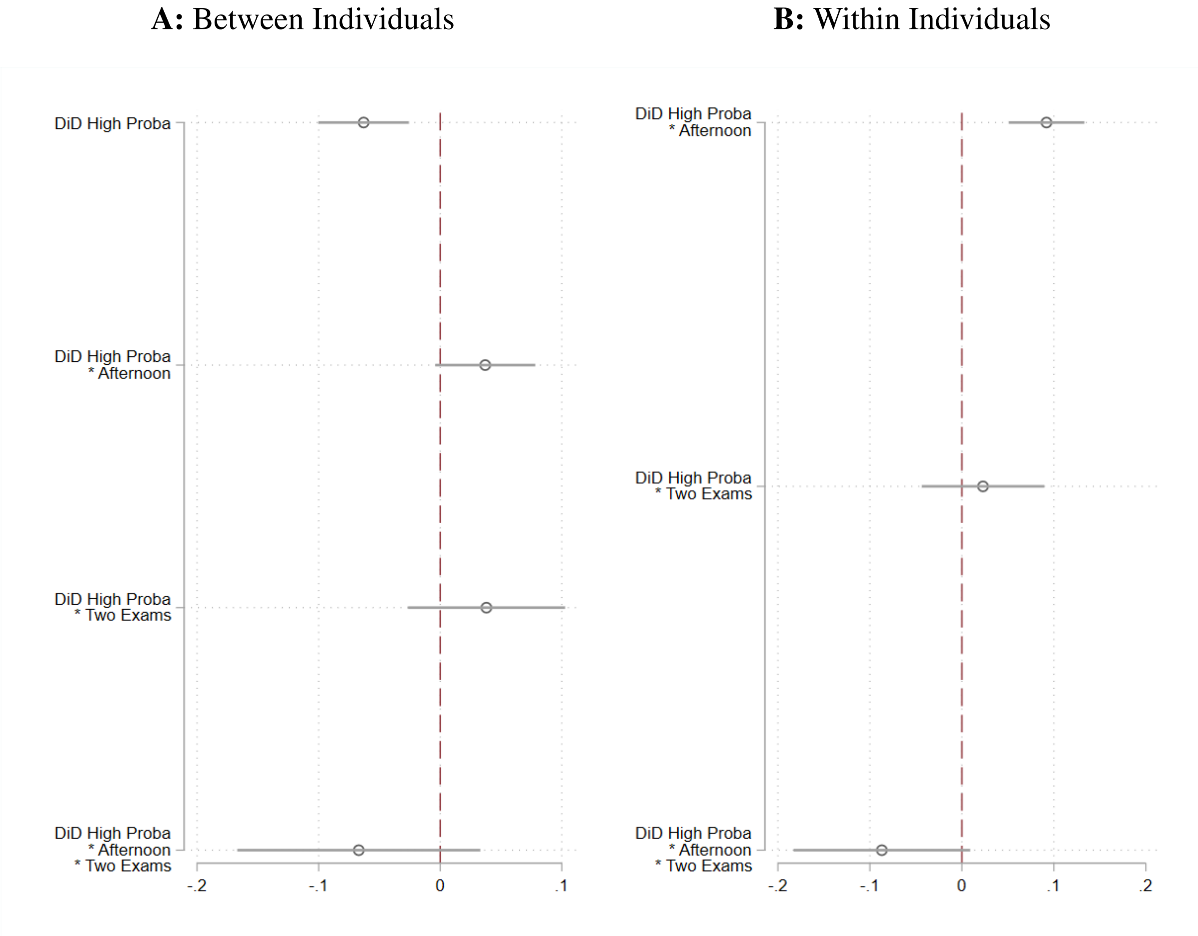
First, we check whether the effect size changes over time. Our assumption is that if the negative effect we find is mainly driven by a lack of sleep, the effect would become more negative over time. If the effect is stronger in the first week of Ramadan, it would be more likely that the decline in cognitive performance is mainly due to fasting. We base these assumptions on the literature on sleep, which clearly shows how sleep deprivation accumulates Van Dongen et al. (2003), leading to worse cognitive functioning as the nights affected by lack of sleep accumulate. Our assumption on fasting affecting the first week more than later weeks is based on the anecdotal evidence that the first few days of fasting are more taxing, and that over time the body adapts. To check the empirical impact of these mechanisms, we decompose the effect of Ramadan on exam grades according to whether an exam falls in the first, second or third week of Ramadan. Figure A13 shows that the effect over the three weeks is quite constant. We also report the coefficient of the additional effect of the exact day into Ramadan when each exam is taken. This is a precisely estimated zero, confirming that the drop in performance is constant over time. This suggests that neither cumulative sleep deprivation nor fasting can alone explain our results. Finally, we also compute an effect for the few exams that were sat before Ramadan in 2019. We find that the performance did not drop for complying students. This means that if there is an anticipation effect, for example because of a change of eating habits to transition smoothly to the Ramadan fast, this is not particularly strong.

Second, we check whether the effect is different for morning and afternoon exams. We hypothesize that if fasting is the main driver of the Ramadan effect, the effect would be smaller in the morning, closer to breakfast. If we expect the lack of sleep to be driving the effect, we would expect the effect to be larger in the morning, as students have no time to sleep in, or to have a midday nap. In Table A9 we show that the coefficient capturing the additional effect of exams held in the afternoon is not significant and is close to zero in magnitude, showing that the timing of the exam does not contribute to the impact of Ramadan. However, the previous findings do not take into account that on some days, students have an exam both in the morning and in the afternoon. When we consider in our model the interaction between timing of the exam and the number of exams taken in one day, we find different results, shown in Figure 10. Exams that are held in the afternoon, when no exam is scheduled in the morning, are affected the least. We believe that this is driven by students being able to rest in the morning.

Finally, the decline in exam results might be driven by additional time constraints that students face during Ramadan. Students often spend more time with their relatives and participate in Ramadan celebrations. This reduces the time available to prepare for exams and could result in lower scores.

We check whether this might play a role in two ways. First, we check whether the effect is different depending on whether a student has a small or large family. We hypothesize that a large family might have more social obligations during Ramadan. We do not find much evidence for this, with only slightly larger effects on outcomes for students with larger families (see Table A11). Second, we focus again on the results split by type of subject. If we believe that our results are driven by time constraints that affect the preparation for the exams but not the exams themselves, we would only expect to find an effect on exams that depend a lot on short-term memory and not so much on other subjects (like languages that use reading comprehension tests as final exams, or math). However, Table A10 shows that not only short-term memory exams, but also math exams - and language exams to a lesser extent - are affected. Math exams are affected the most, which suggests that our effect is not (solely) driven by less preparation time due to social activities. These two findings together suggest that time constraints are not a significant contributor to the negative impact of Ramadan on secondary education outcomes of students with a high Ramadan-compliance probability.

Figure 10: Effect on final grade on days when there are two exams



Notes: The figure displays the point estimates and 95% confidence intervals for the coefficients estimated using equation (4) with (on the right) and without (on the left) individual fixed effects, based on the subject level data. The dependent variable is the grade obtained in the exam (after retakes) standardized at subject-track-year level. *Afternoon* is a dummy that takes value 1 if an exam is held in the afternoon. *DiD High Proba* is a dummy equal to one if the exam is taken during the Ramadan month by a student with high probability of Ramadan compliance. *Two Exams* is dummy that takes value 1 if an exam is held on a day in which a student has to sit two exams. The coefficient of interest is the triple interaction between these dummies, which captures the additional impact of sitting an exam in the afternoon after a student already took another exam in the morning. All controls are included, as in Table 1. We include school-track-exam subject and treatment type-exam subject-afternoon-two day exam fixed effects. Robust standard errors are clustered at the school-track-subject-year level.

7.2 Policy Implications

Under-performing in exams, and especially failing exams, has long-term consequences. Students who do not achieve a passing grade have to repeat the year or go into adult education to obtain a diploma. Table A12 shows that students that have a high Ramadan-compliance probability are 12.4% more likely to repeat the year if their final exams coincided with Ramadan. The costs of this are considerable, as the student will enter the labor market a year later and suffer from a year of earnings loss. Meulen (2023) estimates that, in the Netherlands, students who marginally fail the high school graduation exam earn 3000 euro less per year at the age 28 than those who marginally pass the exam. The government also bears more costs if a student fails to graduate: it spends 7500 euro for an additional year of school for a student (Van Vuuren and Van der Wiel, 2017).

Students have different options when they fail their exams, depending on their age. At age 18, they are no longer obliged to obtain a so-called "start qualification", which is a senior general secondary education or pre-university education diploma, or a diploma in secondary vocational education. We split our analysis between students who are age 18 on the 1st of September after the final exams, when secondary and tertiary education normally start again, and those who are not. We find that our impact on students repeating the year (or going into adult education) is driven by students younger than 18, while students older than 18 are 22.5% more likely to drop out if their final exams overlapped with Ramadan (see Table A13).

Figure A13 shows that in what we have defined as week zero, the week before the beginning of Ramadan, we find no significant change in grades for students with a high Ramadan-compliance probability. The most straightforward policy implication would therefore be to move exams entirely so that they do not overlap with Ramadan, or take place (right) after Ramadan (in case there are ongoing effects of Ramadan). However, as exams take place over several weeks, it might not be (politically) feasible to fully avoid the month of Ramadan. A second, and milder, policy option is to take into account that the impact is largest for students with two exams on one day (see Figure 10). To mitigate the impact of Ramadan compliance on secondary education outcomes, it would therefore be best only schedule one exam per day.

8 Conclusion

Using a unique natural experiment combined with high quality administrative data, we causally estimate the impact of Ramadan observance on student performance during high stakes exams. Our main finding is that there is a large and significant negative impact on average grades and pass probability for this already disadvantaged population. Specifically, we find that those most likely to fast obtained grades 5.8% of a standard deviation (SD) lower during Ramadan. These are large effects which are quantitatively comparable to taking an exam during heat waves (1 SD higher temperature (Park, 2022)) or during pollution spikes (1.5 SD higher (Ebenstein et al., 2016)). It is however smaller than the effect size (10% of a SD) that Oosterbeek and van der Klaauw (2013) find when Ramadan overlaps with a university course *before* the exam. In the

Dutch context, it is equivalent to having a teacher with 6 years less experience before an exam (Gerritsen et al., 2017) or twice the effect of students having legal cannabis access (Marie and Zölitz, 2017).

This drop in exam performance during Ramadan also strongly impacted pass rates. Because the religious obligations of affected minority students were not accommodated, there was a large increase by 16.4% in the pre-existing achievement gap with their non-Muslim peers. These negative impacts on education outcomes were especially bad for low achieving students, and almost entirely driven by those in religiously segregated schools. For those in schools with the highest quartile of Ramadan-complying students, the achievement gap increased by a quarter. There were also potential longer term effects for affected students as 22.5% more of those above 18, who could legally drop out, ended up doing so.

The simplest remedy might be to change the timing of national exams to accommodate this religious minority, but it has proven practically and politically difficult to implement.[33] In our context, the exploration as to whether the Ramadan impact stemming from fasting or sleeping yields the interesting finding that the worst outcomes are found for the afternoon exam when there are two exams on the same day. Adapting the schedule to remove two-exam days is a simple change that would attenuate Ramadan effects.

