

# The Effect of Physical Activity on Student Performance in College: An Experimental Evaluation

## AUTHORS

**Hans Fricke**

Stanford University

**Michael Lechner**

University of St. Gallen

**Andreas Steinmayr**

University of Munich

## ABSTRACT

What is the role of physical activity in the process of human capital accumulation? Brain research provides growing evidence of the importance of physical activity for various aspects of cognitive functions. An increasingly sedentary lifestyle could thus be not only harmful to population health, but also disrupt human capital accumulation. This paper analyzes the effects of on-campus recreational sports and exercise on educational outcomes of university students. To identify causal effects, we randomize financial incentives to encourage students' participation in on-campus sports and exercise. The incentives increased participation frequency by 0.26 times per week (47%) and improved grades by 0.14 standard deviations. This effect is primarily driven by male students and students at higher quantiles of the grade distribution. Results from survey data suggest that students substitute off-campus with on-campus physical activities during the day but do not significantly increase the overall frequency. Our findings suggest that students spend more time on campus and are better able to integrate studying and exercising, which may enhance the effectiveness of studying and thus improve student performance.

**Acknowledgements:** We thank Heinz Hofstetter and Felix van den Berg for invaluable support with the administrative data of the University of St. Gallen, Nora Zogg and Peggy van der Wallen for administering our surveys, Daniel Weber, Viola Rutar, and Pavel Obraztsov for outstanding research assistance. We are grateful for valuable remarks by Joshua Angrist, Eric Bettinger, Yvonne Giesing, Anna Gumpert, Angela Johnson, Raaele Miniaci, Hessel Oosterbeek, Jane Rochmes, and by seminar participants at the LaborWorkshop in Laax, Stanford University, Tinbergen Institute in Rotterdam, University of Nottingham, University College London, and the 2017 AEFP Annual Conference. We are grateful for financial support from Migros. Hans Fricke received funding from the Swiss National Science Foundation through grant P1SGP1 158810. The usual disclaimer applies.

## VERSION

May 2017

**Suggested citation:** Fricke, H., Lechner, M., & Steinmayr, A. (2017). The Effect of Physical Activity on Student Performance in College: An Experimental Evaluation (CEPA Working Paper No.17-03). Retrieved from Stanford Center for Education Policy Analysis: <http://cepa.stanford.edu/wp17-03>

# The Effect of Physical Activity on Student Performance in College: An Experimental Evaluation

Hans Fricke, Michael Lechner, and Andreas Steinmayr\*

May 16, 2017

**Abstract** What is the role of physical activity in human capital accumulation? Brain research provides evidence of the importance of physical activity for cognitive functions. An increasingly sedentary lifestyle could thus disrupt human capital accumulation. We analyze effects of on-campus recreational sports and exercise on educational outcomes of university students. We randomize financial incentives to encourage students' participation in on-campus sports and exercise. Incentives increased participation frequency by 47% and improved grades by 0.14 standard deviations. Students primarily substitute off-campus with on-campus physical activities and are better able to integrate studying and exercising, which may enhance the effectiveness of studying.

**Keywords** Sports, physical activity, human capital, student achievement, randomized experiment

**JEL Classifications** C93, I12, I18, I23, J24

We thank Heinz Hofstetter and Felix van den Berg for invaluable support with the administrative data of the University of St. Gallen, Nora Zogg and Peggy van der Wallen for administering our surveys, Daniel Weber, Viola Rutar, and Pavel Obraztsov for outstanding research assistance. We are grateful for valuable remarks by Joshua Angrist, Eric Bettinger, Yvonne Giesing, Anna Gumpert, Angela Johnson, Raffaele Miniaci, Hessel Oosterbeek, Jane Rochmes, and by seminar participants at the Labor Workshop in Laax, Stanford University, Tinbergen Institute in Rotterdam, University of Nottingham, University College London, and the 2017 AEFPP Annual Conference. We are grateful for financial support from Migros. Hans Fricke received funding from the Swiss National Science Foundation through grant P1SGP1\_158810. The usual disclaimer applies.

---

Contact: Hans Fricke: Stanford University, hfricke@stanford.edu; Michael Lechner: University of St. Gallen, IZA, CESifo, and CEPR, michael.lechner@unisg.ch; Andreas Steinmayr: University of Munich (LMU), IFW Kiel, and IZA, andreas.steinmayr@econ.lmu.de

# 1 Introduction

Evidence from neuroscience and related disciplines suggests that physical exercise could contribute to the formation of human capital. This literature shows that physical activity alters functional and structural properties of certain brain structures, which leads to learning and skills acquisition (Hillman, Erickson, and Kramer, 2008; Hötting and Röder, 2013). Intervention studies have discovered both immediate and persistent effects of exercising on cognitive performance. Physical activity persistently increases executive functions, attention, memory, and speed of processing weeks and months after interventions (Hötting and Röder, 2013). Furthermore, physical activity has immediate effects and, for example, enhances memory storage and retrieval during and shortly after exercising (Lambourne and Tomporowski, 2010). Hillman et al. (2009) find that even single bouts of exercise may improve the cognitive control of attention and support cognitive health.

A sedentary lifestyle might thus not only be problematic from a health perspective but might also impede the process of human capital accumulation.<sup>1</sup> This is particularly true for students, as attending lectures and studying involve hardly any physical activities. Physical inactivity as an obstacle to human capital accumulation has been largely ignored in economic research, especially compared to the vast literature on the effects of education and other environmental factors on human capital formation (for summaries on factors see Cunha, Heckman, Lochner, and Masterov, 2006; Currie and Almond, 2011).

We conduct an experimental study to identify the causal effect of physical activity on educational outcomes of university students. We use an encouragement design<sup>2</sup> and randomize incentives to participate in on-campus recreational sports and exercise among two cohorts of freshmen students at a Swiss university. Students can earn up to 200 CHF (approximately 220 USD) if they use the recreational programs or facilities twice per week during their first year. Prior studies have found that similar financial incentives over shorter periods can raise exercise levels of college students (Charness and Gneezy, 2009; Acland and Levy, 2015).<sup>3</sup> Hence, the random assignment of the financial incentives can be used as an instrument for actual participation.

Moreover, we investigate several potential mechanisms through which increased exercise on-campus could influence educational outcomes. First, we investigate characteristics of activities beyond frequency, including location, timing, and type. Studies in neuroscience suggest that these dimensions might be relevant. For instance, studying might be more effective if

---

<sup>1</sup>Research in other fields has extensively documented the positive health effects of physical activity and exercise (see for example U.S. Department of Health and Human Services, 1996; Strong et al., 2005; Warburton, Nicol, and Bredin, 2006; Janssen and Leblanc, 2010)

<sup>2</sup>Research designs using incentives to identify a causal effect have been used before in the literature. For example, Katz, Kling, and Liebman (2001) use a related design to identify the causal effects of neighborhood characteristics on individual outcomes.

<sup>3</sup>See Royer, Stehr, and Sydnor (2015); Carrera, Royer, Stehr, and Sydnor (2017) for evidence of the effects of financial incentives and commitment contracts on exercise for employees of a Fortune-500 company.

preceded by an exercise session since increased cognitive performance has been shown to occur immediately after exercising (Hillman et al., 2008; Hötting and Röder, 2013). Second, we explore students' use of time. Time spent exercising requires a reduction in time devoted to other activities. Grades might deteriorate if less time is spent studying (Stinebrickner and Stinebrickner, 2008). On the other hand, if on-campus sports replace other potentially harmful leisure activities such as drinking, grades might be positively affected.<sup>4</sup> Third, we investigate several self-reported measures of health and emotional well-being. Benefits of physical activity on these dimensions have been extensively documented.<sup>5</sup> These benefits could lead to gains in productivity and therefore improve educational outcomes.

Our study draws on three unique data sources to answer the above questions. First, we collect comprehensive on-campus sports and exercise information with a student ID scanner system at the entrance of the university sports facilities. Second, we use administrative student records for educational outcomes. Third, we administer online surveys at the end of each semester to obtain information on sports and exercise off campus, on students' use of time, and on health and well-being. These surveys allow us to shed light on the mechanisms through which educational outcomes may be affected. All three data sources are linked using a unique student identifier.

Our results suggest a substantial impact of on-campus exercise on educational outcomes. Cash incentives raised on-campus sports by 47% or 0.26 facility visits per week in the first cohort. Incentives were less effective at increasing participation in the second cohort. Thus, we restrict our analysis of educational outcomes to the first cohort. The increased activity improved grades of the treatment group by on average 0.14 standard deviations compared to those of the control group. This translates into an improvement in grades of 0.42 standard deviations for a one standard deviation increase in on-campus exercise frequency. Additional evidence from surveys suggests that the incentives did not raise the overall level of physical activity but led to a substitution of on-campus activity for off-campus activity. On-campus activity increased mostly in the afternoon. The survey data also provides evidence that students spend more time in class. Taken together, these findings indicate that integrating studying and exercising during the day may enhance the productivity of study time and thus improve student performance. In line with the fact that overall activity does not seem to be affected, health outcomes and emotional well-being do not appear to explain educational gains. Educational gains are primarily driven by male students and students at higher quantiles of the grade distribution.

This study relates to a growing economic literature on the effects of sports and physical activity on educational and labor market outcomes. Existing results suggest positive associations of physical activity with cognitive and non-cognitive development of children (Felfe et al.,

---

<sup>4</sup>Several studies suggest adverse effects of binge drinking on academic achievement in college. See for example DeSimone (2007) and Lindo, Swensen, and Waddell (2013)

<sup>5</sup>See for example U.S. Department of Health and Human Services (1996), Strong et al. (2005), Warburton et al. (2006), Janssen and Leblanc (2010), Felfe, Lechner, and Steinmayr (2016), and Lees and Hopkins (2013).

2016), with school grades (Lipscomb, 2007; Rees and Sabia, 2010; Pfeifer and Cornelißen, 2010), as well as with labor market success of adults (Kosteas, 2012; Lechner, 2009; Barron, Ewing, and Waddell, 2000; Pfeifer and Cornelißen, 2010; Stevenson, 2010; Eide and Ronan, 2001; Ewing, 2007; Rooth, 2011).<sup>6</sup> A limitation of the above studies is how they tackle the problem of self-selection into physical activity. Individuals who practice sports or exercise differ from those that do not (Schneider and Becker, 2005; Farrell and Shields, 2002). These differences might be related to other life outcomes as well, confounding the effect of sports on educational and labor market outcomes. Existing studies used instrumental variable, selection on observable, or panel data methods to address this endogeneity problem. However, concerns regarding a causal interpretation of the findings remain.

This study makes three contributions to the literature. First, to the best of our knowledge, we are the first to use an experimental design to identify the causal effect of recreational sports and exercise on educational outcomes of university students. Second, we provide a comprehensive picture of this effect and the channels through which these effects operate. This is possible with our novel data that combines detailed information on exercise behavior with administrative student data and rich survey data. Third, we add to the literature on the effectiveness of financial incentives to form exercise habits, providing results of a longer lasting intervention.

The remainder of this paper is structured as follows. Section 2 introduces the institutional setting and describes the intervention and randomization design. Section 3 describes the data and presents descriptive statistics. Section 5 presents the results and Section 6 concludes.