Another possibility would be for religious leaders to give permission to their members to (exceptionally) adapt to the rules of the country where they live. In our case, it would involve Islamic authorities producing *fatwas* which would allow Ramadan postponements for those undertaking high stakes exams. Such exemptions exist for health reasons but not educational ones and those living in non-Muslim countries have received specific guidelines as to how to observe Ramadan in difficult contexts. [34] As these requests would need to come from impacted Muslims, the clear findings of our paper might help to make this argument. There is otherwise recent robust evidence that awareness of the negative impact of Ramadan observance on productivity can lead to changes in the fasting behaviour of those concerned (Wang et al., 2022). The results presented here may help inform individuals who have to decide between their religious obligations and their requirement to perform specific civic duties.

[33]There is precedent elsewhere in the world, however, as in 2020 the Seventh Day Adventists were able to force the Brazilian government, after a long legal battle that took them all the way to the Supreme Court, to allow them to take selective exams ('concursos') on days other than Saturday, which is their most holy day.

[34]Various fatwas were declared in the mid-2010s to give permission to Muslims living in northern countries, such as Iceland where the day could last up to 22 hours, to follow the fasting schedule of other (more southern) countries.

References

Adolphus, Katie, Clare L Lawton, Claire L Champ, and Louise Dye, "The effects of breakfast and breakfast composition on cognition in children and adolescents: a systematic review," *Advances in Nutrition*, 2016, *7* (3), 590S–612S.

Almond, Douglas and Bhashkar Mazumder, "Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy," *American Economic Journal: Applied Economics*, 2011, *3* (4), 56–85.

—, —, and Reyn Van Ewijk, "In utero Ramadan exposure and children's academic performance," *The Economic Journal*, 2015, *125* (589), 1501–1533.

Bahammam, Ahmed, "Sleep pattern, daytime sleepiness, and eating habits during the month of Ramadan," *Sleep and Hypnosis*, 2003, *5*, 165–174.

—, "Does Ramadan fasting affect sleep?," *International journal of clinical practice*, 2006, *60* (12), 1631–1637.

Bahammam, Ahmed S, Abdulrhman M Alaseem, Abdulmajid A Alzakri, and Munir M Sharif, "The effects of Ramadan fasting on sleep patterns and daytime sleepiness: An objective assessment," *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences*, 2013, *18* (2), 127.

Bertoli, Paola, Veronica Grembi, and Judit Vall-Castello, "The Ramadan effect in the workplace," *Unpublished Manuscript*, 2022.

Blanden, Jo, Matthias Doepke, and Jan Stuhler, "Educational inequality," Technical Report, National Bureau of Economic Research 2022.

Campante, Filipe and David Yanagizawa-Drott, "Does religion affect economic growth and happiness?," *The Quarterly Journal of Economics*, 2015, *130* (2), 615–658.

Colussi, Tommaso, Ingo E. Isphording, and Nico Pestel, "Minority salience and political extremism," *American Economic Journal: Applied Economics*, 2021, *13* (3), 237–71.

Csipo, Tamas, Agnes Lipecz, Cameron Owens, Peter Mukli, Jonathan W Perry, Stefano Tarantini, Priya Balasubramanian, Ádám Nyúl-Tóth, Valeriya Yabluchanska, Farzaneh A Sorond et al., "Sleep deprivation impairs cognitive performance, alters task-associated cerebral blood flow and decreases cortical neurovascular coupling-related hemodynamic responses," *Scientific reports*, 2021, *11* (1), 20994.

Curcio, Giuseppe, Michele Ferrara, and Luigi De Gennaro, "Sleep loss, learning capacity and academic performance," *Sleep medicine reviews*, 2006, *10* (5), 323–337.

Dongen, Hans PA Van, Greg Maislin, Janet M Mullington, and David F Dinges, "The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation," *Sleep*, 2003, *26* (2), 117–126.

Ebenstein, Avraham, Victor Lavy, and Sefi Roth, "The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution," *American Economic Journal: Applied Economics*, 2016, *8* (4), 36–65.

Ewijk, Reyn Van, "Long-term health effects on the next generation of Ramadan fasting during pregnancy," *Journal of health economics*, 2011, *30* (6), 1246–1260.

Gerritsen, Sander, Erik Plug, and Dinand Webbink, "Teacher quality and student achievement: Evidence from a sample of Dutch twins," *Journal of applied econometrics*, 2017, *32* (3), 643–660.

Hornung, Erik, Guido Schwerdt, and Maurizio Strazzeri, "Religious practice and student performance: Evidence from Ramadan fasting," *Journal of economic behavior & organization*, 2023, *205*, 100–119.

Huijink, Willem, *De religieuze beleving van moslims in Nederland* 2018.

Kökkizil, Melike, "Religiosity and gender gap in STEM performances," *Unpublished manuscript*, 2022.

Lim, Julian and David F. Dinges, "A meta-analysis of the impact of short-term sleep deprivation on cognitive variables," *Psychological bulletin*, 2010, *136* (3), 375–389.

Lugo, Luis, Alan Cooperman, Erin O'Connell, and Sandra Stencel, "The future of the global Muslim population," *Pew Research Center, USA*, 2011, pp. 1–209.

Margolis, Stephen A and Richard L Reed, "Effect of religious practices of Ramadan on sleep and perceived sleepiness of medical students," *Teaching and learning in medicine*, 2004, *16* (2), 145–149.

Marie, Olivier and Ulf Zölitz, ""High" achievers? cannabis access and academic performance," *The Review of Economic Studies*, 2017, *84* (3), 1210–1237.

Montero, Eduardo and Dean Yang, "Religious festivals and economic development: Evidence from the timing of mexican saint day festivals," *American Economic Review*, 2022, *112* (10), 3176–3214.

Najafabadi, Mahboubeh Ghayour, Laya Rahbar Nikoukar, Amir Memari, Hamed Ekhtiari, and Sara Beygi, "Does ramadan fasting adversely affect cognitive function in young females?," *Scientifica*, 2015, *2015.*

Nuryakin, Chaikal, Pyan A Muchtar, Natanael WG Massie, and Sean Hambali, "Having exams during Ramadan: The case of Indonesia," *Economics & Human Biology*, 2022, *47*, 101183.

Oosterbeek, Hessel and Bas van der Klaauw, "Ramadan, fasting and educational outcomes," *Economics of Education Review*, 2013, *34*, 219–226.

Park, R Jisung, "Hot temperature and high-stakes performance," *Journal of Human Resources*, 2022, *57* (2), 400–434.

Schofield, Heather, "Ramadan fasting and agricultural output," *Unpublished manuscript*, 2020.

ter Meulen, Simon, "Long-Term Effects of Grade Retention," *CESifo Working Paper*, 2023.

Tian, Ho-Heng, Abdul-Rashid Aziz, Weileen Png, Mohamed Faizul Wahid, Donald Yeo, and Ai-Li Constance Png, "Effects of fasting during Ramadan month on cognitive function in Muslim athletes," *Asian Journal of Sports Medicine*, 2011, *2* (3), 145.

Vuuren, Daniel Van and Karen Van der Wiel, *Zittenblijven in het primair en voortgezet onderwijs. Een inventarisatie van de voor- en nadelen* CPB Policy Brief 2015/01 2017.

Wang, Ao, Shaoda Wang, and Xiaoyang Ye, "When Information Conflicts with Obligations: the Role of Motivated Cognition," Technical Report, National Bureau of Economic Research 2022.

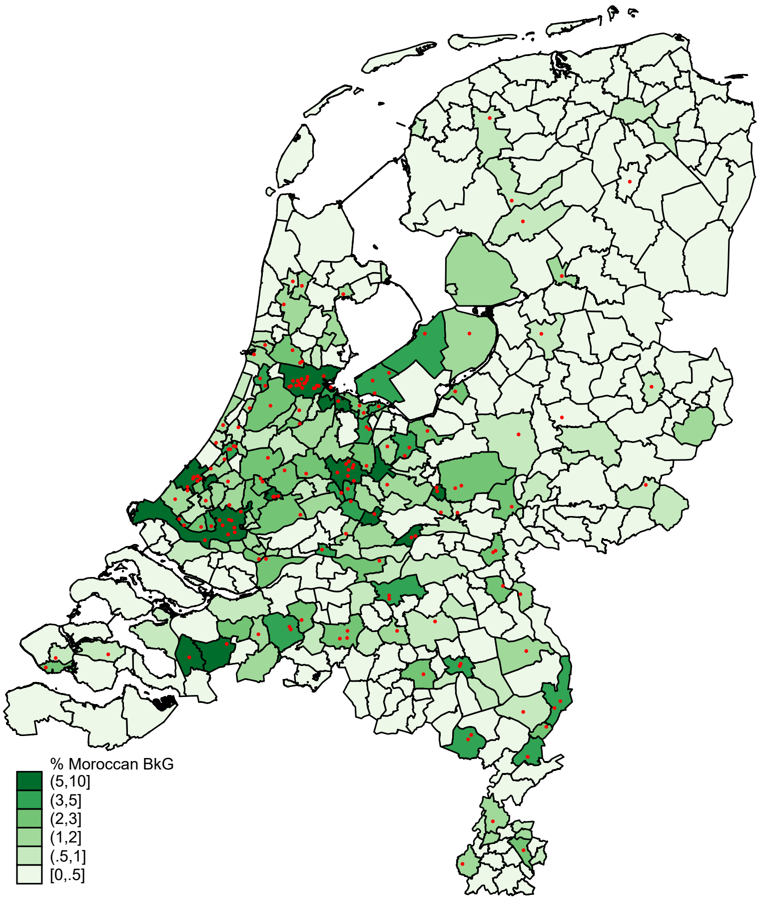
Wittbrodt, Matthew T and Melinda Millard-Stafford, "Dehydration impairs cognitive performance: a meta-analysis," *Med Sci Sports Exerc*, 2018, *50* (11), 2360–2368.

Yasin, Wardah Mohd, Muhammad Muzaffar Ali Khan Khattak, Nik Mazlan Mamat, and Wan Azdie Mohd Abu Bakar, "Does religious fasting affect cognitive performance?," *Nutrition & Food Science*, 2013.

Appendix

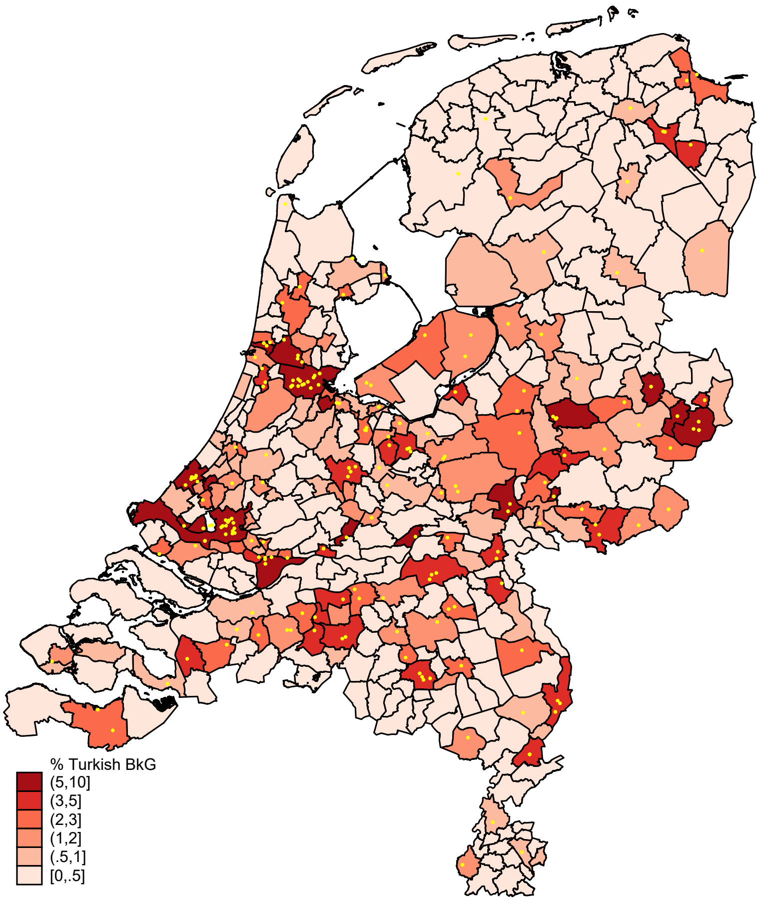
A Additional Tables and Figures

Figure A1: The share of residents with a Moroccan migration background per municipality and the location of mosques in the Netherlands



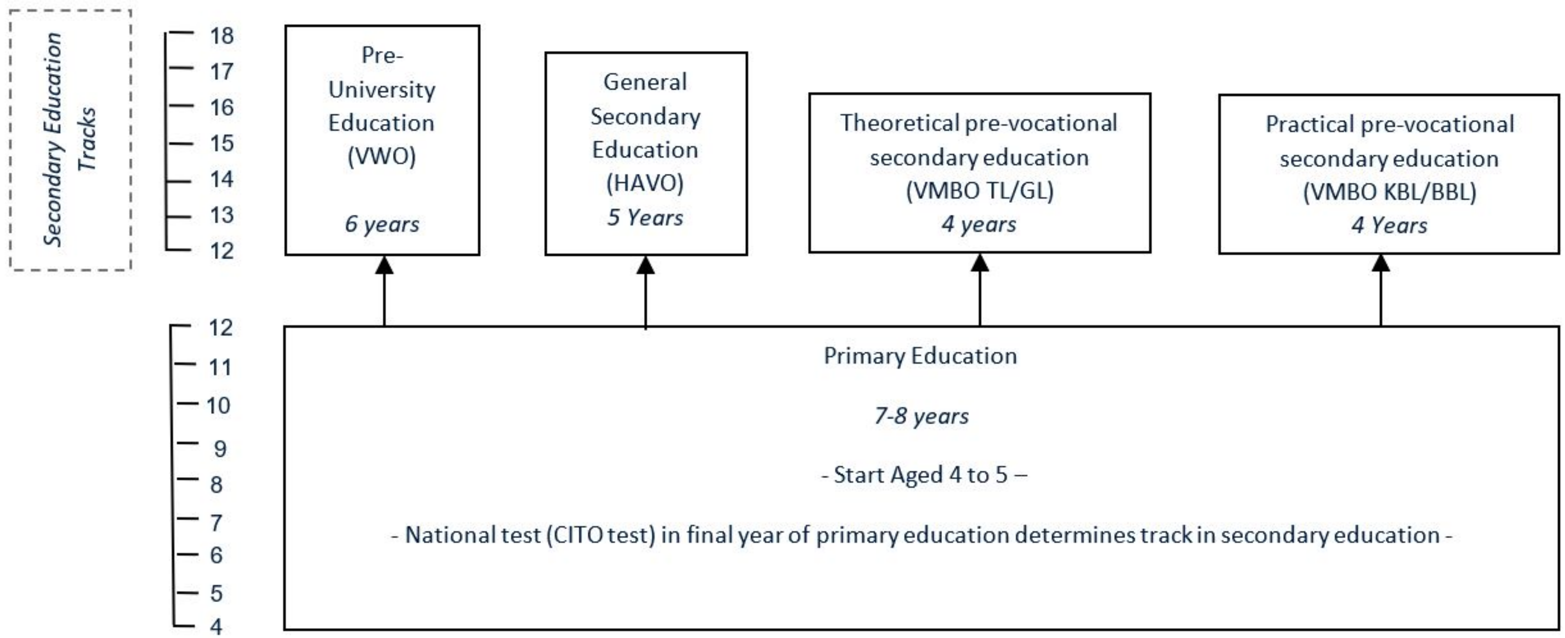
Notes: Municipality level proportion of people with a Moroccan migration background in 2017 in the Netherlands. The locations of mosques where the language of the sermon is Moroccan are indicated with red dots. Source: CBS Statline, statline.cbs.nl; Moskeewijzer, https://moskeewijzer.nl/.

Figure A2: The share of municipalities with a Turkish migration background and the location of mosques in the Netherlands



Notes: Municipality level proportion of people with a Turkish migration background in 2017 in the Netherlands. The locations of mosques where the language of the sermon is Turkish are indicated with yellow dots. Source: CBS Statline, statline.cbs.nl; Moskeewijzer, https://moskeewijzer.nl/.

Table A1: Dutch education system



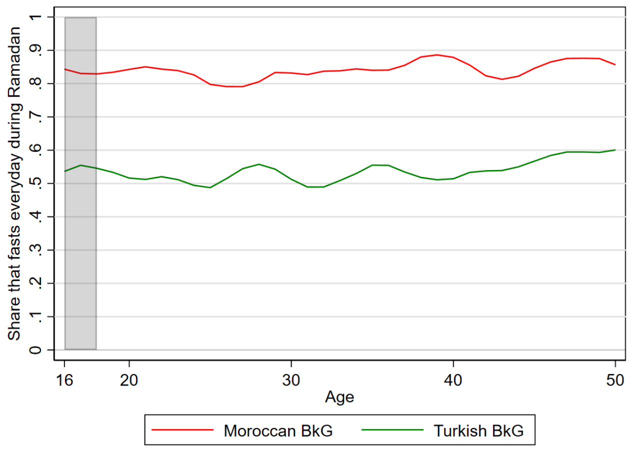
Notes: Detailed graphical overview of the education system from primary education to higher education in the Netherlands.

Table A2: Compliance with measures of Muslim religiosity by migration background

	SIM Respondents	
	Moroccan Migration BkG	Turkish Migration BkG
	(1)	(2)
Prays At Least Every Day	0.796 (.403)	0.376 (.484)
Visits Mosque At Least Every Week	0.367 (.482)	0.344 (.475)
Eats Halal Every Day	0.877 (.328)	0.732 (.443)
Finds Religion Important	0.901 (.299)	0.801 (.399)
Islam Rules **Should be Followed Completely**	0.601 (.490)	0.501 (.500)
Observations	2,677	2,910

Notes: Self-reported compliance with measures of Muslim religiosity based on SIM (*Survey Integratie Minderheden*) data of respondents with a Moroccan migration background on the left and Turkish migration background on the right. Source: SIM data in 2006, 2011, 2015 and 2020.

Figure A3: Fasting behavior of SIM respondents with a Moroccan or Turkish BkG by age (SIM survey 2006-2020)



Notes: Graphical representation of the average share of respondents with a Moroccan migration background (in red) or Turkish migration background (in green) that report to fast everyday during Ramadan by age at the time of the survey. Source: SIM (*Survey Integratie Minderheden*) data on 2006, 2011, 2015 and 2020.

Table A3: Characteristics of SIM Survey Respondents and Exam Takers Sample

	Moroccan Migration BkG		Turkish Migration BkG	
	SIM	Exam takers	SIM	Exam takers
	(1)	(2)	(3)	(4)
Age	38.180	16.496	38.328	16.544
	(15.317)	(0.976)	(15.225)	(0.983)
Ratio under 45 neighbourhood	0.615	0.608	0.609	0.597
	(0.089)	(0.090)	(0.091)	(0.090)
Birth over fertility neighbourhood	0.028	0.028	0.027	0.027
	(0.007)	(0.008)	(0.008)	(0.009)
Average household size neighbourhood	2.107	2.165	2.145	2.188
	(0.303)	(0.316)	(0.321)	(0.310)
Share assistance benefits neighbourhood	0.059	0.058	0.058	0.056
	(0.034)	(0.034)	(0.035)	(0.035)
Share Moroccan BkG neighbourhood	0.107	0.114	0.076	0.074
	(0.091)	(0.096)	(0.078)	(0.082)
Share Turkish BkG neighbourhood	0.078	0.079	0.111	0.102
	(0.075)	(0.073)	(0.092)	(0.085)
Distance to Moroccan mosque	2,068.481	1,975.036	3,393.643	3,988.605
	(3,621.228)	(3,543.813)	(5,882.818)	(6,447.188)
Distance to Turkish Mosque	2,203.873	2,270.903	1,485.046	1,741.672
	(3,001.709)	(2,881.584)	(1,889.499)	(2,140.501)
Observations	2,104	30,413	2,208	29,546

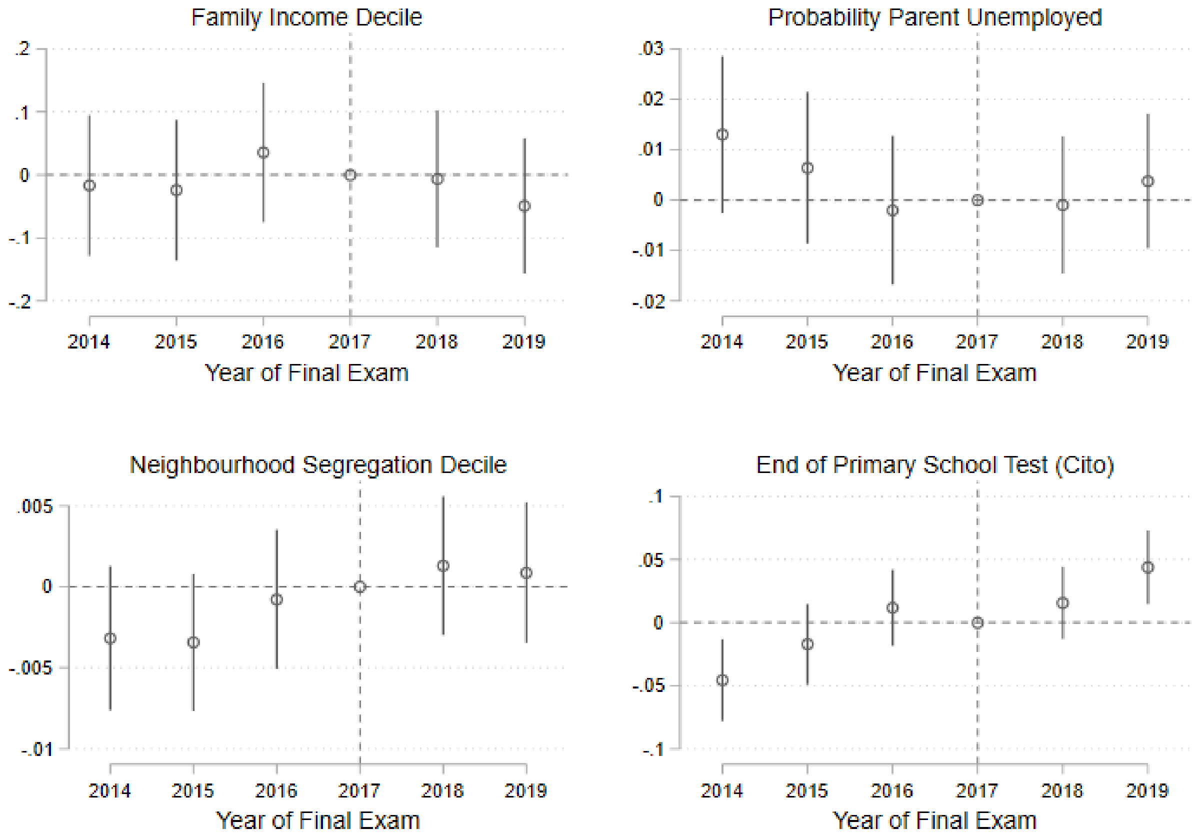
Notes: Balance table of characteristics of the sample of SIM (Survey Integratie Minderheden) respondents (in 2006-2020) with a Turkish or Moroccan migration background (that we use to train our ML model to predict Ramadan compliance with) and the sample of exam takers (in 2014-2019) with a Turkish or Moroccan migration background that we use in our analysis.