## 2 Institutional setting

### 2.1 General background

We conducted the experiment at the University of St. Gallen during the first year of the incoming cohorts in 2013 and 2014. The University of St. Gallen is a public university in Switzerland and is one of the main universities to offer undergraduate programs in business, economics, law, and related fields. It is considered one of the top business schools in Europe.<sup>7</sup>

The first-year curriculum is almost identical for all undergraduate students regardless of their major. In this first year, the curriculum is strictly organized and mandatory. The curriculum centers around two courses in each of the three core fields: business, economics, and law. Appendix A.6 provides a complete overview of the first-year curriculum. If students fail this year, they can attempt the entire curriculum a second time or leave the university. About 48%

---

<sup>6</sup>College sports have only been analyzed by a few studies in the US (Long and Caudill, 1991; Henderson, Olbrecht, and Polachek, 2006; Caudill and Long, 2010). They find that college athletes earn more in some occupations. However, intercollegiate sports take place at an almost professional level and are thus substantially different from on-campus recreational sports.

<sup>7</sup>St. Gallen awards about 30% of all undergraduate degrees in Business Administration and Economics in Switzerland (Tafreschi and Thiemann, 2016). The Financial Times ranked the University of St. Gallen 4th in the European Business School Ranking in 2015.

of students do not pass the first year on the first attempt, which points to the highly competitive nature of the curriculum.

The University of St. Gallen offers a large variety of recreational sports free of charge. Most of these activities are concentrated in one large complex on campus. This complex includes a gym and several outdoor and indoor fields for team sports. Moreover, the university offers an extensive number of instructed classes, such as dance, group fitness, and martial arts in these facilities. For a complete list, see Appendix A.7. Henceforth, we refer to the complex as the “gym.” The university only offers a few activities off-campus, such as climbing, swimming, and sailing, if the corresponding facilities are not available on-campus.

## 2.2 The experiment

We use an encouragement design to identify the effects of physical activity on educational outcomes. That is, we randomize financial incentives to increase participation in recreational sports and exercise on campus.<sup>8</sup> Financial incentives were randomized among new undergraduate students in 2013 and 2014 who answered a baseline survey. We implemented a blocked randomization design based on pre-treatment characteristics.<sup>9</sup> All new undergraduate students in these two years received a survey invitation in August, one month before students started the program. The response rates were 42% of 1,340 students in the 2013 cohort and 58% of 1,293 students in the 2014 cohort.<sup>10</sup>

The incentives were structured as follows: Students were offered an initial endowment of CHF 100 (equals approximately USD 110 at the time of the experiment) in each semester. Therefore, students could earn a total of CHF 200 (CHF 100 per semester). The CHF 100 corresponds to CHF 10 per week (over 10 weeks). If students participated in activities in the gym twice per week during every week, they received the entire amount. Each week the endowment was reduced by CHF 5 if they participated only once that week, and by CHF 10 if they did not participate at all that week. In each cohort, we provided the incentives in two semesters. Incentives were provided in 10 weeks out of the 14-week semester. These 10 weeks covered the third and the last week of the semester. We did not provide incentives during a two-week break in the middle of the semester as most students are not in St. Gallen.<sup>11</sup>

---

<sup>8</sup>This research design differs from those used, for example, in neuroscience to study the chronic effects of exercise. These studies usually assign incentives to the treatment group to follow a relatively strict training plan, whereas we only assign incentives to increase participation. Thus, our treatment likely has a smaller effect on amount and quality of exercise than designs with an assigned training plan. However, incentive designs can be readily implemented on a larger scale.

<sup>9</sup>Students were grouped into 13 blocks in 2013 and 12 blocks in 2014 based on available pre-treatment characteristics such as gender and nationality. In all blocks, approximately half of the students were randomly assigned to treatment and control group. Table A.1 shows mean values of the characteristics used in the randomization, the number of students in each block, and the effective treatment probability.

<sup>10</sup>See Table A.2 in the appendix for descriptive statistics of students in the randomization sample and the remaining students by cohort.

<sup>11</sup>In a recent discussion paper, Carrera et al. (2017) show that such constant incentives are particularly useful to increase physical activity at the intensive margin.

To possibly obtain larger effects on sports and exercise participation, we structured the incentives appealing to students' loss aversion. That is, we framed the incentives in a way such that students would lose money if they did not exercise instead of receiving money if they did (Kahneman and Tversky, 1979). Such incentives have proven to be effective in our pilot study. Note that by appealing to loss aversion, we deviate from the structure of previous studies that used financial incentives to increase exercise (Charness and Gneezy, 2009; Acland and Levy, 2015; Royer et al., 2015).

We invited students in the treatment groups to participate in a pilot program to foster participation in on-campus recreational activities. In the first week of the semester we sent students a letter with the invitation and a personalized credit card-sized card to check participation.<sup>12</sup> In addition, we notified students with an email that they had been selected for the pilot program.

Participation was checked by the university sports staff and course instructors. Every time students participated in activities in the gym, the staff or the course instructors handed them a sticker. The personalized cards had two marked slots per week for these stickers. In order to minimize cheating, we instructed gym staff and course instructors to make sure that students wore gym clothes or showed clear signs of physical exhaustion and to check if the name on the card corresponded to the name of the university ID.

At the outset of the following semester, we paid the remaining endowment (max. CHF 100) as a voucher of the university caterer Migros. The voucher could be redeemed in all university cafes and cafeterias as well as in retail stores owned by the same company.<sup>13</sup> Furthermore, the voucher could be exchanged for cash in the main cafeteria and hence should have been valued by the students close to the cash amount.<sup>14</sup>

Since randomization insures independence of the incentives and student characteristics, we can estimate the reduced form effect of receiving the incentives on student outcomes with a straightforward means comparison:

$$y_i = \alpha_0 + \alpha_1 \text{incentives}_i + \epsilon_i, \quad (1)$$

where  $y_i$  is the outcome of interest of student  $i$ . Primary measures for on-campus physical activity are the average number of gym visits per week in the first and second semester, and in the first year overall. For educational outcomes, we report individual grades in the core courses: business, economics, and law. All students have to take a course in each of these fields in the first and second semester. In addition, we report the effect on the weighted average grade in each semester and in the first year. As a measure of persistence in the university, we

---

<sup>12</sup>Appendix A.9 shows the card.

<sup>13</sup>Meals are not covered by tuition fees, nor does the university provide meal plans.

<sup>14</sup>Both cohorts redeemed a similar share of stickers in the first semester (28% in the 2013 cohort and 29% in the 2014 cohort). The 2013 cohort redeemed a higher share of stickers in the second semester than the 2014 cohort (31% vs. 25%). Figure A.3 in the appendix shows the distribution of the number of stickers among students who redeemed the sticker card. In both cohorts, the majority of students redeeming their voucher received the full amount of CHF 100.

use an indicator for whether or not a student passes the first year. For easier interpretation, we standardize all grades to have mean zero and standard deviation one within cohort and randomization sample. See Table A.3 in the appendix for standard deviations of the different grades. The university assigns the lowest grade possible to students who do not show up to the exams.<sup>15</sup> We keep this assignment but present also the estimated effect on the probability of having missed at least one exam in the respective term. The outcome measures to investigate potential channels are explained in Section 5.3.  $incentives_i$  is a binary indicator that takes the value of one if the student received the incentives and zero otherwise.  $\alpha_0$  corresponds to the mean level of the outcome variable in the control group.  $\alpha_1$  is the mean difference between the treatment and control groups and thus measures the effect of the incentives. This effect is an Intention to Treat effect (ITT).<sup>16</sup>

Under a set of further assumptions, the incentives can be used as an instrumental variable for participating in sports at the university. First, the incentives have to significantly increase students' on-campus sports participation (instrument relevance). Sections 4 show that this assumption holds for the 2013 cohort. Second, the incentives must have a non-negative effect on each student, i.e. students should only increase their participation in response to the incentives (monotonicity). Given that the treatment rewards only increases in physical activity this monotonicity assumption is unlikely to be violated.<sup>17</sup> Furthermore, implications of monotonicity are tested in Appendix A.8 and suggest that monotonicity is plausible. Third, any effect of the incentives on outcomes has to work through the effect on exercise behavior (exclusion restriction). A violation of this assumption could occur, for instance, if the money students can earn with their participation directly affects their academic performance. However, the incentives are paid out after the exams. Thus, students would need to change their behavior in anticipation of the additional income. Such a behavior seems unlikely as the maximum amount of CHF 100 per semester is small relative to the overall living expenses, which are estimated at CHF 12,720 per semester in 2013.<sup>18</sup> Another concern could be that students feel more welcome, integrated, or connected with the university because of the experiment. If the mere receipt of the email and the letter creates this feeling, the exclusion restriction would be violated. Yet, such a link is unlikely as all students including the control group receive a lot of information including a letter and a booklet with all offers from the university sports administration. If students feel

---

<sup>15</sup>Grades are assigned on a scale of one to six with six being the highest score.

<sup>16</sup>One potential concern is a violation of the Stable Unit Treatment Value Assumption (SUTVA) (Rubin, 1980). The treatment status of one student could potentially have an effect on exercise behavior and indirectly on other outcomes of other students. For example, if treated students increase their exercise level and motivate some untreated peers to join them, the activity level in the control group might also go up. While we can not rule out such spillover effects, positive spillover effects to the control group would bias our effect towards zero.

<sup>17</sup>This should at least have been true when the incentives were provided. If financial incentives crowd out intrinsic motivation, there could indeed be a negative long-run effect on activity. This, however, does not seem to be the case in recent studies (Charness and Gneezy, 2009; Acland and Levy, 2015; Royer et al., 2015).

<sup>18</sup>Calculations are done by the university and based on the Social and Economic Conditions of Student Life Survey (SSEE) collected by Swiss Statistics in 2013.



welcome because they actually participate in recreational activities, this effect is not a threat to our identification strategy as it works through exercise and is thus part of the treatment effect.

Unfortunately, we do not know the true relationship between incentives, exercise, and educational outcomes. Thus, the specification of the treatment variable is to some extent arbitrary. While using weekly gym visits as treatment is an obvious choice, it is possible that the general fitness level is relevant or that only exercise directly before the exams or study sessions matters. The results furthermore suggest substitution between sports on and off campus (see Section 5.3). In that case, the characteristics rather than the quantity of the activities would be the relevant treatment. For these reasons, we provide the ITT estimates as main results.

We present Two-Stage Least Squares (TSLS) estimates of the following IV model in Section 5.2:

$$y_i = \beta_0 + \beta_1 gymvisits_i + v_i, \quad (2)$$

$$gymvisits_i = \gamma_0 + \gamma_1 incentives_i + \vartheta_i, \quad (3)$$

where  $gymvisits_i$  corresponds to the average number of visits per week.  $\beta_1$  identifies the marginal effect of one gym visit per week if the linear model is true. Otherwise, Frölich (2007) shows that  $\beta_1$  can be interpreted as an average local treatment effect over the various complier populations characterized by their initial level of gym visits and their increase in visits.<sup>19</sup>

For all estimates, we show heteroskedasticity robust Huber-White standard errors. Additionally, we implement a permutation procedure for the ITT specifications that does not rely on asymptotic approximations. This procedure provides exact p-values for the test of the null hypothesis of no treatment effect against the two-sided alternative for the ITT parameter. The procedure estimates the ITT coefficient for 1000 possible combinations of treatment assignments within randomization clusters. Once the distribution of treatment effects is estimated, we can test at which percentile of the distribution the actual treatment effect falls.<sup>20</sup>

## 3 Data and description of student body

### 3.1 Data

This study makes use of three different data sources. All three data sources can be merged via a unique identifier. First, on-campus recreational sports and exercise is recorded via an ID card

---

<sup>19</sup>Angrist and Imbens (1995) propose the interpretation of an average per unit treatment effect. However, this interpretation depends on the assumption that increases are not overlapping, e.g. no one can increase visits from zero to twice per week while some increase visits from zero to once per week, which is unlikely to hold in our setting.