Table A4: Composition of treated group

	Students with	
	Moroccan Migration BkG	Turkish Migration BkG
	(1)	(2)
Moroc BkG	1.000	0.000
High Ramadan Proba	0.998	0.250
Continuous Ramadan Proba	0.883	0.565
Observations	30,413	29,546

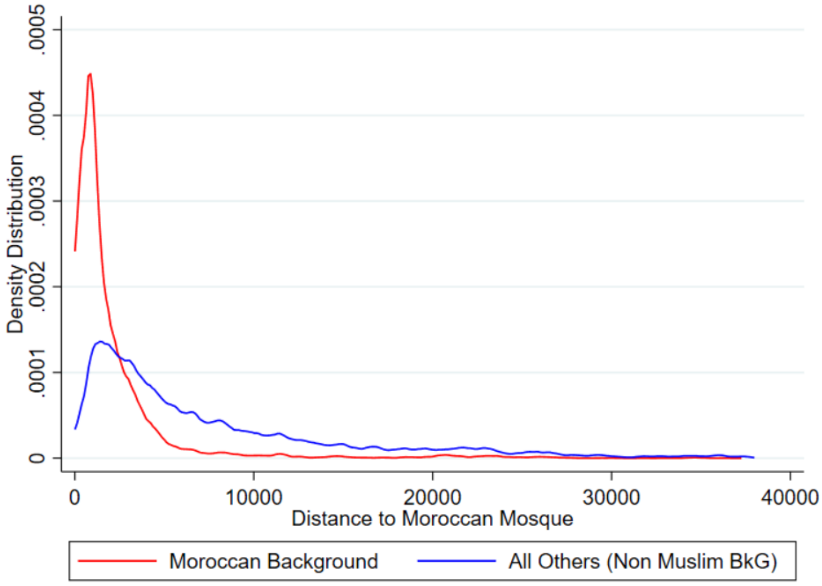
Notes: Share of students in our sample by migration background that are ‘treated’ (compliant to the Ramadan fast) by treated specification.

Figure A4: DiD with student characteristics as outcomes and students with high Ramadan probability as treated group



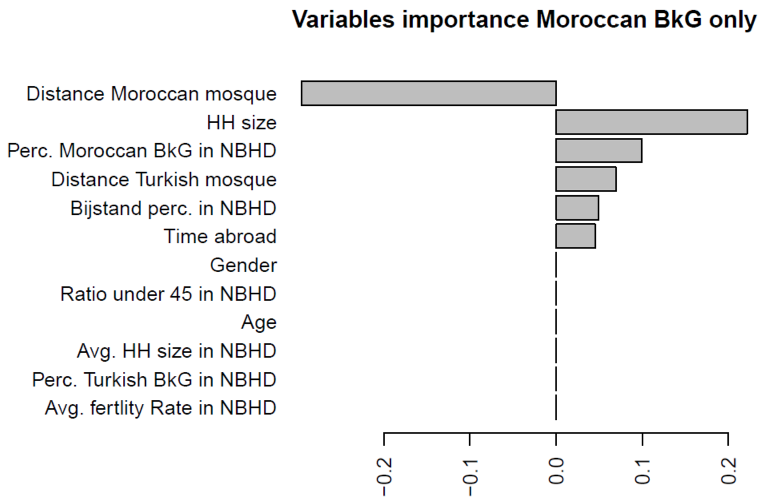
Notes: The figures plot the point estimates and 99% confidence intervals for the coefficients (β_t) that are estimated using our main specification (equation (1)). Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. Family income is divided into deciles, probability parent unemployed is measured as having at least one parent being unemployed, neighbourhood segregation is measured as the share of people with a Moroccan or Turkish migration background in your neighbourhood and divided into deciles, the end of primary school test (CITO) is standardized by CITO year with mean zero and standard deviation one. Controls included are: gender, age at exam, end of primary school test score (standardized by year), continuous exam grade (standardized by track and exam year), moved up from lower track or failed the final exams last year, number of siblings, a dummy for being Moroccan or Turkish, time spent abroad, parental income, unemployment or illness benefits, and on the neighbourhood level: the share receiving assistance benefits, the average household size, the ratio under 45 years old, fertility rate, share with Turkish migration background, share with Moroccan migration background. We exclude a control variable if it overlaps with the outcome variable. School-subtrack and exam year fixed effects are included. Robust standard errors are clustered at the school-track-subtrack-year level. The horizontal dashed line marks the zero or no effect.

Figure A5: Distance to Moroccan mosque



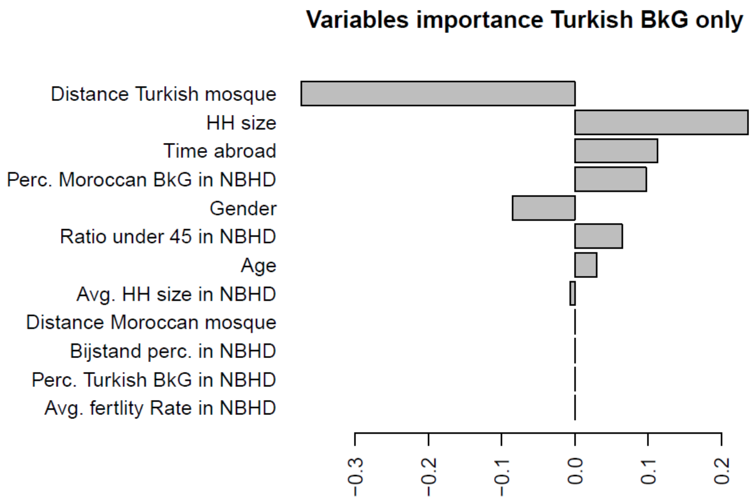
Notes: Density distribution of the distance from the neighbourhood of the student to the neighbourhood of the Moroccan mosque for students with a Moroccan migration background (red) and students with a non-Muslim migration background or no migration background (blue). Source: Moskeewijzer, https://moskeewijzer.nl/.

Figure A6: Penalized Logit model coefficients: Moroccan BkG



Notes: Importance of variables for prediction of Ramadan compliance probability as predicted with a penalized logit ML model including only SIM respondents with a Moroccan migration background. Variables used in the model are the share of the respondent's neighbourhood having a Moroccan neighbourhood or Turkish neighbourhood, age of the respondent, time spent abroad, gender of the respondent, average household size in the respondent's neighbourhood, distance to Moroccan or Turkish mosque, share of the neighbourhood using welfare (*bijstand*), fertility in the neighbourhood and the share of the respondent's neighbourhood below 45 years old.

Figure A7: Penalized Logit model coefficients: Turkish BkG



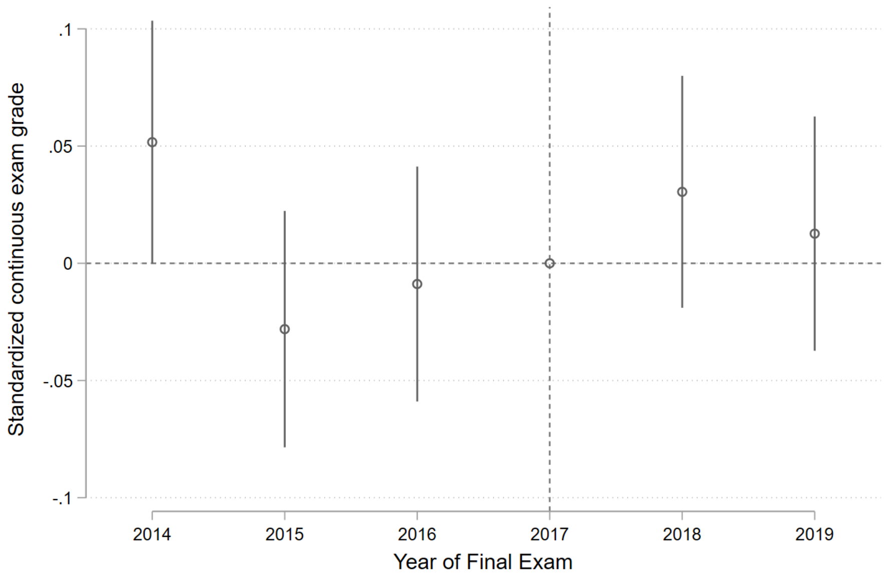
Notes: Importance of variables for prediction of Ramadan compliance probability as predicted with a penalized logit ML model including only SIM respondents with a Turkish migration background. Variables used in the model are the share of the respondent's neighbourhood having a Moroccan neighbourhood or Turkish neighbourhood, age of the respondent, time spent abroad, gender of the respondent, average household size in the respondent's neighbourhood, distance to Moroccan or Turkish mosque, share of the neighbourhood using welfare (*bijstand*), fertility in the neighbourhood and the share of the respondent's neighbourhood below 45 years old.