<sup>20</sup>For the estimations including covariates, we implement the procedure suggested in Rosenbaum (2002). In short, we regress the outcome on the covariates and then use the residuals from this regression to conduct the same permutation exercise as above.

scanner at the entrance of the sports facilities. The staff at the entrance ensures that students scan their ID card when entering the facilities. The scanner saves the student ID number, and the current time and date. Students then select the type of activity they plan to do on a touchscreen. As mentioned before, these facilities cover all on-campus activities and most activities provided by the university. Thus, the scanner data provide us with very accurate information about the frequency and the type of exercise on campus, for both treatment and control group.

Second, we use administrative student records of the university. These data contain information on courses and grades, enrollment, drop-out, and some socio-demographic characteristics such as gender, age, nationality, and region of high school.<sup>21</sup>

Third, we administer online surveys to the students before they enter the university (baseline) as well as after their first and second semester (follow-up). The surveys describe the incoming freshmen cohort, measure exercise outside the university, and collect information on student time-use, well-being, health status, as well as student health-related lifestyle.<sup>22</sup>

Table 1 shows response rates to the first and second follow-up surveys by cohort. The response rates for both follow-up surveys are relatively low but somewhat higher in the 2013 cohort than in the 2014 cohort, 49% vs. 40% in the first follow-up survey and 56% vs. 28% in the second follow-up survey. One potential explanation for the low participation is that students were surveyed for a different unrelated study and were asked to provide course evaluations at a similar time. The stark difference in the response rates in the second follow-up survey (56% vs. 28%) is caused by an additional intervention to raise response rates in the 2013 cohort.<sup>23</sup> Table 1 also provides evidence that response behavior does not seem to be affected by the incentives as differences in response rates between the treated and control group are negligible and insignificant.

## 3.2 Description of student body

Table 1 shows descriptive statistics for the students in our sample for the cohorts 2013 and 2014, respectively.<sup>24</sup> In both cohorts, students are on average approximately 20 years old and around 40% of students are female. Only 6% are enrolled in the law specific first-year curriculum, which includes two additional law courses instead of mathematics. However, there are also several differences between the cohorts. In particular, the share of Swiss students decreases from 80% to 72% between the cohorts. Relatedly, the share of students who have to take the

---

<sup>21</sup>Several recent papers have made use of these administrative data in other contexts (Fricke, Grogger, and Steinmayr, 2015b; Tafreschi and Thiemann, 2016).

<sup>22</sup>We are not able to link the response of the 2014 baseline survey to the administrative data.

<sup>23</sup>Students were randomly offered to enter a cash lottery with different values conditional on survey response. See Fricke, Frölich, Huber, and Lechner (2015a) for details.

<sup>24</sup>Table A.2 in the appendix provides descriptive statistics of the randomization sample (students that answered the baseline survey) and the excluded sample (students that did not answer the baseline survey).

entry examination, foreign nationals without a Swiss high school degree, increases from 16% to 22%.<sup>25</sup> About 90% of students in both cohorts speak German as their native language, but a substantial share of the students, 19% and 17% respectively, choose the English track, which has the same curriculum but is taught entirely in English. These descriptive statistics suggest that the students are representative of the students at University of St.Gallen and largely of students at business focused universities in general.

---

<sup>25</sup>While all Swiss students with a high school degree are admitted to the University of St. Gallen, international students have to take an admissions test. Because of this initial screening, international students on average receive better grades in the first-year and are more likely to pass the first year on the first attempt.

Table 1: Descriptive statistics by treatment status

	<i>Cohort 2013</i>				<i>Cohort 2014</i>			
	Total	Control	Treated	p-value	Total	Control	Treated	p-value
Female (0/1)	0.37	0.37	0.38	0.87	0.40	0.40	0.40	1.00
Age (years)	19.91	19.99	19.83	0.23	19.95	19.96	19.95	0.94
Student aid contribution (0/1)	0.06	0.08	0.05	0.11	0.08	0.09	0.06	0.13
English track (0/1)	0.19	0.19	0.18	0.78	0.17	0.15	0.19	0.24
Law track (0/1)	0.06	0.05	0.06	0.61	0.06	0.06	0.06	0.88
Swiss national (0/1)	0.80	0.80	0.79	0.80	0.72	0.72	0.72	0.87
Entry exam (0/1)	0.16	0.17	0.16	0.69	0.22	0.21	0.22	0.93
German mother tongue (0/1)	0.90	0.91	0.88	0.29	0.88	0.89	0.87	0.51
High school degree from								
Canton St.Gallen (0/1)	0.15	0.16	0.14	0.60	0.13	0.14	0.12	0.45
Canton Zurich (0/1)	0.14	0.14	0.13	0.77	0.17	0.17	0.16	0.70
Other German speaking Canton (0/1)	0.46	0.46	0.46	0.96	0.39	0.38	0.40	0.60
Non-German speaking Canton (0/1)	0.08	0.06	0.09	0.22	0.08	0.09	0.08	0.69
Non-Swiss institution (0/1)	0.17	0.18	0.17	0.87	0.23	0.22	0.24	0.55
Gym visits (avg per week)	0.35	0.37	0.32	0.41	0.43	0.46	0.39	0.26
Most frequent activity pre treatment								
Gym (0/1)	0.13	0.14	0.13	0.59	0.12	0.14	0.11	0.19
Other (0/1)	0.19	0.17	0.21	0.26	0.23	0.21	0.24	0.30
None (0/1)	0.68	0.69	0.67	0.59	0.65	0.65	0.65	1.00
Follow-up survey 1 (0/1)	0.49	0.50	0.48	0.77	0.40	0.39	0.41	0.60
Follow-up survey 2 (0/1)	0.56	0.54	0.58	0.34	0.28	0.28	0.28	0.94
Observations	567	282	285		754	377	377	

Notes: The table presents the means of the covariates by treatment status separately for the cohorts 2013 and 2014. p-values stem from a t-test on the equality of means. “Age” corresponds to the age at enrollment. “Student aid contribution” indicates whether or not a student donated to the student aid fund with her first tuition payment. “English track” indicates whether or not a student is enrolled into the English first-year curriculum, which covers the same topics but is taught entirely in English. “Law track” indicates whether or not a student is enrolled into the law specific first-year curriculum, which requires two additional law courses instead of mathematics. “Entry exam” indicates if a student had to take an admissions test. All non-Swiss students without a Swiss high school degree have to take an admissions test. “Gym visits pre-treatment” is the average number of visits per week in the first two weeks of the semester before incentives start. “Most frequent activity pre-treatment” indicate the most frequent type of activity during the first two weeks during the semester. “Follow-up survey 1” indicates participation in the follow-up survey at the end of the first semester. “Follow-up survey 2” indicates participation in the follow-up survey at the end of the second semester.

Table 1 also provides reassurance that the randomization was successful. Means of all observed student characteristics are very similar between those students who received the incentives and those who did not. T-tests fail to reject the null hypothesis that the mean differences in observed characteristics between both groups are equal to zero in all tests. Given the good balance, controlling for covariates does not significantly change our estimates. See Appendix A.4.

## 4 Effects on participation frequency

The experimental design relies on the effectiveness of the incentives to increase students' participation in on-campus recreational sports and exercise. Figure 1 shows weekly mean participation frequencies over the academic year for both cohorts. The gray bars correspond to the participation frequencies for students who did not receive the incentives. In both cohorts, these students visited the facilities around 0.6 times per week in the first term and around 0.5 times in the second term.

The dashed lines represents weekly treatment effects of the incentives on participation frequencies. These effects differ substantially between cohorts. The incentives increase participation of the treatment group in both semesters for the 2013 cohort by about 50%. In comparison, treatment effects are substantially weaker for the 2014 cohort. This is particularly true for the second semester. In the second half of the second semester, weekly treatment effects are statistically indistinguishable from zero. These differences raise the question of whether the incentives increased activity sufficiently to estimate the effect on educational outcomes for the 2014 cohort.<sup>26</sup>

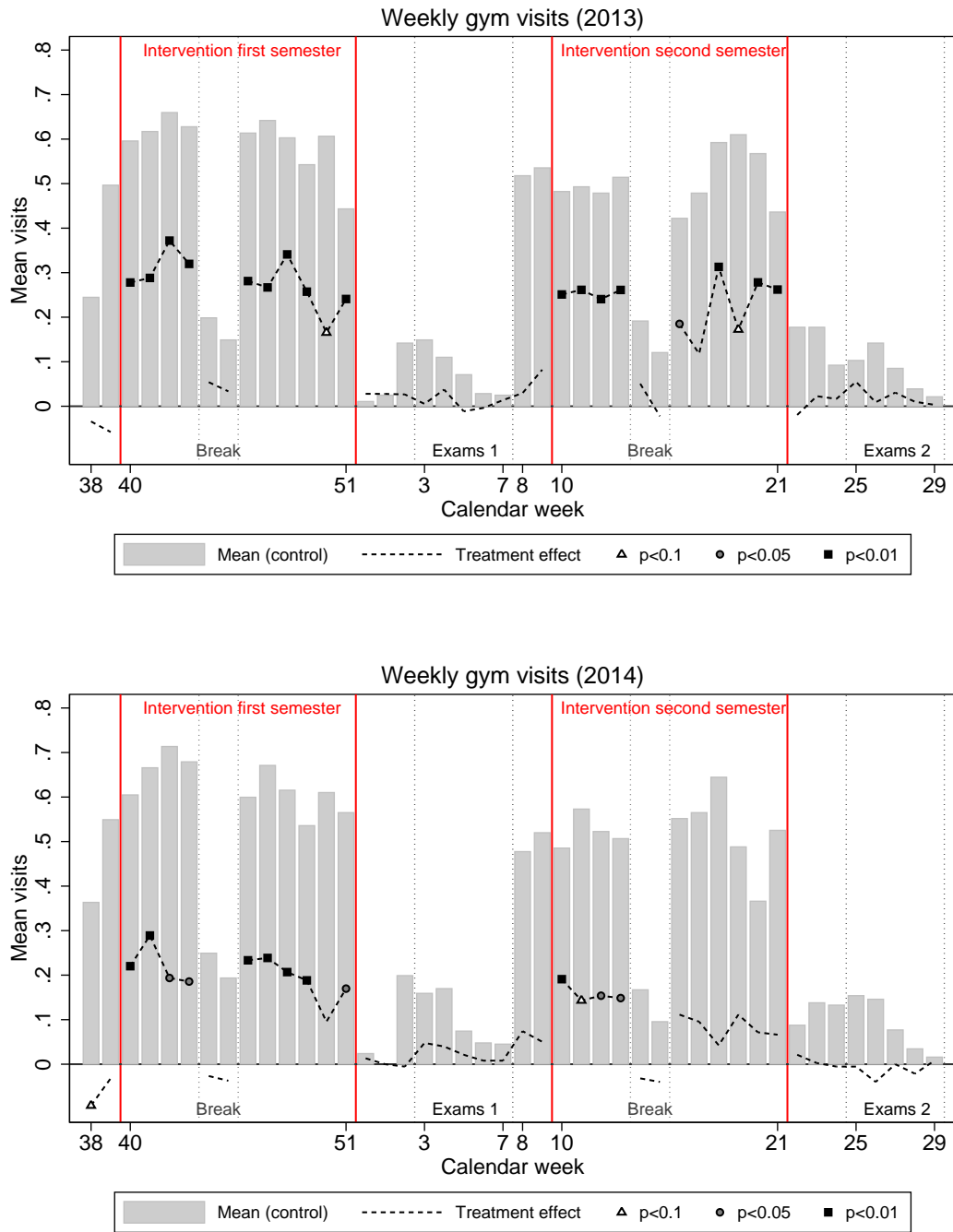
Table 2 shows first stage regression results on aggregate measures of weekly gym visits for both cohorts to address this question systematically. The results confirm the substantial differences in the effectiveness of the incentives between cohorts. While for both cohorts mean visits per week of the treated group increased significantly in both semesters, the magnitude of these effects is much smaller for the 2014 cohort. For the 2013 cohort, treated students visited the gym 0.26 times more per week over the academic year, which represents a 47% increase in comparison to the control group. In the 2014 cohort, the treated students visited the gym only 0.16 times or 28% more often per week over the academic year. Moreover, the F-statistic is well above the suggested threshold for IV analyses of 10 in the 2013 cohort but below this threshold in the 2014 cohort. Therefore, we report the results for the 2013 cohort as main results. Results for the 2014 cohort can be found in Appendix A.2. Table A.5 in the appendix shows estimates using the pooled sample.<sup>27</sup>

---

<sup>26</sup>We see no treatment effects in the semester breaks without incentives, in which the participation level of the control group is also substantially lower. In addition, Appendix A.1 shows that the incentives do not seem to affect activity beyond the treatment periods during the second year of the 2013 cohort. These results are in line with those of Acland and Levy (2015) who show that after the first-semester break newly formed activity habits disappear.

<sup>27</sup>Appendix A.3 shows that time varying observable characteristics such as weather or overall gym utilization cannot explain the difference in participation between cohorts.

Figure 1: Effect of incentives on gym usage by week



Notes: The figure depicts weekly mean visits of the control group, and the treatment effects of the incentives. p-values stem from a t-test on the equality of means. Calendar weeks 38 to 51 correspond to the first semester and calendar weeks 8 to 21 to the second semester. The respective exam periods are from calendar week 3 to 7 and 25 to 29. In the middle of each semester, a two-week break without classes takes place. No incentives are provided in these breaks.

Table 2: Effect of incentives on gym usage

	(1)	(2)	(3)
	Visits total	Visits 1st sem	Visits 2nd sem
Cohort 2013			
Cash incentives	0.26*** (0.06) [0.00]	0.28*** (0.07) [0.00]	0.23*** (0.07) [0.00]
Constant	0.55*** (0.04)	0.60*** (0.04)	0.51*** (0.05)
F-statistic	16.00	16.45	10.99
Observations	567	567	567
Cohort 2014			
Cash incentives	0.16*** (0.05) [0.00]	0.20*** (0.06) [0.00]	0.11** (0.06) [0.06]
Constant	0.57*** (0.04)	0.63*** (0.04)	0.52*** (0.04)
F-statistic	8.51	10.83	3.88
Observations	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Visits total” corresponds to average visits per week during the 20 treatment weeks in the first year. “Visits 1st sem” corresponds to average visits per week during the 10 treatment weeks in the first semester. “Visits 2nd sem” corresponds to average visits per week during the 10 treatment weeks in the second semester.

## 5 Results

### 5.1 Educational outcomes

We now turn to the effects on educational outcomes - our primary outcomes of interest. All grade measures are standardized to have mean zero and standard deviation one within cohort. We find significant positive ITT effects on the business and law grades (Table 3, Columns (5) and (7)). The effect sizes are 0.15 standard deviations in the first semester and slightly larger in the second semester. The estimated coefficients for economics are only slightly smaller but statistically insignificant at the 10% level. Coefficients for average semester grades are also close to 0.15 standard deviations but marginally insignificant. The estimated coefficient on the likelihood that at least one grade is missing is negative, i.e. the incentives decrease the probability of a missing grade by five percentage points. As mentioned before, grades are missing when students do not take the exam. The estimated effects on the probability to pass the first year are close to zero and insignificant.

These findings suggest that the incentives and the resulting increase in exercise lead to substantial educational gains. Yet, it seems that only the part of the grade distribution that is not at the margin of passing the first year is affected. Quantile effects in Section 5.4 provide

Table 3: Effect of incentives on educational outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
<i>First semester</i>							
Cash incentives			-0.05*	0.13	0.15*	0.10	0.15*
			(0.03)	(0.08)	(0.08)	(0.08)	(0.08)
			[0.09]	[0.11]	[0.08]	[0.22]	[0.08]
Observations			567	567	567	567	567
<i>Second semester</i>							
Cash incentives	0.00	0.14*	-0.04	0.14	0.16*	0.13	0.17**
	(0.04)	(0.08)	(0.04)	(0.08)	(0.08)	(0.08)	(0.08)
	[1.00]	[0.10]	[0.33]	[0.11]	[0.07]	[0.13]	[0.05]
Observations	567	567	567	567	567	567	567

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first-year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first-year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. Results not including individuals with missing grades are reported in Table A.4 in the appendix.

further evidence for this conclusion.

We provide further results excluding students with missing grades in Table A.4 in the appendix. Under the assumption that students who were induced by the treatment to take the exam have on average lower ability than students who would have taken the exam anyway, the correlation between treatment status and student ability in the observed sample would be negative. The observed coefficient would thus be biased downward and should be interpreted as a lower bound. We still observe sizable positive coefficients. However, only the effect on the business grade in the second semester is statistically significant. The analysis of quantile effects in Section 5.4 excluding students with missing grades show positive and significant effects at higher quantiles of the grade distribution.

## 5.2 Rescaling the effect using TSLS

We use the incentives as an instrument and the weekly number of gym visits as the treatment in an instrumental variables analysis (see Section 2.2) for a discussion on the effect size. For interpretability, we also standardize the treatment variable to have mean zero and standard deviation one within cohort. Therefore, the obtained coefficient can be interpreted as the effect of a one standard deviation increase in average weekly gym visits. A standard deviation increase corresponds to 0.78, 0.84, and 0.85 visits per week, overall, in the first semester, and



in the second semester, respectively.<sup>28</sup> The results in Table 2 correspond to the first stage in the TSLS estimation (see Equation 3).

Table 4: IV estimates of effect of gym visits per week on grades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
	<i>First semester</i>						
Visits 1st sem.			-0.14	0.39	0.43*	0.31	0.44*
			(0.09)	(0.25)	(0.26)	(0.25)	(0.26)
Observations			567	567	567	567	567
	<i>Second semester</i>						
Visits total	0.00	0.42*	-0.13	0.41*	0.47*	0.39	0.50**
	(0.12)	(0.25)	(0.11)	(0.25)	(0.25)	(0.25)	(0.25)
Observations	567	567	567	567	567	567	567

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first-year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first-year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. The treatment variables are standardized to have mean zero and standard deviation one within cohort and randomization sample.

Table 4 presents the second stage results. We find that a one standard deviation increase in the number of weekly visits increases grades by about 0.42 standard deviations, which is a substantial gain. However, as discussed before, these coefficients should be interpreted with caution as the linear model might not represent the relationship between exercise and educational outcomes correctly.

<sup>28</sup>See Table A.3 in the appendix.

## 5.3 Mechanisms

Understanding the relevant mechanisms is important to assess external validity and policy implications of the results. To this end, we investigate the effects on several intermediate outcomes that could explain the link between exercise on campus and the positive effects on student educational outcomes. First, we analyze whether the incentives also changed characteristics of exercise beyond increasing the frequency. Second, we investigate how the incentives affect students' time-use in general. Finally, we investigate how the incentives affect health, health-related behavior, and well-being.

Several outcomes in this section stem from surveys conducted at the end of each semester. Since participation is voluntary, the number of observations is lower than in the analysis of the educational outcomes. Nonetheless, response behavior does not seem to be affected by the incentives as differences in response rates between the treated and control group are negligible and insignificant (see Table 1).

### 5.3.1 Characteristics of sports and exercise

Changes in the characteristics of students' physical activity beyond frequency could be one possible explanation for the improved educational outcomes. Studies in neuroscience suggest that the type and the timing of the physical activity might be relevant. For instance, since increased cognitive performance has been shown to occur immediately after exercising (Lambourne and Tomporowski, 2010), studying might be more effective if preceded by exercise. Most evidence from neuroscience points to positive effects of cardio-exercises on brain performance; moreover, resistance exercises (e.g., lifting weights) have also been shown to have positive effects. Note that other types of activity, such as team sports, have not been studied in this literature.

Table 5 explores this explanation. Columns (1) and (2) show self-reported activity overall and on campus measured by days that students exercised in a normal week. The estimated effects for overall activity are small and insignificant in both semesters. In contrast, on-campus activity is consistently positive in both semesters. The group that received incentives is 0.32 and 0.21 days more active in the first and second semester. However, only the estimated effect in the first semester is significantly different from zero at the 10% level. The results on self-reported on-campus activity are in line with the results using the scanner data. These results suggest that students substitute off-campus activities with on-campus activities.

This substitution could suggest that students are more likely to exercise in between classes or studying. We analyze at what time students increased on-campus exercise to explore this possibility. Columns (6)-(8) show the effects on weekly gym visits in the morning (from 7 to noon), afternoon (noon to 5 pm), and evening (5 pm to 10 pm). Treated students were more active over the entire day with the largest increases of 0.09 and 0.11 additional visits per week during the afternoon. Figure A.2 in the appendix depicts these effects for each hour. These results suggest that many students exercised during the day because of the incentives. This

integration of exercise into the class and study schedule could have increased learning efficiency in line with the evidence on the positive effects of exercise on cognitive functioning.

Table 5: Effect of incentives on characteristic of sport and exercise

	Self-reported activity			Sports type		Time of day		
	(1) Overall	(2) On-campus	(3) High intensity	(4) Weights/gym	(5) Other	(6) Morning	(7) Afternoon	(8) Evening
<i>First semester</i>								
Cash incentives	-0.02 (0.19) [0.89]	0.32* (0.16) [0.05]	0.04 (0.06) [0.56]	0.21*** (0.07) [0.00]	0.07 (0.05) [0.11]	0.08*** (0.03) [0.00]	0.11*** (0.03) [0.00]	0.08** (0.03) [0.01]
Constant	2.53*** (0.14)	1.51*** (0.12)	0.48*** (0.05)	0.33*** (0.04)	0.27*** (0.03)	0.15*** (0.02)	0.24*** (0.02)	0.20*** (0.02)
Observations	278	278	251	567	567	567	567	567
<i>Second semester</i>								
Cash incentives	-0.10 (0.18) [0.54]	0.21 (0.15) [0.18]	-0.12** (0.06) [0.04]	0.19*** (0.07) [0.01]	0.04 (0.05) [0.36]	0.07** (0.03) [0.02]	0.12*** (0.03) [0.00]	0.04 (0.03) [0.15]
Constant	2.85*** (0.14)	1.51*** (0.12)	0.59*** (0.04)	0.28*** (0.04)	0.23*** (0.03)	0.15*** (0.02)	0.19*** (0.02)	0.17*** (0.02)
Observations	315	315	296	567	567	567	567	567

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Overall days of sport” corresponds to the number of days a student does sports or exercise in a normal week during the semester. “On-campus days of sport” corresponds to the number of days a students does sports or exercise on-campus in a normal week during the semester. “High intensity” indicates whether or not a student reports to generally be exhausted with strong perspiration and breathlessness. “Weights”/gym” corresponds to weekly gym visits in this category. Students declared the category when scanning their student ID at the gym entrance. “Other” corresponds to activities other than weights/gym. “Time of day” corresponds to mean weekly gym visits in the morning (7 am to noon), afternoon (noon to 5 pm), and evening (5 pm to 10 pm).