Table A5: Impact of Ramadan on education outcomes of students with a high Ramadan probability

	Secondary School Outcomes					
	High Ramadan Proba		**Moroc BkG**		**Cont. Ramadan Proba**	
	Final	Pass	Final	Pass	Final	Pass
	(1)	(2)	(3)	(4)	(5)	(6)
Ramadan Exam Years	- 0.058***	- 0.011***	- 0.059***	- 0.009**	- 0.052***	- 0.011**
***Treated**	(.010)	(.004)	(.011)	(.004)	(.011)	(.004)
School-Subtrack f.e.	Yes	Yes	Yes	Yes	Yes	Yes
Exam year f.e.	Yes	Yes	Yes	Yes	Yes	Yes
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean outcome before	6.10	86.8%	6.13	87.4%	6.08	86.0%
Share treated	4.7%	4.7%	3.9%	3.9%	5.5%	5.5%
Observations	776,219	776,284	746,697	746,742	776,219	776,284

Notes: The table shows point estimates for the coefficient β estimated using equation (2) for the 3 different treatment group specifications. Final is the average of the central exam grades standardized with mean zero and standard deviation one by track and final exam year. Pass is a dummy indicating that a student passed the final exams (including retakes) and graduated from secondary education. The difference in number of observations between the analyses with Final and Pass as outcomes can be explained by students not taking the exam during the first period, but during the first or second retake due to e.g. illness. Mean outcome before is the average outcome for treated students in the pre-treatment years 2014 to 2017. Controls included are: gender, age at exam, end of primary school test score (standardized by year), continuous exam grade (standardized by track and exam year), moved up from lower track or failed the final exams last year, number of siblings, a dummy for being Moroccan or Turkish, time spent abroad, parental income, unemployment or illness benefits, and on the neighbourhood level: the share receiving assistance benefits, the average household size, the ratio under 45 years old, fertility rate, share with Turkish migration background, share with Moroccan migration background. Robust standard errors are clustered at the school-track-subtrack-year level in parentheses. $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A8: Impact of Ramadan on continuous exam grade of students with a high Ramadan probability



Notes: The figures plot the point estimates and 99% confidence intervals for the coefficients (β_t) that are estimated using our main specification (equation (1)). Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. Continuous exam grade is measured as is the average continuous exam grade standardized with mean zero and standard deviation one by track and exam year. The continuous exams took place before the central final exams and hence did not overlap with Ramadan, this test serves as a robustness check and we should not observe any significant impact of Ramadan on this outcome. All controls are included, as in Table A5. School-subtrack and exam year fixed effects are included. Robust standard errors are clustered at the school-track-subtrack-year level. The horizontal dashed line marks the zero or no effect.

Table A6: Impact of Ramadan on secondary school outcomes controlling for sibling and family fixed effects

	Secondary School Outcomes			
	Incl. Sibling f.e.		**Incl. Family f.e.**	
	Final	Pass	Final	Pass
	(1)	(2)	(3)	(4)
Ramadan Exam Years	- 0.059***	- 0.014**	- 0.068***	- 0.018***
*** High Ramadan Proba**	(.014)	(.007)	(.013)	(.005)
School-Subtrack f.e.	Yes	Yes	Yes	Yes
Exam year f.e.	Yes	Yes	Yes	Yes
Sibling f.e.	Yes	Yes	No	No
Family f.e.	No	No	Yes	Yes
All controls	Yes	Yes	Yes	Yes
Mean outcome before	6.03	81.8%	6.07	84.9%
Share treated	4.6%	4.6%	5.4%	5.4%
Observations	408,017	408,060	493,477	493,528

Notes: The table shows point estimates for the coefficient β estimated using equation (2). Treated students in this table are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. In Column 1 and 2 we include sibling fixed effects, in Column 3 and 4 family fixed effects (comparing students not only to their siblings but also to their cousins). Final is the average of the central exam grades standardized with mean zero and standard deviation one by track and final exam year. Pass is a dummy indicating that a student passed the final exams and graduated from secondary education. Mean outcome before is the average outcome for treated students in the pre-treatment years 2014 to 2017. All controls are included, as in Table A5. Robust standard errors are clustered at the school-track-subtrack-year level in parentheses. $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

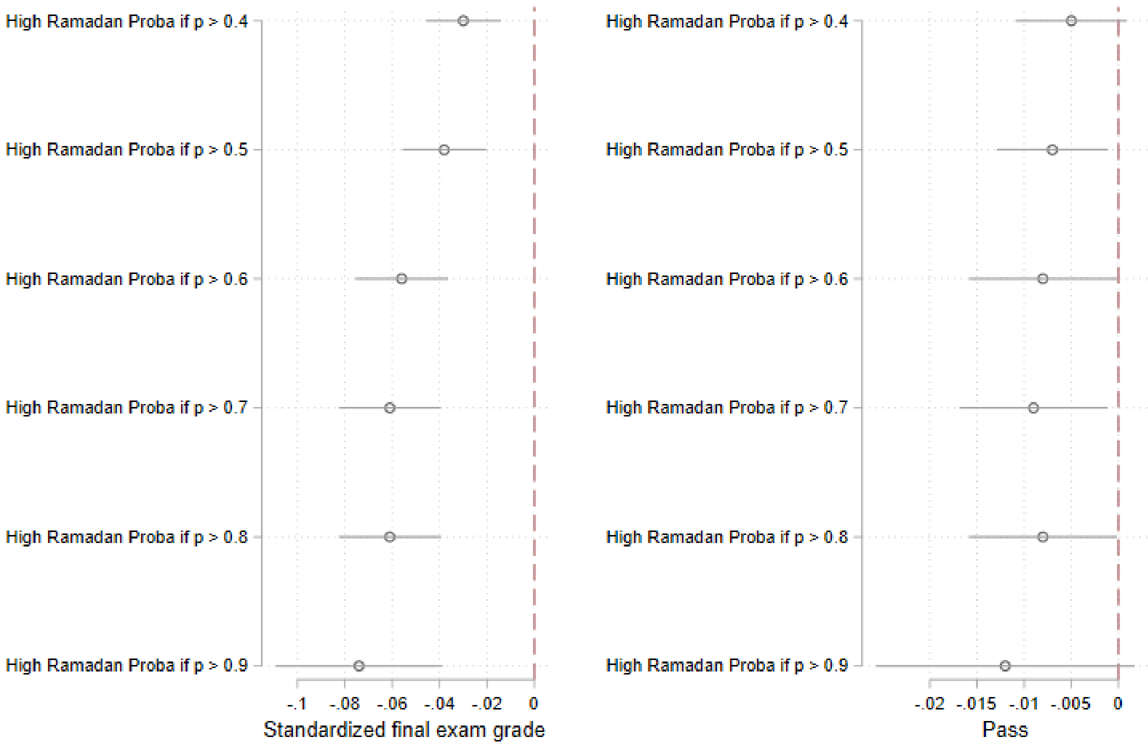
Table A7: Placebo impact of Ramadan on secondary school outcomes

	Secondary School Outcomes	
	Final	Pass
	(1)	(2)
Panel A: Ramadan Probability Based on Within-Turk ML Model		
Ramadan Exam Years	- 0.001	- 0.000
***High Ramadan Proba**	(.008)	(.003)
Mean exam grade	6.02	84.7%
Share treated	4.9%	4.9%
Observations	707,501	707,561
School-Subtrack f.e.	Yes	Yes
Exam year f.e.	Yes	Yes
All controls	Yes	Yes
Panel B: Ramadan Probability Based on Within-Moroc ML Model		
Ramadan Exam Years	0.013	- 0.000
***High Ramadan Proba**	(.008)	(.003)
Mean exam grade	6.13	87.4%
Share treated	4.9%	4.9%
Observations	707,501	707,561
School-Subtrack f.e.	Yes	Yes
Exam year f.e.	Yes	Yes
All controls	Yes	Yes

Notes: The table shows point estimates for the coefficient β estimated using equation (2). In Panel A, 'treated' students are defined as students with a non-Muslim migration background or without a migration background that are classified as high Ramadan probability according to predictions made by our within-Turk ML model. In Panel B, 'treated' students are defined as students with a non-Muslim migration background or without a migration background that are classified as high Ramadan probability according to predictions made by our within-Moroccan ML model. Final is the average of the central exam grades standardized with mean zero and standard deviation one by track and final exam year. Pass is a dummy indicating that a student passed the final exams and graduated from secondary education. Mean outcome before is the average outcome for treated students in the pre-treatment years 2014 to 2017. All controls are included, as in Table 1. Robust standard errors are clustered at the school-track-subtrack-year level in parentheses. $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A9: Impact of Ramadan on secondary education outcomes of students with a high Ramadan probability at different thresholds

A: Final exam grade **B:** Pass



Notes: The figures plot the point estimates and 99% confidence intervals for the coefficients (β_t) that are estimated using our main specification (equation (1)). Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. We define 'high' Ramadan probability based on multiple thresholds, ranging from a probability of 0.4 or higher to a probability of 0.9 or higher. The exams in years 2014 to 2017 did not overlap with Ramadan (and hence we would not expect an effect here), while exams in 2018 and 2019 did overlap with Ramadan. Final exam grade is measured as the average of the central exam grades standardized with mean zero and standard deviation one by track and exam year. Probability of overall pass is measured by a dummy that indicates whether the student passed the final exams (and graduated from secondary education). All controls are included, as in Table A5. School-subtrack and exam year fixed effects are included. Robust standard errors are clustered at the school-track-subtrack-year level. The horizontal dashed line marks the zero or no effect.

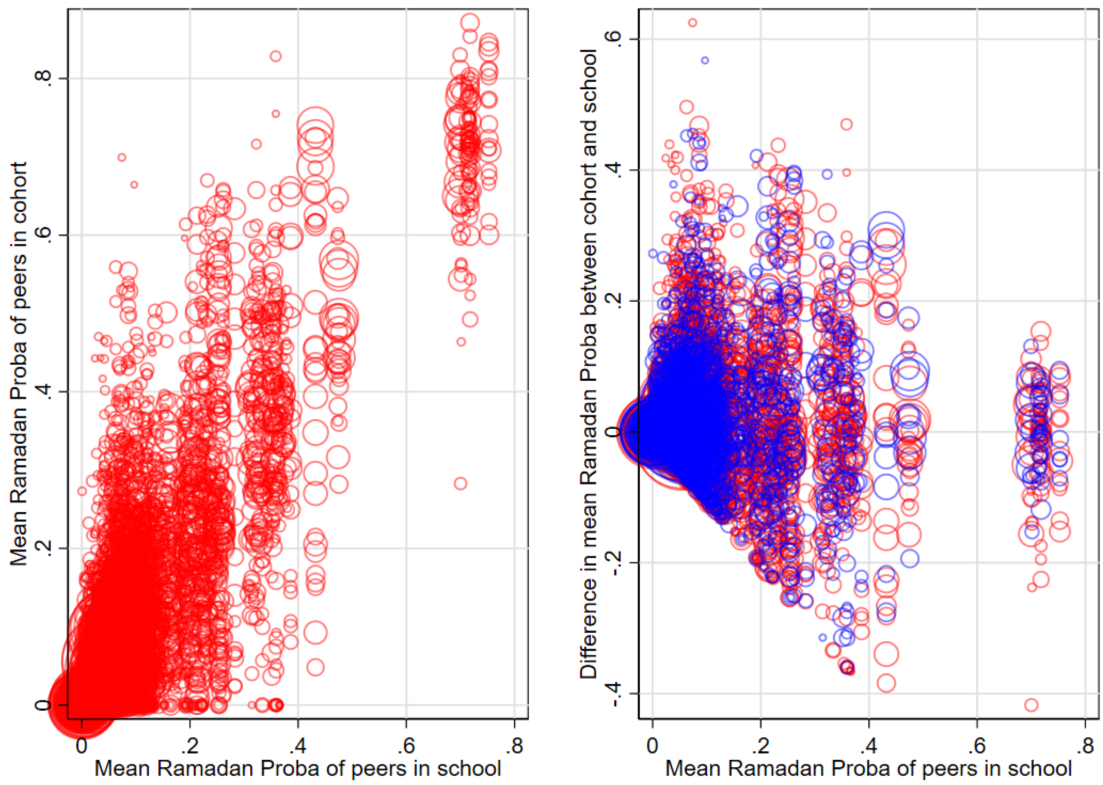
Table A8: Impact of Ramadan at subject level