Moreover, we explore which type of activity increased because of the incentives. Columns (4) and (5) show the effects on weekly gym visits by the category that students declared at the student id scanner. We look at the effects for the category “Weights/gym”, which corresponds to by far the most frequent category and includes individual fitness activities such as lifting weights, running on treadmills, rowing, etc. that also have been shown to affect brain activity. We pool other categories such as team sports, group fitness classes, etc. The results suggest that the increase in exercise is mostly driven by an increase in individual fitness activities on-campus. In this category, students visited the facilities 0.21 and 0.19 more often per week, whereas the effect on other types is much smaller and insignificant.

Finally, intensity of exercise might matter for the effect on educational outcomes. Column (3) investigates if the incentives changed intensity of exercise. Exercise is classified as ‘High intensity’ if students report to generally being exhausted with strong perspiration and breathlessness as opposed to minor or no exhaustion with little or no perspiration and breathlessness. The results do not depict a consistent pattern. While the estimate for the first semester is slightly positive but insignificant, the estimate in the second semester is negative and significantly different from zero, suggesting that the treatment leads to fewer high-intensity activities.

### 5.3.2 Time-use

Increased exercise might also change educational outcomes if it shifts how students allocate their time. The results of the previous section suggest that student’s overall activity does not change but that students integrate exercise into their day on campus. Consequently, students may use less time for physical activity overall and may spend more time on campus. If this excess time is used productively, it may explain the positive effects of the incentives on student’s grades.

Table 6: Effect of incentives on general time use

	Study related			Work & household		Leisure			
	(1) Class	(2) Studying	(3) Commute	(4) Work	(5) Chores	(6) Sport	(7) Digital Media	(8) Friends	(9) Other leisure
	<i>First semester</i>								
Cash incentives	0.26*	0.31	0.23**	0.17*	-0.11	0.01	-0.07	-0.15	-0.15
	(0.16)	(0.19)	(0.09)	(0.10)	(0.10)	(0.09)	(0.12)	(0.09)	(0.09)
	[0.09]	[0.10]	[0.01]	[0.10]	[0.28]	[0.89]	[0.59]	[0.10]	[0.12]
Constant	4.20***	2.21***	0.71***	0.23***	1.18***	0.92***	1.38***	1.23***	0.66***
	(0.10)	(0.13)	(0.05)	(0.05)	(0.08)	(0.06)	(0.09)	(0.07)	(0.07)
Observations	278	278	275	273	278	277	276	276	266
	<i>Second semester</i>								
Cash incentives	0.23	0.05	0.28***	0.08	0.03	0.15	0.08	-0.03	-0.10
	(0.15)	(0.20)	(0.10)	(0.11)	(0.09)	(0.09)	(0.11)	(0.11)	(0.14)
	[0.14]	[0.79]	[0.01]	[0.47]	[0.73]	[0.11]	[0.43]	[0.81]	[0.47]
Constant	3.42***	2.61***	0.79***	0.37***	1.05***	0.91***	1.42***	1.38***	0.88***
	(0.11)	(0.14)	(0.06)	(0.06)	(0.05)	(0.05)	(0.08)	(0.09)	(0.10)
Observations	315	311	308	308	315	309	311	308	296

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. Outcomes correspond to self-reported hours spent on the respective activities on a regular weekday.

We estimate the effects of the incentives on how many hours students engage in a specific activity on a regular weekday to explore this possibility. The outcome measures are self-reported in our online survey.<sup>29</sup> Table 6 shows the results. Students who received the incentives spend 0.26 and 0.23 hours (or 15.6 and 13.8 minutes) more in class (Column (1)) per day in the first and second semester. Yet, only the first estimate is significantly different from zero. Moreover, students who received the incentives spend significantly more time commuting (Column (3)) per day in both semesters (0.23 and 0.28 hours or 13.8 and 16.8 minutes). The incentives do not increase time spent with sports or exercise, which is in line with the previous results on exercise frequency.

These results suggest that students in fact spend more time on campus. The positive effects on hours commuting and hours in class could indicate that students commute more often to campus and go to class more frequently. Hence, the observed performance gains might be explained by a combination of students following class material more closely and increased learning efficiency.

<sup>29</sup>Results for time-use during the weekend can be found in Table A.8 in the appendix.

### 5.3.3 Health and well-being

Positive health benefits of physical activity have been extensively documented. Increases in health and well-being, and resulting gains in productivity could thus explain the educational gains in the treated group. Our outcomes for self-reported student health are an indicator for being in good or very good health, BMI, and an indicator for having felt physical pain often during the semester. For health-related behavior, we use indicators for healthy nutrition, smoking, and regular alcohol consumption. Healthy nutrition is classified as whether or not a student cares strongly or very strongly about a healthy diet. A student smokes if she smokes at the time of the survey. A student drinks regularly if she reports to drink either beer, wine, liquor, or mixed drinks regularly as opposed to every once in a while, rarely, or never.

Table 7 shows the results for health outcomes. We do not find significant effects on self-reported health and BMI but do on the probability of having felt physical pain recently.<sup>30</sup> The group that received the incentives is five percentage points more likely to have often felt physical pain. The effect is significant at the 10 percent level in the first and second semester. Most likely, this effect is the result of physical pain felt during or after exercising. We do not find any spillovers of more on-campus sports on other dimensions of health-related behavior, such as healthier nutrition, less smoking, or lower alcohol consumption (columns (4)-(6)). We also look at a set of measures for well-being, see Table A.7 in the appendix. However, we do not find any systematic effects for these outcome dimensions.

Overall, these results suggest that better health is unlikely to be the main driver of the improved educational outcomes. This conclusion is in line with the finding that students substitute off-campus activity with on-campus activity. Health and well-being would more likely be affected by an overall increase in physical activity. However, we acknowledge that measures of self-reported health may not be as effective as objective measures to detect small changes in students' health.

## 5.4 Effect Heterogeneity

In this section, we assess potential effect heterogeneity. First, we consider differences in the effects according to gender. Table 8 shows first stage effects separately for male and female students. The effects on overall activity (Column (1)) and by semester (Columns (2) and (3)) for female and male students are very similar and statistically significant. However, we find substantial heterogeneity when looking at the effects at different times of the day. Male students increase their activity primarily in the afternoon, while female students primarily increase participation in the evening (Columns (5) and (6)). The two groups increase activity similarly in morning hours (Column (4)).

Table 9 shows the ITT effects of the incentives on grades by gender. Male students experience

---

<sup>30</sup>This outcome is based on a five-point scale going from always to never. The indicator takes one if students choose 1 or 2 and zero otherwise.

Table 7: Effect of incentives on health and health behavior

	Health			Health behavior		
	(1) Good health	(2) Pain	(3) BMI	(4) Healthy nutrition	(5) Smokes	(6) Drinks regularly
<i>First semester</i>						
Cash incentives	-0.04 (0.05) [0.42]	0.05* (0.03) [0.09]	0.23 (0.43) [0.61]	0.06 (0.06) [0.31]	-0.01 (0.04) [0.75]	0.02 (0.05) [0.74]
Constant	0.81*** (0.03)	0.04** (0.02)	21.90*** (0.25)	0.43*** (0.04)	0.14*** (0.03)	0.24*** (0.04)
Observations	278	276	273	277	277	277
<i>Second semester</i>						
Cash incentives	-0.03 (0.04) [0.48]	0.05* (0.03) [0.09]	0.07 (0.28) [0.80]	0.00 (0.06) [0.99]	0.02 (0.04) [0.56]	0.02 (0.05) [0.68]
Constant	0.83*** (0.03)	0.05*** (0.02)	21.81*** (0.20)	0.50*** (0.04)	0.11*** (0.03)	0.25*** (0.04)
Observations	313	313	314	315	313	313

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Good health” indicates whether or not a student reports to be in good or very good health. “Pain” indicates whether a student often felt physical pain during the semester. It is based on a five-point scale going from always to never. The indicator takes one if students choose 1 or 2 and zero otherwise. “BMI” is calculated as follows:  $BMI_i = \frac{weight_i}{(height_i/100)^2}$  where  $weight_i$  is student  $i$ 's self-reported weight in kilogrammes and self-reported  $height_i$  the height in centimeter. “Healthy nutrition” indicates whether or not a student cares strongly or very strongly about a healthy diet. “Smokes” indicates whether or not a student smokes at the time of the survey. “Drinks regularly” indicates whether or not a student reports to drink either beer, wine, liquor, or mixed drinks regularly.

a positive effect on grades. The estimated effects are stronger with about 0.2 standard deviations and significant for all subjects in the second semester. Female students exhibit lower and insignificant effects.

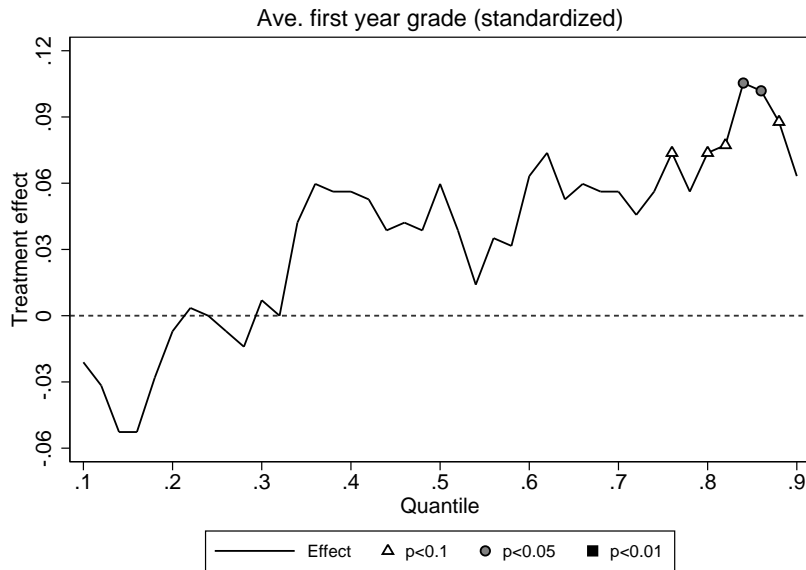
Female and male students similarly increase overall activity but only the grades of male students are positively affected by this increase. The argument that physical activity increases cognitive functioning after the exercise offers a tentative explanation for this discrepancy. Since male students increase exercise mostly during the day, they are likely to go to class or study afterward. Female students, however, increase activity predominantly in the evening, possibly after class or studying. Hence, we would expect an increase in physical activity to be more beneficial for male students than for female students. Another potential explanation can be found in Quadlin (2016). She suggests that female students spend more time engaged in academically oriented activities than male students in the beginning of their studies. While our time-use results by gender are too imprecise to be conclusive, they also suggest that female students spend more time studying and in class than their male counterparts. Hence, an improved learning efficiency or an increase in class attendance could have a lower marginal impact on the performance of female students.

Table 8: Effect of incentives on gym usage by student sex

	Semester			Time of day		
	(1) Visits total	(2) 1st sem	(3) 2nd sem	(4) Morning	(5) Afternoon	(6) Evening
<i>Male</i>						
Cash incentives	0.25*** (0.08) [0.00]	0.26*** (0.09) [0.00]	0.24*** (0.09) [0.01]	0.08** (0.04) [0.04]	0.14*** (0.04) [0.00]	0.03 (0.04) [0.36]
Constant	0.53*** (0.05)	0.60*** (0.06)	0.46*** (0.05)	0.16*** (0.02)	0.20*** (0.02)	0.17*** (0.02)
F-statistic	8.65	8.05	6.83	4.10	13.48	0.86
Observations	356	356	356	356	356	356
<i>Female</i>						
Cash incentives	0.27*** (0.10) [0.01]	0.32*** (0.10) [0.00]	0.23** (0.11) [0.04]	0.08** (0.04) [0.02]	0.08 (0.05) [0.13]	0.11** (0.05) [0.02]
Constant	0.58*** (0.07)	0.58*** (0.07)	0.58*** (0.08)	0.13*** (0.02)	0.25*** (0.04)	0.21*** (0.03)
F-statistic	7.58	9.17	4.08	5.43	2.27	5.89
Observations	211	211	211	211	211	211

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Visits total” corresponds to average visits per week during the 20 treatment weeks in the first year. “1st sem” corresponds to average visits per week during the 10 treatment weeks in the first semester. “2nd sem.” corresponds to average visits per week during the 10 treatment weeks in the second semester. “Time of day” corresponds to mean weekly gym visits in the morning (7 am to noon), afternoon (noon to 5 pm), and evening (5 pm to 10 pm).

Figure 2: Quantile reduced form effects on average first-year grade



Notes: The graphs depict the estimated coefficients at the corresponding quantiles. The outcome is the average grade over the first year course work weighted by the number of credits for each course. Missing grades are treated as missing. See Figure A.4 in the appendix for quantile regressions with missing grades imputed with the lowest grade.

Table 9: Effect of incentives on educational outcomes by student sex

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
<i>Male</i>							
<i>First semester</i>							
Cash incentives			-0.07** (0.03) [0.04]	0.17 (0.11) [0.11]	0.16 (0.10) [0.15]	0.12 (0.11) [0.27]	0.22** (0.10) [0.03]
Observations			356	356	356	356	356
<i>Second semester</i>							
Cash incentives	0.03 (0.05) [0.58]	0.19* (0.11) [0.08]	-0.07 (0.05) [0.16]	0.19* (0.11) [0.07]	0.24** (0.11) [0.03]	0.21** (0.11) [0.06]	0.21** (0.10) [0.05]
Observations	356	356	356	356	356	356	356
<i>Female</i>							
<i>First semester</i>							
Cash incentives			-0.00 (0.05) [1.00]	0.06 (0.14) [0.66]	0.13 (0.14) [0.39]	0.08 (0.13) [0.55]	0.02 (0.15) [0.88]
Observations			211	211	211	211	211
<i>Second semester</i>							
Cash incentives	-0.05 (0.07) [0.49]	0.05 (0.14) [0.69]	0.01 (0.06) [0.88]	0.05 (0.14) [0.72]	0.02 (0.14) [0.90]	-0.00 (0.13) [1.00]	0.09 (0.14) [0.52]
Observations	211	211	211	211	211	211	211

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first-year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first-year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization.

Next, we explore effect heterogeneity at different quantiles of the grade distribution to understand why average grades improve but persistence is unaffected. We focus on the weighted average grade in the first year. Figure A.6 shows the effects at quantiles ranging from the 10% to the 90%-quantile. Note that we exclude students with missing grades that are assigned the lowest grade by the university and thus drive the effect at the lower quantiles (see Figure A.4). The effects are increasing for higher quantiles. At the right tail of the grade distribution, the effects are approximately 0.1 standard deviations and significantly different from zero. These results suggest that students at the upper part of the grade distribution profit most from the intervention. Students at the margin of passing seem unaffected. Figure A.5 in the appendix shows quantile results for the average first- and second-semester grades.

## 6 Conclusion

Sport and exercise become increasingly important in modern societies as people’s lives are becoming more sedentary. University students would not be very active if they did not engage



in recreational physical activity as they spend most of their time sitting in lectures or studying. Even though evidence from neuroscience suggests positive short- and long-term effects of exercising on brain functioning, little is known about productivity enhancing effects of physical activity.

This paper contributes to the literature by identifying the effects of physical activity on students' educational performance. To be able to credibly identify causal effects, we randomize financial incentives to participate in on-campus recreational sports and exercise among first-year students at a Swiss university. First, we document a positive effect of the incentives on weekly usage of on-campus sports facilities, that is much stronger in our first cohort. In this cohort, the incentives increase weekly usage of on-campus sports facilities by roughly 47% from 0.55 to 0.81 visits per week.

For this cohort, we find a strong ITT effect of the incentives on grades. Estimates show an improvement in grades by about 0.15 standard deviations. Rescaling the effect using TSLS suggests a one standard deviation increase in the number of weekly gym visits increases average grades by about 0.42 standard deviations. Overall exercise frequency is unaffected while the frequency of on-campus exercise increases. Therefore, students appear to substitute other activities with on-campus activities. On-campus activity increases over the entire day but the increase is strongest in the afternoon. The incentives seem to trigger students to spend more time on campus and to integrate studying and exercise. These results indicate that integrating studying and exercising during the day may enhance the productivity of study time and thus improve students' performance. Furthermore, we find that estimated effects on self-reported health, health behavior, and well-being are mostly small and insignificant.

Our study suggests that physical activity is an under-explored factor in the human capital production function. In particular, on-campus exercise can contribute to the educational objectives of universities. Our results further indicate that not only the amount but also type and timing of physical activity might be important determinants of the effect on student performance.

This finding relates to the policy question of whether or not universities should foster physical activities of their students. The public has expressed increasing concerns that universities are spending excessively on recreational facilities at the expense of academic investments (Rubin, 2014; Lane, 2014). Jacob, McCall, and Stange (2013) document that non-profit 4-year colleges in the US spend on average 50 cents on recreational amenities for each dollar spent on academics. Against this critique universities, seem more interested in the benefits of these amenities (Bachman, 2014). While our results cannot speak to the benefits of extravagant facilities such as water parks or to the relative benefits compared to academic spending, they suggest that providing basic opportunities for students to exercise on campus contributes to the educational objectives of universities.

## References

- ACLAND, D. AND M. R. LEVY (2015): “Naiveté, Projection Bias, and Habit Formation in Gym Attendance,” *Management Science*, 61, 146–160.
- ANGRIST, J. D. AND G. W. IMBENS (1995): “Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity,” *Journal of the American Statistical Association*, 90, 431–442.
- BACHMAN, R. (2014): “Colleges Are Tracking When Students Work Out at Rec Centers,” *The Washington Post*.
- BARRON, J. M., B. T. EWING, AND G. R. WADDELL (2000): “The Effects of High School Athletic Participation on Education and Labor Market Outcomes,” *Review of Economics and Statistics*, 82, 409–421.
- CARRERA, M., H. ROYER, M. F. STEHR, AND J. R. SYDNOR (2017): “The Structure of Health Incentives: Evidence from a Field Experiment,” NBER Working Paper 23188, National Bureau of Economic Research.
- CAUDILL, S. B. AND J. E. LONG (2010): “Do Former Athletes Make Better Managers? Evidence from a Partially Adaptive Grouped-data Regression Model,” *Empirical Economics*, 39, 275–290.
- CHARNESS, G. AND U. GNEEZY (2009): “Incentives to Exercise,” *Econometrica*, 77, 909–931.
- CUNHA, F., J. J. HECKMAN, L. LOCHNER, AND D. V. MASTEROV (2006): “Interpreting the Evidence on Life Cycle Skill Formation,” *Handbook of the Economics of Education*, 1, 697–812.
- CURRIE, J. AND D. ALMOND (2011): “Human Capital Development Before Age Five,” *Handbook of Labor Economics*, 4, 1315–1486.
- DESIMONE, J. (2007): “Fraternity Membership and Binge Drinking,” *Journal of Health Economics*, 26, 950–967.
- EIDE, E. R. AND N. RONAN (2001): “Is Participation in High School Athletics an Investment or a Consumption Good? Evidence from High School and Beyond,” *Economics of Education Review*, 20, 431–442.
- EWING, B. T. (2007): “The Labor Market Effects of High School Athletic Participation Evidence From Wage and Fringe Benefit Differentials,” *Journal of Sports Economics*, 8, 255–265.
- FARRELL, L. AND M. A. SHIELDS (2002): “Investigating the Economic and Demographic Determinants of Sporting Participation in England,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 165, 335–348.

- FELFE, C., M. LECHNER, AND A. STEINMAYR (2016): “Sports and Child Development,” *PLOS ONE*, 11, 1–23.
- FRICKE, H., M. FRÖLICH, M. HUBER, AND M. LECHNER (2015a): “Endogeneity and Non-Response Bias in Treatment Evaluation: Nonparametric Identification of Causal Effects by Instruments,” IZA Discussion Paper 9428, Institute for the Study of Labor (IZA).
- FRICKE, H., J. GROGGER, AND A. STEINMAYR (2015b): “Does Exposure to Economics Bring New Majors to the Field? Evidence from a Natural Experiment,” NBER Working Paper 21130, National Bureau of Economic Research.
- FRÖLICH, M. (2007): “Nonparametric IV Estimation of Local Average Treatment Effects with Covariates,” *Journal of Econometrics*, 139, 35–75.
- HENDERSON, D. J., A. OLBRECHT, AND S. W. POLACHEK (2006): “Do Former College Athletes Earn More at Work? A Nonparametric Assessment,” *Journal of Human Resources*, 41, 558–577.
- HILLMAN, C. H., K. I. ERICKSON, AND A. F. KRAMER (2008): “Be Smart, Exercise Your Heart: Exercise Effects on Brain and Cognition,” *Nature Reviews Neuroscience*, 9, 58–65.
- HILLMAN, C. H., M. B. PONTIFEX, L. B. RAINE, D. M. CASTELLI, E. E. HALL, AND A. F. KRAMER (2009): “The Effect of Acute Treadmill Walking on Cognitive Control and Academic Achievement in Preadolescent Children,” *Neuroscience*, 159, 1044–1054.
- HÖTTING, K. AND B. RÖDER (2013): “Beneficial Effects of Physical Exercise on Neuroplasticity and Cognition,” *Neuroscience & Biobehavioral Reviews*, 37, 2243–2257.
- JACOB, B., B. MCCALL, AND K. M. STANGE (2013): “College as Country Club: Do Colleges Cater to Students’ Preferences for Consumption?” NBER Working Paper 18745, National Bureau of Economic Research.
- JANSSEN, I. AND A. G. LEBLANC (2010): “Systematic Review of the Health benefits of Physical Activity and Fitness in School-aged Children and Youth,” *The International Journal of Behavioral Nutrition and Physical Activity*, 7, 40.
- KAHNEMAN, D. AND A. TVERSKY (1979): “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica*, 47, 263–291.
- KATZ, L., J. KLING, AND J. LIEBMAN (2001): “Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment,” *Quarterly Journal of Economics*, 116, 607–654.
- KOSTEAS, V. D. (2012): “The Effect of Exercise on Earnings: Evidence from the NLSY,” *Journal of Labor Research*, 33, 225–250.