	Secondary School Outcomes		
	High Ramadan Proba	**Moroc BkG**	**Cont. Ramadan Proba**
	Final	Final	Final
	(1)	(2)	(3)
Ramadan Exam Years	- 0.032***	- 0.037***	- 0.028***
***Treated**	(.006)	(.006)	(.007)
School-Track-Subject f.e.	Yes	Yes	Yes
Exam year f.e.	Yes	Yes	Yes
Treated-Subject-Afternoon-Two day f.e.	Yes	Yes	Yes
All controls	Yes	Yes	Yes
Mean outcome before	6.13	6.17	6.11
Share treated	4.13%	3.40%	4.83%
Observations	4,731,843	4,573,531	4,731,846

Notes: The table shows point estimates for the coefficient β estimated using equation (3) for the 3 different treatment group specifications, based on the subject level data. The dependent variable is the grade obtained in the exam (after retakes) standardized at subject-track-year level. All controls are included, as in Table A5. Mean outcome before is the average outcome for treated students in the pre-treatment years 2014 to 2017. Robust standard errors are clustered at the school-track-subject-year level in parentheses. $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A10: Variation in Ramadan compliance probability across and within schools

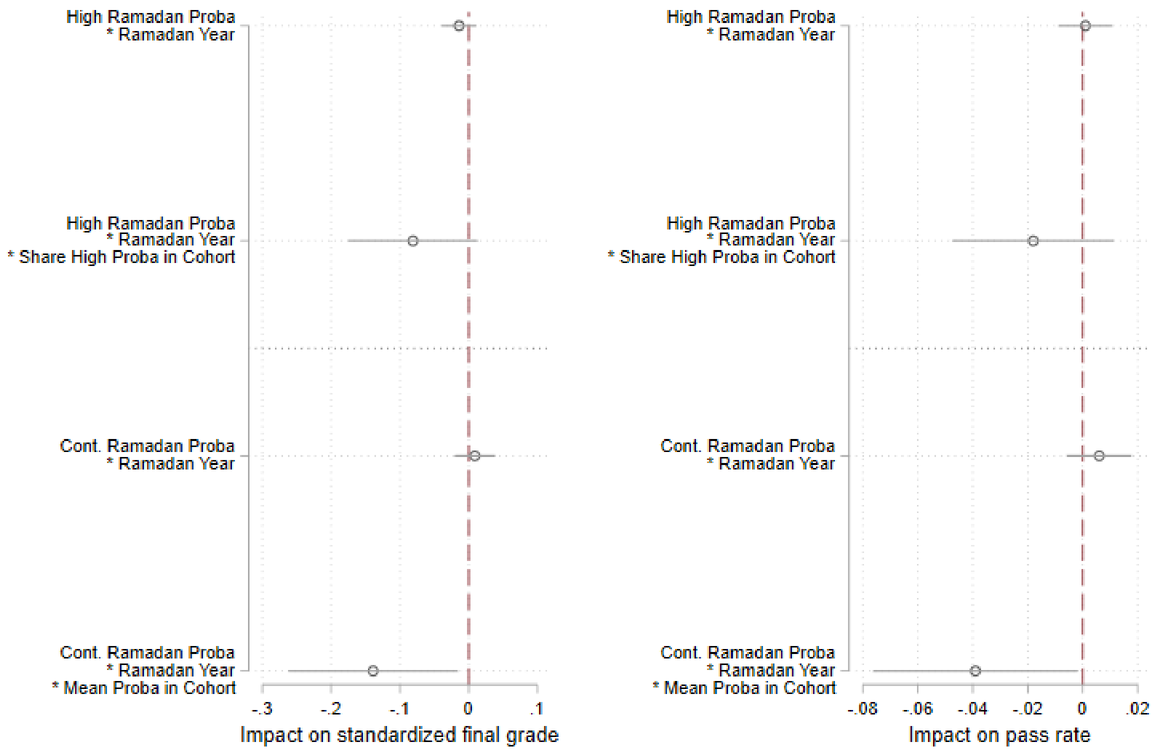
A: Probability in schools and cohorts **B:** Difference in probability within schools



Notes: The left figure plots the mean Ramadan compliance probability in a student's school on the x-axis (leave-one-out-mean in same school) and the mean proability in a student's cohort on the y-axis (leave-one-out-mean in same subtrack-track-school-exam year) for the years 2014 to 2019. The size of the circles is determined by size of cohort. The right figure plots the (leave-one-out) mean continuous Ramadan probability of peers in school (leave-one-out-mean in same school) on the x-axis and the difference between the (leave-one-out-mean) continuous Ramadan probability in the cohort and in the school on the y-axis. The size of the circles is determined by size of cohort. The red circles define the cohorts that took their final exams in 2014-2017, while the blue circles define the cohorts that took their final exams in 2018-2019 (when Ramadan overlapped with the final exams).

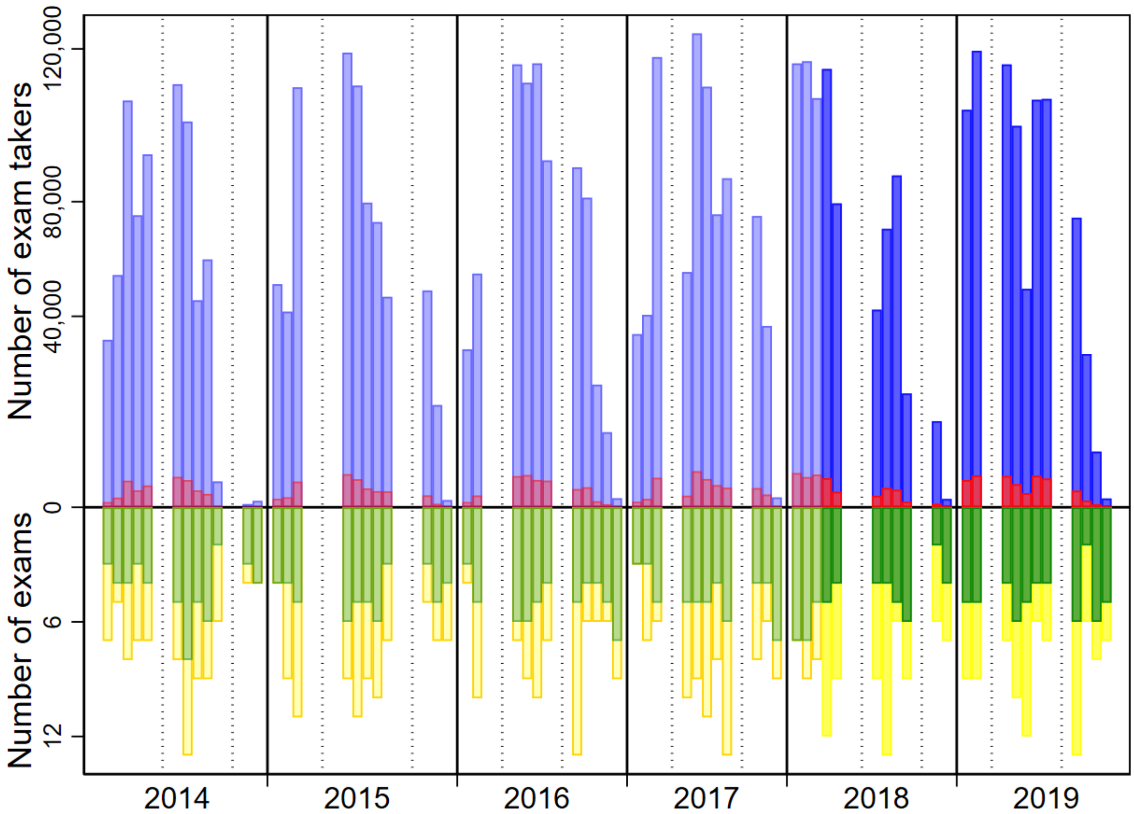
Figure A11: Heterogeneity in impact of Ramadan on secondary education outcomes by Ramadan probability of peers

A: Standardized final grade **B:** Pass rate



Notes: The figures plot the point estimates and 95% confidence intervals for the coefficient β that is estimated using equation (2). In the upper panels, treated students in are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. In the bottom panels, treated students are defined continuously by the Ramadan probability. We interact β with the average share of a student's peers that also have a high Ramadan probability (in the upper panel) or the average Ramadan probability of a student's peers (in the bottom panel). Final exam grade is measured as the average of the central exam grades standardized with mean zero and standard deviation one by track and exam year. Probability of overall pass is measured by a dummy that indicates whether the student passed the final exams (and graduated from secondary education). All controls are included, as in Table A5. School-subtrack and exam year fixed effects are included. Robust standard errors are clustered at the school-track-subtrack-year level. The horizontal dashed line marks the zero or no effect.

Figure A12: Number of exam takers and exams by timing and migration background



Notes: The figure plots on the upper panel the number of exam takers, with in red the number of students with a Moroccan or Turkish migration background and in blue the total number of exam takers on that day. In the bottom panel, it displays the number of exams per day, with in green the number of exams that take place in the afternoon and in yellow the number of exams that take place in the morning. The darker colors indicate exam days that took place during Ramadan.

Table A9: Impact of Ramadan on education outcomes of students with a high Ramadan probability

	Secondary School Outcomes
	Final grade
	(1)
DiD	- 0.040***
	(.014)
DiD *Afternoon	0.010
	(.015)
School-Track-Subject f.e.	Yes
Treated-Subject-Afternoon-Two day f.e.	Yes
Exam year f.e.	Yes
All controls	Yes
Mean outcome before	6.13
Share treated	4.6%
Observations	4,731,843

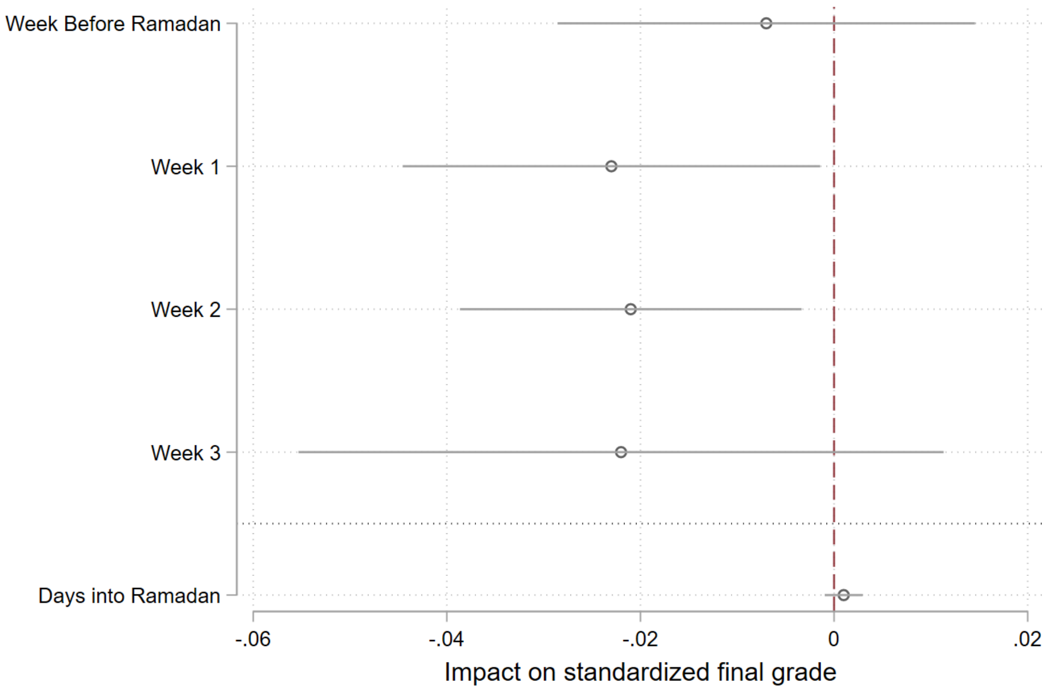
Notes: The table shows point estimates for the coefficient β and for the interaction term between β and a dummy that takes value 1 if the exam took place in the afternoon estimated using equation (3), based on the subject level data. Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. The dependent variable is the grade obtained in an exam (after retakes) standardized at subject-track-year level. Mean outcome before is the average outcome for treated students in the pre-treatment years 2014 to 2017. All controls are included, as in Table A5. Robust standard errors are clustered at the school-track-subject-year level in parentheses. $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: DiD estimates by type of subject using high Ramadan probability as treated

	Secondary School Outcomes		
	Memory-based	Languages	Math
	Final	Final	Final
	(1)	(2)	(3)
Ramadan Exam Years	- 0.042***	- 0.024***	- 0.089***
*** High Ramadan Proba**	(.014)	(.010)	(.025)
School-Subtrack f.e.	Yes	Yes	Yes
Exam year f.e.	Yes	Yes	Yes
All controls	Yes	Yes	Yes
Mean exam grade	6.04	6.19	6.21
Share treated	4.1%	4.5%	4.2%
Observations	867,603	1,930,649	702,153
No. students	583,126	729,274	659,989

Notes: The table shows point estimates for the coefficient β estimated using equation (3), for three groups of exams: languages (reading comprehension tests in Dutch, English, German and French), more memory-based exams (history, geography and biology), and math. The dependent variable is the grade obtained in an exam (after retakes) standardized at subject-track-year level. We include school-track-exam subject and treatment type-exam subject-afternoon-two day exam fixed effects. Robust standard errors are clustered at the school-subtrack-year level in parentheses. All controls are included, as in Table A5. $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A13: Effect on final grade by Ramadan week



Notes: This figure plots the point estimates and 99% confidence intervals for the coefficients estimated using equation (5) and equation (6). The coefficients *Week* 1, *Week* 2 and *Week* 3 measure the impact of Ramadan exposure during the first, second or third week of Ramadan on the standardized result of the central exam, using data at subject level. *Week Before Ramdan* refers to the week before the beginning of Ramadan (which is observed only in 2018). *Days into Ramadan* captures the marginal effect of taking exams after an additional day of Ramadan exposure, as measured by the interaction term β_2 in equation (6). Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. All controls are included, as in Table A5. We include school-track-exam subject and treatment type-exam subject-afternoon-two day exam fixed effects. Robust standard errors are clustered at the school-track-subject-year level in parentheses. The horizontal dashed line marks the zero or no effect.

Table A11: DiD estimates by family size using high Ramadan probability as treated

	Secondary School Outcomes			
	Small family Final (1)	Big family Final (2)	Small family Pass (3)	Big family Pass (4)
Ramadan Exam Years	-0.054***	- 0.065***	- 0.010**	- 0.012**
*** High Ramadan Proba**	(.013)	(.013)	(.005)	(.005)
School-Subtrack f.e.	Yes	Yes	Yes	Yes
Exam year f.e.	Yes	Yes	Yes	Yes
Mean exam grade	6.12	6.07	87.2%	86.3%
Share treated	4.5%	5.0%	4.5%	5.0%
All controls	Yes	Yes	Yes	Yes
Observations	423,221	352,488	423,261	352,513

Notes: The table shows point estimates for the coefficient β estimated using equation (2). Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. We split our sample between students with a small (below-median) family and students with a big (above-median) family. Family size is measured as extended family size, including siblings, parents, grandparents, aunts, uncles and cousins. Final is the average of the central exam grades standardized with mean zero and standard deviation one by track and final exam year. Pass is a dummy indicating that a student passed the final exams and graduated from secondary education. Mean outcome before is the average outcome for treated students in the pre-treatment years 2014 to 2017. All controls are included, as in Table A5. Robust standard errors are clustered at the school-track-subtrack-year level in parentheses. $^*p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Impact of Ramadan on Longer-Term Outcomes

	Longer Term Outcomes		
	Grade retention (1)	Adult education (2)	Drop out (3)
Ramadan Exam Years	0.007***	0.004	0.004
***High Ramadan Proba**	(.003)	(.002)	(.004)
School-Subtrack f.e.	Yes	Yes	Yes
Exam year f.e.	Yes	Yes	Yes
All controls	Yes	Yes	Yes
Share treated	4.70%	4.70%	4.70%
Mean outcome	5.60%	4.94%	5.01%
Observations	776,284	776,284	776,284

Notes: The table shows point estimates for the coefficient β estimated using equation (2). Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. Grade retention is a dummy the takes value one if the student repeats the final year in the year after the final exam. Adult education is a dummy that takes value one if the student goes into adult education in the year after the final exam. Drop out is a dummy that takes value one if the student is not registered for any education in the year after the final exam. Mean outcome before is the average outcome for treated students in the pre-treatment years 2014 to 2017. All controls are included, as in Table A5. Robust standard errors are clustered at the school-track-subtrack-year level in parentheses. $^*p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Impact of Ramadan on Longer-Term Outcomes by Age

	Longer Term Outcomes					
	Younger than 18			**Older than 18**		
	Grade retention	Adult education	Drop out	Grade retention	Adult education	Drop out
	(1)	(2)	(3)	(4)	(5)	(6)
Ramadan Exam Years	0.006**	0.004*	0.001	0.008	0.000	0.017**
***High Ramadan Proba**	(.003)	(.002)	(.004)	(.005)	(.008)	(.008)
School-Subtrack f.e.	Yes	Yes	Yes	Yes	Yes	Yes
Exam year f.e.	Yes	Yes	Yes	Yes	Yes	Yes
All controls	Yes	Yes	Yes	Yes	Yes	Yes
Share treated	5.32%	5.32%	5.32%	3.34%	3.34%	3.34%
Mean outcome	5.86%	2.18%	4.26%	4.68%	14.5%	7.59%
Observations	536,927	536,927	536,927	238,111	238,111	238,111

Notes: The table shows point estimates for the coefficient β estimated using equation (2). Treated students in this figure are students with a Moroccan or Turkish migration background that have a high Ramadan probability, as predicted by our machine learning model. We split our sample into students that are younger or older than 18 years old at the 1st of September after the final exam. Grade retention is a dummy the takes value one if the student repeats the final year in the year after the final exam. Adult education is a dummy that takes value one if the student goes into adult education in the year after the final exam. Drop out is a dummy that takes value one if the student does not register for any education in the year after the final exam. Mean outcome before is the average outcome for treated students in the pre-treatment years 2014 to 2017. All controls are included, as in Table A5. Robust standard errors are clustered at the school-track-subtrack-year level in parentheses. $^{*}p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Data Appendix

B.1 Construction of sample

We use administrative data from Statistics Netherlands which contains information on all individuals who are registered in a municipality by 1995.[35] We start with the registry of exam takers (*EXAMVOTAB*), select all exam takers between 2014 and 2019 (where exam year is the school year plus one, because students start the final school year in September of i.e. 2013 and take the exam in May 2014) and drop entries for which the final grade is missing,[36] which gives us a sample of 983,837 individuals (of whom some take more than one exam, so we have a total of 1,048,400 observations). We match them to the track and sub-track they are in using the education reference book (*OPLEIDINGSNRREFV26*).

B.2 Construction of variables

We match the exam takers in our sample to their demographic characteristics (country of birth, country of origin, country of birth of parents, gender, birth date and migration generation) in the registry of persons (*GBAPERSOONTAB*). The parent-child registry (*KINDOUDERTAB*) is used to link the exam takers in our sample to their parents such that we can determine outcomes for the parents. We match parents to their demographic characteristics (country of birth, country of origin, birth date and country of birth of parents) in the registry of persons. We also use parents to identify the rest of the (extended) family to use in a family fixed effects analysis. We first use the identifying number of the mother to identify an exam taker's siblings, then we use it to identify the grandparents. We use the grandparents to identify aunts and uncles and then link those aunts and uncles to their children to identify cousins. We also do this for the father to identify the father's side of the family. Due to the problem of overlapping families, the family fixed effects are based on the mother of the father or, if this is not available, the mother of the mother, or, if this is not available the mother (only siblings) or father (only siblings). Treated students suffer from more limitations than control students, because in order to identify the extended family we need to identify the grandparents, which are often not available in the

[35]The administrative data from Statistics Netherlands is available at a remote-access facility after signing a confidentiality agreement.

[36]The number of missing observations in the exam years 2016 and 2019 is unusually high, but these entries refer to students that were already pre-registered in the exam grade dataset in the year before the final year for policy reasons, and not to students in the final year taking the exams.

administrative data for students with a migraton background. Therefore, while for students in our control group we can identify for 95% the extended family (and compare the exam taker not only with siblings but also with cousins), we can only do this for 77% of the students in our treatment group.

Country of origin is defined as the country of birth if a person is born abroad or, if born in the Netherlands, the country of birth of the mother, or, if the mother is also born in the Netherlands, the country of birth of the father. We construct a new variable that also considers third generation migrants, by replacing the country of origin with the country of origin of the mother if the country of origin of the exam taker is the Netherlands, or the country of origin of the father if the country of origin of the mother is the Netherlands. We classify exam takers as Muslim if this new country of origin is a Muslim majority country (see Table A14).[37]

We use birth date from the registry of persons to define the age at exam (using the 1st of May of the exam year). We add information from the registry with the information on the individual's obtained test score at the standardized test at the end of primary education (*CITOTAB*). This registry is only available from 2006 to 2019, so given that exam takers took the CITO test 4 to 6 years before their final exam the exam takers to the CITOTAB 2006-2019. We can match 75.4% of the exam takers to their obtained CITO test score. This leaves us with a sample of 841,347 observations (789,259 unique students).

We aggregate the new country of origin variable at the school level to create a variable with the proportion of Moroccans in each school and track. We then divide this variable in four quantiles.

We then match the parents to data on labour income (*SECMWERKNDGAMNDBEDRAG-BUS*), unemployment benefits (*SECMWERKLMNDBEDRAGBUS*), income from profits (*SECM-ZLFMNDBEDRAGBUS*), and illness benefits (*SECMZIEKTEAOMNDBEDRAGBUS*). We aggregate the income data to yearly income data and only keep data on income in the year before the exam was taken. We generate a dummy variable that equals one if one of the parents receives unemployment benefits and do the same for illness benefits. We sum all sources of income to create a variable on total parental income and divide this in year-specific deciles.