- LAMBOURNE, K. AND P. TOMPOROWSKI (2010): “The Effect of Exercise-induced Arousal on Cognitive Task Performance: A Meta-regression Analysis,” *Brain Research*, 1341, 12–24.
- LANE, C. (2014): “Colleges Drift Away from Their Academic Priorities,” *The Washington Post*, 2014-09-24.
- LECHNER, M. (2009): “Long-run Labour Market and Health Effects of Individual Sports Activities,” *Journal of Health Economics*, 28, 839–854.
- LEES, C. AND J. HOPKINS (2013): “Effect of Aerobic Exercise on Cognition, Academic Achievement, and Psychosocial Function in Children: A Systematic Review of Randomized Control Trials,” *Preventing Chronic Disease*, 10, 1–8.
- LINDO, J. M., I. D. SWENSEN, AND G. R. WADDELL (2013): “Alcohol and Student Performance: Estimating the Effect of Legal Access,” *Journal of Health Economics*, 32, 22–32.
- LIPSCOMB, S. (2007): “Secondary School Extracurricular Involvement and Academic Achievement: A Fixed Effects Approach,” *Economics of Education Review*, 26, 463–472.
- LONG, J. E. AND S. B. CAUDILL (1991): “The Impact of Participation in Intercollegiate Athletics on Income and Graduation,” *Review of Economics and Statistics*, 73, 525–531.
- PFEIFER, C. AND T. CORNELISSEN (2010): “The Impact of Participation in Sports on Educational Attainment—New Evidence from Germany,” *Economics of Education Review*, 29, 94–103.
- QUADLIN, N. Y. (2016): “Gender and Time Use in College: Converging or Diverging Pathways?” *Gender & Society*, 30, 361–385.
- REES, D. I. AND J. J. SABIA (2010): “Sports Participation and Academic Performance: Evidence from the National Longitudinal Study of Adolescent Health,” *Economics of Education Review*, 29, 751–759.
- ROOTH, D.-O. (2011): “Work Out or Out of Work — The Labor Market return to Physical Fitness and Leisure sports Activities,” *Labour Economics*, 18, 399–409.
- ROSENBAUM, P. R. (2002): “Covariance Adjustment in Randomized Experiments and Observational Studies,” *Statistical Science*, 17, 286–327.
- ROYER, H., M. F. STEHR, AND J. R. SYDNOR (2015): “Incentives, Commitments and Habit Formation in Exercise: Evidence from a Field Experiment with Workers at a Fortune-500 Company,” *American Economic Journal: Applied Economics*, 7, 51–84.
- RUBIN, C. (2014): “College Recreation Now Includes Pool Parties and River Rides,” *The New York Times*, 2014-09-19.

- RUBIN, D. B. (1980): “Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment,” *Journal of the American Statistical Association*, 75, 591–593.
- SCHNEIDER, S. AND S. BECKER (2005): “Prevalence of Physical Activity among the Working Population and Correlation with Work-Related Factors: Results from the First German National Health Survey,” *Journal of Occupational Health*, 47, 414–423.
- STEVENSON, B. (2010): “Beyond the Classroom: Using Title IX to Measure the Return to High School Sports,” *Review of Economics and Statistics*, 92, 284–301.
- STINEBRICKNER, R. AND T. R. STINEBRICKNER (2008): “The Causal Effect of Studying on Academic Performance,” *The B.E. Journal of Economic Analysis & Policy*, 8.
- STRONG, W. B., R. M. MALINA, C. J. BLIMKIE, S. R. DANIELS, R. K. DISHMAN, B. GUTIN, A. C. HERGENROEDER, A. MUST, P. A. NIXON, J. M. PIVARNIK, T. ROWLAND, S. TROST, AND F. TRUDEAU (2005): “Evidence Based Physical Activity for School-age Youth,” *The Journal of Pediatrics*, 146, 732–737.
- TAFRESCHI, D. AND P. THIEMANN (2016): “Doing It Twice, Getting It Right? The Effects of Grade Retention and Course Repetition in Higher Education,” *Economics of Education Review*, 55, 198–219.
- U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES (1996): *Physical Activity and Health: A Report of the Surgeon General*, Atlanta (GA): Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion.
- WARBURTON, D. E. R., C. W. NICOL, AND S. S. D. BREDIN (2006): “Health Benefits of Physical Activity: The Evidence,” *Canadian Medical Association Journal*, 174, 801–809.

# A Appendix

## A.1 Further descriptives statistics and results

Table A.1: Student characteristics by randomization blocks

Block no	Covariates							Treat. prob.	N
	Female	Swiss	Entry examen	German mother tongue	Law track	Sporty	Healthy		
<i>Cohort 13</i>									
1	0	0	0.67	1	0	0	0	0.56	9
2	0	0	0.8	0.98	0.02	0	1	0.51	41
3	0	0	0.86	0.89	0	1	0.94	0.49	35
4	0	1	0	0.96	0.04	0	0	0.5	28
5	0	1	0	0.91	0	0	1	0.5	129
6	0	1	0	1	1	0	1	0.5	10
7	0	1	0	0.86	0	1	0.98	0.5	103
8	1	0	0.79	0.95	0.11	0	0.79	0.53	19
9	1	0	0.82	1	0	1	0.91	0.55	11
10	0.97	1	0	0.9	0.13	0	0	0.5	30
11	1	1	0	0.8	0	0	1	0.51	90
12	1	1	0	0.86	1	0	1	0.43	7
13	1	1	0	0.93	0.15	1	0.91	0.51	55
<i>Cohort 14</i>									
1	0	0	0.4	0.89	0.02	-	-	0.51	47
2	0	0	1	0.95	0.01	-	-	0.5	100
3	0	1	0	0	0	-	-	0.51	41
4	0	1	0	1	0	-	-	0.5	245
5	0	1	0	1	1	-	-	0.47	19
6	1	0	0.56	0	0	-	-	0.56	9
7	1	0	0.36	1	0.2	-	-	0.52	25
8	1	0	1	0	0	-	-	0.5	6
9	1	0	0.96	1	0	-	-	0.48	25
10	1	1	0	0	0	-	-	0.5	26
11	1	1	0	1	0	-	-	0.5	194
12	1	1	0	1	1	-	-	0.47	17

Notes: The table presents the means of student characteristics used to create the randomization blocks. In the 2013 cohort, the covariates used were “Female”, “Swiss”, “Entry Exam”, “Sporty”, and “Healthy”. In the 2014 cohort, the covariates used were “Female”, “Swiss”, “Entry Exam”, “German mother tongue”, and “Law track”. “Law track” indicates whether or not a student is enrolled into the law specific first year curriculum, which requires two additional law courses instead of mathematics. “Entry exam” indicates if students had to take an admissions test. All non-Swiss students without a Swiss high school degree have to take an admissions test. “Sporty” and “Healthy” are based on the baseline survey in August 2013. “Sporty” is a binary indicator for whether student reported to exercise more than 3 times per week before starting the university. “Healthy” is a binary indicator for whether of not the student reported to be in good or very good health. “Treat. prob.” is the share of treated students in the respective block. In the case of odd numbers, the probability slightly deviates from 50%. The share of female students in block 10 in the 2013 cohort does not equal 100% because of data updating for one student after randomization.

Table A.2: Characteristics of included and excluded students

	<i>Cohort 2013</i>			<i>Cohort 2014</i>		
	Total	Excluded	Included	Total	Excluded	Included
Female (0/1)	0.34	0.31	0.37	0.35	0.28	0.40
Age (years)	20.05	20.17	19.91	20.01	20.08	19.95
Student aid contribution (0/1)	0.07	0.08	0.06	0.08	0.09	0.08
English track (0/1)	0.19	0.20	0.19	0.20	0.24	0.17
Law track (0/1)	0.06	0.06	0.06	0.05	0.04	0.06
Swiss national (0/1)	0.80	0.80	0.80	0.75	0.78	0.72
Entry exam (0/1)	0.14	0.12	0.16	0.19	0.15	0.22
German mother tongue (0/1)	0.85	0.81	0.90	0.84	0.79	0.88
High school degree from						
Canton St.Gallen (0/1)	0.14	0.13	0.15	0.12	0.11	0.13
Canton Zurich (0/1)	0.15	0.17	0.14	0.17	0.17	0.17
Other German speaking Canton (0/1)	0.43	0.40	0.46	0.38	0.37	0.39
Non-German speaking Canton (0/1)	0.10	0.12	0.08	0.11	0.15	0.08
Non-Swiss institution (0/1)	0.17	0.17	0.17	0.22	0.20	0.23
Gym visits pre treatment (avg per week)	0.32	0.30	0.35	0.42	0.40	0.43
Most frequent activity pre treatment						
Gym (0/1)	0.13	0.12	0.13	0.13	0.13	0.12
Other (0/1)	0.16	0.15	0.19	0.21	0.18	0.23
None (0/1)	0.71	0.73	0.68	0.67	0.69	0.65
Follow-up survey 1 (0/1)	0.27	0.12	0.49	0.30	0.16	0.40
Follow-up survey 2 (0/1)	0.35	0.20	0.56	0.20	0.10	0.28
Observations	1340	773	567	1293	539	754

Notes: The table presents the means of the covariates of students who were not selected (did not answer the baseline survey) for the experiment and of students who were selected (answered the baseline survey). “Student aid contribution” indicates whether or not a student donated to the student aid fund with their first tuition payment. “English track” indicates whether or not a student is enrolled into the English first year curriculum, which covers the same topics but is taught entirely in English. “Law track” indicates whether or not a student is enrolled into the law specific first year curriculum, which requires two additional law courses instead of mathematics. “Entry exam” indicates if students had to take an admissions test. All non-Swiss students without a Swiss high school degree have to take an admissions test. “Gym visits pre-treatment” is the average number of visits per week in the first two weeks of the semester. “Most frequent activity pre-treatment” indicate the most frequent type of activity during the first two weeks during the semester. “Follow-up survey 1” indicates participation in the follow-up survey at the end of the first semester. “Follow-up survey 2” indicates participation in the follow-up survey at the end of the second semester.