In our analysis we make use of different spatial data: we exploit several characteristics of the students' neighbourhood of residence and we compute (approximate) distances between the students' residences and the closest mosque and primary school. In order to link students with

[37](We are excluding Indonesia, even though this is a Muslim majority country, as a very selected group of (mostly non$_{M}uslim)migrantscamefromIndonesiatotheNetherlandsarenotMuslim$.

Table A14: Muslin majority countries (source: worldpopulationreview.com).

Country	Percentage of Muslims
Morocco	99
Somalia	100
Afghanistan	100
Tunisia	99
Iran	99
Mauritania	100
Yemen	99
Iraq	97
Maldives	100
Comoros	98
Niger	99
Turkey	100
Algeria	99
Azerbaijan	97
Jordan	97
Uzbekistan	88
Djibouti	94
Libya	97
Senegal	96
Pakistan	97
Tajikistan	98
Gambia	96
Egypt	90
Mali	94
Turkmenistan	93
Syria	87
Saudi Arabia	99
Bangladesh	89
Sudan	100
Kyrgyzstan	90
Indonesia	87
Oman	86
Guinea	89
Albania	57
Sierra Leone	77
Brunei	79
United Arab Emirates	76
Kazakhstan	70
Kuwait	75
Bahrain	74
Malaysia	61
Qatar	65
Burkina Faso	63
Lebanon	61
Chad	52
Bosnia and Herzegovina	51
Nigeria	50

their neighbourhood of residence[38], in the year in which the final exam is taken, we use the dataset called *GBAADRESOBJECTBUS*. This dataset provides two codes (*soortobjectnummer* and *rinobjectnummer*) that uniquely identify the neighbourhood of residence of each individuals that is registered in the Dutch population registration system. More specifically, we need to match these two codes with an additional code that identifies a neighbourhood (the variable

[38]Spacial data at neighbourhood level (*buurt* in Dutch) represent the smallest geographical unit we could obtain from CBS.

BU2021, available in the dataset called *VSLGWTAB*). This neighbourhood code can be then used to obtain the geospatial data of the neighbourhood (from the *gemeentewijkbuurt* dataset) or its socio-demograpic characteristics (from the *wijk_en_buurtstatistieken* dataset).

We obtained the mosques' geolocation and characteristics from the website *moskeewijzer.nl* in November 2021[39]. Importantly, for each mosque, we know whether it is a Moroccan or Turkish mosque as reported by the website. In order to use these data in our analysis, we asked CBS to encrypt the address of the mosques so that each could get a *soortobjectnummer* and *rinobjectnummer*. We then matched again these codes with the corresponding neighbourhood identifier (*BU2021*). Finally, we computed the distance between the centroids of the students neighbourhood of residence and the closest neighbourhood with a mosque. In doing so, we considered only Turkish mosques for students with a Turkish migration background and only Moroccans mosque for students with a Moroccan backgound.

B.3 Construction of outcome variables

We use two different datasets on outcomes in secondary education: a registry of the exam takers and exam outcomes (*EXAMVOTAB*) and a registry of the exam outcomes at course level (*EXAMVOVAKTAB*).

We start with the exam outcomes at the general level. First, we drop exam takers for which the exam outcomes are missing, while the exam taker did not withdrew from the exam. We only keep the students that passed or failed the final exams or that withdrew from the final exam. We standardize the average grade from the first central exams, the average grade from the central exams including the first and second retake (if applicable and grade improved) and the average grade from the continuous school exam by track with mean zero and standard deviation one. We generate a dummy that takes value one if the student passed the exams (and graduated).

For the data at course level (*EXAMVOVAKTAB*) we consider only the exam takers that are included in our analysis of *EXAMVOTAB*. The dataset contains 39 different exam subjects. We manually classified them into 29 subjects. Those exams that we could not classify into one of these subjects are dropped. The final list of exam subjects included in our analysis is the following: Arabic, Arts, Biology, Chemistry, Dutch, Economics, English, French, Friesian,

[39]The dataset used in our analysis contains the existing mosques in the Netherlands in the year 2021. As we do not know when a mosque was built, we use all the existing mosques in 2021 to compute the students' distance to the closest mosque. This introduces some measurement error as the map of mosques location has most likely changed between 2021 and 2014 (the first year considered in our analysis).

Geography, German, Greek, History, Latin, Management, Math (VMBO level, A, B, C), Music, Philosophy, Physics, Physics and Chemistry (I and II), Russian, Social Studies, Spanish, Technical Design, Turkish. This means that some practical exams and uncommon exams (both mainly in the pre-vocational track) are dropped from the sample. We merge these data with the date and time the exams are held, as reported by the website *www.examenblad.nl*.

B.4 SIM Survey

The *Survey integratie minderheden* (SIM) is a survey conducted by the *Sociaal en Cultureel Planbureau* (SCP) in collaboration with Statistics Netherlands. The survey includes questions on education, labour, social life, cultural integration, religion, health and sports. The interviews were conducted face-to-face, by telephone and online in 2006, 2011, 2015 and 2020.[40] The sample of the survey consists of respondents of 15 years or older of the four important (first and second generation) immigration groups in the Netherlands: Turkey, Morocco, the (former) Dutch Antilles, and Suriname (and in 2015 also Poland and Somalia), and of a control group of people without a migration background. Sample selection consisted of two steps: municipalities were randomly selected, within these municipalities the sample size was determined by municipality size, and people were selected randomly from the different immigration groups in the municipality.

In 2006, the data collection took place from March 2006 to December 2006. The sample consisted of 11,520 people, of whom 2,162 had a Turkish migration background and 2,390 a Moroccan migration background. The response rate was 53% (1,032 Moroccans and 1,132 Turks). In 2011, the data collection took place from November 2010 to June 2011. The sample consisted of 14,178 people, of whom 2,544 had a Turkish migration background and 2,826 a Moroccan migration background. The response rate was 48% (1,395 Moroccans and 1,347 Turks). In 2015, the data collection took place from January 2015 to July 2015. The sample consisted of 15,028 people, of whom 1,858 had a Turkish migration background and 2,082 a Moroccan migration background. The response rate was 45% (920 Moroccans and 951 Turks). In 2020, the data collection took place from September 2020 to Januari 2021.[41] The sample

[40]The interviews were conducted between March and December in 2006, between January to April and November to December in 2011, January to June in 2015 and March 2020 to January 2021. Due to measures against COVID-19 during Ramadan in 2020, mosques were closed and Muslims did not share the iftar meal with friends and neighbours.

[41]The data collection was delayed by the Covid-19 pandemic and took longer than planned because of lockdowns.

consisted of 15,498, of whom 2020 had a Turkish migration background and 2,114 a Moroccan migration background. The response rate was 35.4% for individuals with a Turkish migration background (696 respondents) and 27.5% for individuals with a Moroccan migration background (570 respondents).

We linked the SIM data to the administrative data from Statistics Netherlands. Some respondents are included in more than one survey wave and for those respondents we only keep the first observation. We are left with 5,400 respondents in total of a (first or second) generation Moroccan or Turkish migration background. In total, we have 2,670 respondents with a Moroccan migration background and 2,730 respondents with a Turkish migration background.

B.5 Additional Regressions

In this section we provide details for the regressions that we estimate in our heterogeneity analyses.

In section 6.2, we replicate our main results, using data at the subject level. To do so, we estimate the following regression:

$$\begin{aligned} y_{i,t,z} = \alpha_{p,z} + \theta_{i,z} + \eta_{i,z} + \beta Treated_i * Ramadan\ Year_t + \zeta Treated_i \\ + Exam\ Num_{i,t,z} + \delta X_{i,t} + \varepsilon_{i,t,z} \end{aligned} \tag{3}$$

$y_{i,t,z}$ is the grade obtained by student i, in year t, in subject z, standardized at the subject-track-year level. $Exam\ Num_{i,t,z}$ represents fixed effects for the number of exams that student i in year t has already completed while taking exam z. $\alpha_{p,z}$ are school-track-exam subject fixed effects. $\theta_{i,z}$ are treatment type-exam subject-afternoon- two day exam fixed effects. The controls in the equation are the same as in the previous specifications. Standard errors are clustered on school-subtrack-exam year-exam subject level.

In section 7.1 we estimate the impact of Ramadan on exam performance, based on the timing of the exams and how many exams a student takes in a day. To do so, we employ the following regression:

$$\begin{aligned} y_{i,t,z} = \alpha_{p,z} + \theta_{i,z} + \delta X_{i,t} + \gamma_t + Exam\ Num_{i,t,z} + \beta_1 DiD_{i,t} + \beta_2 Afternoon_z + \beta_3 Two\ Exam_z \\ + \beta_4 DiD_{i,t} * Afternoon_z + \beta_5 DiD_{i,t} * Two\ Exam_{i,z} + \beta_6 Afternoon_z * Two\ Exam_{i,z} \\ + \beta_7 DiD_{i,t} * Two\ Exam_z * Afternoon_z + \zeta Treated_i + \varepsilon_{i,t,z} \end{aligned} \tag{4}$$

$y_{i,t,z}$ is the grade obtained by student i, in year t, in subject z, standardized at the subject-track-year level. DiD_{it} is equal to the interaction between $Treated_i$ and $Ramadan\ Year_t$. $Afternoon_z$ is a dummy equal one if exam z is in the afternoon. $Two\ Exam_{i,z}$ is a dummy equal one if individual i is taking exam z in the day same day of another exam. $Treated_i$ is replaced by individual fixed effects in some specifications. All the other variables are the same as in equation(1).

In section 7.1 we estimate the impact of Ramadan on exam performance, depending on the week of Ramadan in which the exam is held. We estimate the following regression:

$$y_{i,t,z} = \alpha_{p,s,z} + \theta_{i,z} + \omega_{i,z} + \delta X_{i,t} + \gamma_t + Exam\ Num_{i,t,z} + \sum_{t=2014}^{2017} \beta_t Treated_i * Year_t + \zeta Treated_i$$
$$+ \sum_{n=0}^{3} Week_{z,n} + \phi_n Week_{z,n} * Treated_i + \varepsilon_{i,t,z} \quad (5)$$

$y_{i,t,z}$ is the grade obtained by student i, in year t, in subject z, standardized at the subject-track-year level. $Week_{z,n}$ is a dummy taking value on if exam z is held in week n of the Ramadan month. $Week_{z,0}$ is equal to one if exam z is held in 2018 in the week before the beginning of the Ramadan month. The coefficients of interests are ϕ_1, ϕ_2, ϕ_3, which captures the impact of Ramadan exposure during the first, second or third week of the Ramadan month.

To test the possibility that Ramadan exposure affects students differently, depending on whether the exam is taken at the beginning or at the end of the Ramadan month, we also implement the following regression:

$$y_{i,p,s,t} = \alpha_{p,s} + \beta_1 Treated_i * Ramadan\ Year_t + \beta_2 Ramadan\ Days * Treated_i * Ramadan\ Year_t$$
$$+ \delta X_{i,t} + \gamma_t + \varepsilon_{p,s,t} \quad (6)$$

In the equation above *Ramadan Days* is a variable measuring how many days after the beginning of Ramadan an exam is taken. All other variables are the same as in equation (1). The coefficient β_2 measures the marginal effect of taking exams after an additional day of Ramadan exposure.