Table A.3: Summary statistics outcome and treatment variables

	<i>Cohort 2013</i>				<i>Cohort 2014</i>			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Avg 1st sem grade	4.18	0.96	1	5.7	4.05	1.01	1	5.73
Business 1 grade	3.76	0.93	1	5.5	3.75	1.03	1	5.5
Econ 1 grade	4.26	1.17	1	6	4.04	1.2	1	6
Law 1 grade	4.19	1.27	1	6	3.92	1.36	1	6
Avg 2nd sem grade	3.88	1.43	1	5.62	3.65	1.58	1	5.66
Business 2 grade	3.5	1.39	1	5.5	3.31	1.5	1	5.5
Econ 2 grade	3.73	1.48	1	6	3.62	1.66	1	6
Law 2 grade	3.61	1.53	1	6	3.38	1.61	1	6
Overall grade	4.01	1.19	1	5.62	3.81	1.28	1	5.68
Visits 1st sem	0.74	0.84	0	4.7	0.73	0.85	0	4.4
Visits 2nd sem	0.63	0.85	0	7.4	0.58	0.79	0	4.5
Visits total	0.68	0.78	0	5.85	0.65	0.75	0	4.1

Notes: “Avg 1st sem grade” and “Avg 2nd sem grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business 1 grade”, “Econ 1 grade”, “Law 1 grade”, “Business 2 grade”, “Econ 2 grade”, and “Law 2 grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. Grades are not standardized and range from 1 to 6 where 6 is the highest grade. Missing grades are set to one. “Visits 1st sem” corresponds to average visits per week during the 10 treatment weeks in the first semester. “Visits 2nd sem” corresponds to average visits per week during the 10 treatment weeks in the second semester. “Visits total” corresponds to average visits per week during the 20 treatment weeks in the first year.

Table A.4: Effect of incentives on educational outcomes (missing grades not included - 2013)

	(1)	(2)	(3)	(4)	(5)
	Overall grade	Avg semester grade	Business grade	Economics grade	Law grade
<i>First semester</i>					
Cash incentives	0.09 (0.10) [0.38]	0.04 (0.09) [0.65]	0.09 (0.09) [0.31]	0.03 (0.09) [0.76]	0.10 (0.09) [0.26]
Observations	404	500	531	535	514
<i>Second semester</i>					
Cash incentives	0.09 (0.10) [0.34]	0.14 (0.10) [0.14]	0.17* (0.09) [0.06]	0.10 (0.09) [0.28]	0.14 (0.09) [0.14]
Observations	404	410	452	456	445

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missing grades are not included.



Table A.5: Effect of incentives on educational outcomes for both cohorts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
Both cohorts							
<i>First semester</i>							
Cash incentives			-0.04**	0.08	0.10*	0.07	0.05
			(0.02)	(0.05)	(0.05)	(0.06)	(0.06)
			[0.03]	[0.16]	[0.07]	[0.15]	[0.32]
Observations			1321	1321	1321	1321	1321
<i>Second semester</i>							
Cash incentives	0.00	0.06	-0.01	0.05	0.05	0.05	0.06
	(0.03)	(0.06)	(0.03)	(0.06)	(0.06)	(0.06)	(0.06)
	[0.86]	[0.27]	[0.86]	[0.38]	[0.36]	[0.36]	[0.30]
Observations	1321	1321	1321	1321	1321	1321	1321

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization.

Table A.6: Effect of incentives on longterm outcomes (2013)

	(1)	(2)	(3)	(4)	(5)	(6)
	Avg 3rd sem	Avg 3rd sem (first year)	Avg 3rd sem (major)	Avg 4th sem	Avg 4th sem (first year)	Avg 4th sem (major)
Cash incentives	0.08	0.10	0.14	0.08	0.23	0.12
	(0.09)	(0.20)	(0.10)	(0.09)	(0.23)	(0.10)
	[0.33]	[0.64]	[0.17]	[0.37]	[0.31]	[0.24]
Constant	-0.04	-0.05	-0.07	-0.04	-0.12	-0.06
	(0.06)	(0.14)	(0.07)	(0.07)	(0.21)	(0.07)
Observations	486	97	393	486	80	391

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Avg 3rd sem” corresponds to weighted average grade in the third semester. “Avg 4th sem” corresponds to weighted average grade in the fourth semester. The averages are calculated over all attempted courses (Columns (1) and (4)), overall courses still pertaining to the first year curriculum if students did not pass the first year on the first attempt (Columns (2) and (5)), and overall courses in the major specific curriculum (Columns (3) and (6)). All grades are standardized to have mean zero and standard deviation one.

Table A.7: Effect of incentives on well-being (2013)

	Felt often...				Satisfied with...		
	(1) Pressure	(2) Energy	(3) Sad	(4) Calm	(5) Health	(6) Sleep	(7) Leisure
<i>First semester</i>							
Cash incentives	-0.03 (0.06) [0.68]	0.08 (0.06) [0.19]	0.04 (0.05) [0.46]	-0.04 (0.06) [0.53]	-0.08 (0.05) [0.15]	-0.02 (0.06) [0.80]	0.09 (0.06) [0.14]
Constant	0.49*** (0.04)	0.29*** (0.04)	0.26*** (0.04)	0.43*** (0.04)	0.76*** (0.04)	0.46*** (0.04)	0.39*** (0.04)
Observations	277	277	278	278	278	278	278
<i>Second semester</i>							
Cash incentives	-0.00 (0.06) [0.92]	-0.02 (0.05) [0.69]	0.03 (0.05) [0.62]	-0.00 (0.06) [0.97]	-0.04 (0.05) [0.42]	-0.06 (0.06) [0.27]	-0.01 (0.06) [0.90]
Constant	0.52*** (0.04)	0.35*** (0.04)	0.27*** (0.04)	0.40*** (0.04)	0.75*** (0.04)	0.51*** (0.04)	0.48*** (0.04)
Observations	315	314	314	315	315	314	315

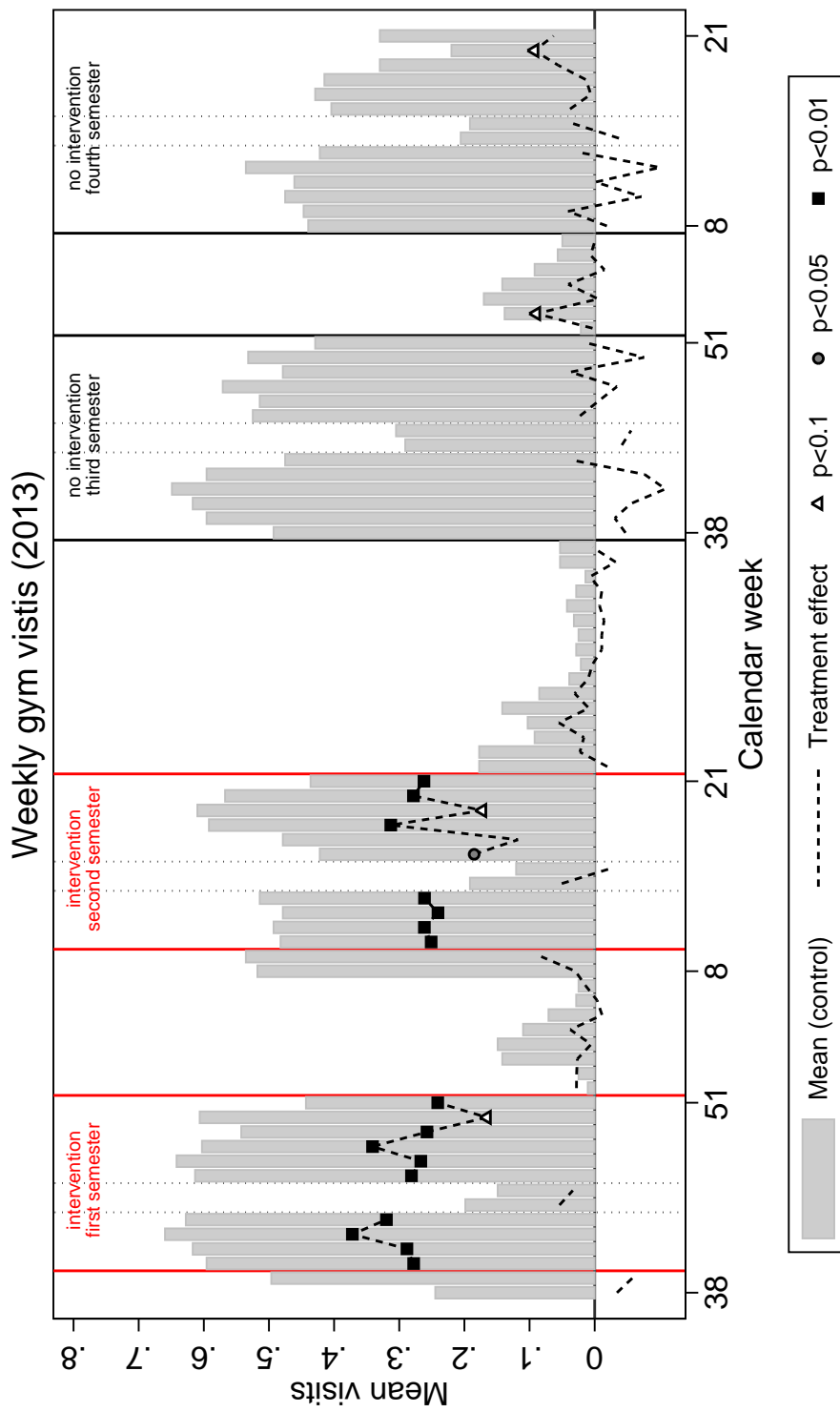
Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Pressure” indicates whether a student often felt pressure during the semester. “Energy” indicates whether a student often felt energetic during the semester. “Sad” indicates whether a student often felt sad during the semester. “Calm” indicates whether a student often felt calm during the semester. All four outcomes are based on a five-point scale going from ‘always’ to ‘never’. The respective indicator takes one if students answer 1 or 2 and zero otherwise. “Health”, “Sleep”, and “Leisure” indicate whether a student is satisfied in the respective domain. All three outcomes are based on a ten point scale ranging from ‘not at all satisfied’ to ‘completely satisfied’. The indicators take the value one if a student answers 8 or above.

Table A.8: Effect of incentives on general time use on weekends (2013)

	Study related			Work & household		Leisure			
	(1) Class	(2) Studying	(3) Commute	(4) Work	(5) Chores	(6) Sport	(7) Digital Media	(8) Friends	(9) Other leisure
<i>First semester</i>									
Cash incentives	0.05 (0.05) [0.44]	0.32 (0.23) [0.17]	0.21 (0.19) [0.26]	0.04 (0.19) [0.82]	0.15 (0.13) [0.24]	-0.14 (0.10) [0.18]	-0.00 (0.18) [0.99]	0.08 (0.20) [0.71]	0.11 (0.19) [0.60]
Constant	0.00 (0.00)	2.65*** (0.16)	0.77*** (0.13)	0.64*** (0.13)	1.35*** (0.08)	1.04*** (0.08)	2.21*** (0.11)	3.13*** (0.14)	1.65*** (0.14)
Observations	269	274	271	268	275	273	270	274	266
<i>Second semester</i>									
Cash incentives	0.32 (0.23) [0.17]	0.08 (0.09) [0.39]	0.04 (0.17) [0.81]	0.10 (0.17) [0.56]	-0.04 (0.13) [0.74]	-0.07 (0.11) [0.54]	0.03 (0.14) [0.85]	0.03 (0.19) [0.90]	0.00 (0.22) [0.98]
Constant	2.69*** (0.16)	0.09** (0.05)	0.78*** (0.12)	0.65*** (0.12)	1.34*** (0.08)	1.16*** (0.07)	2.15*** (0.10)	3.05*** (0.13)	2.06*** (0.17)
Observations	312	307	307	307	311	308	307	308	297

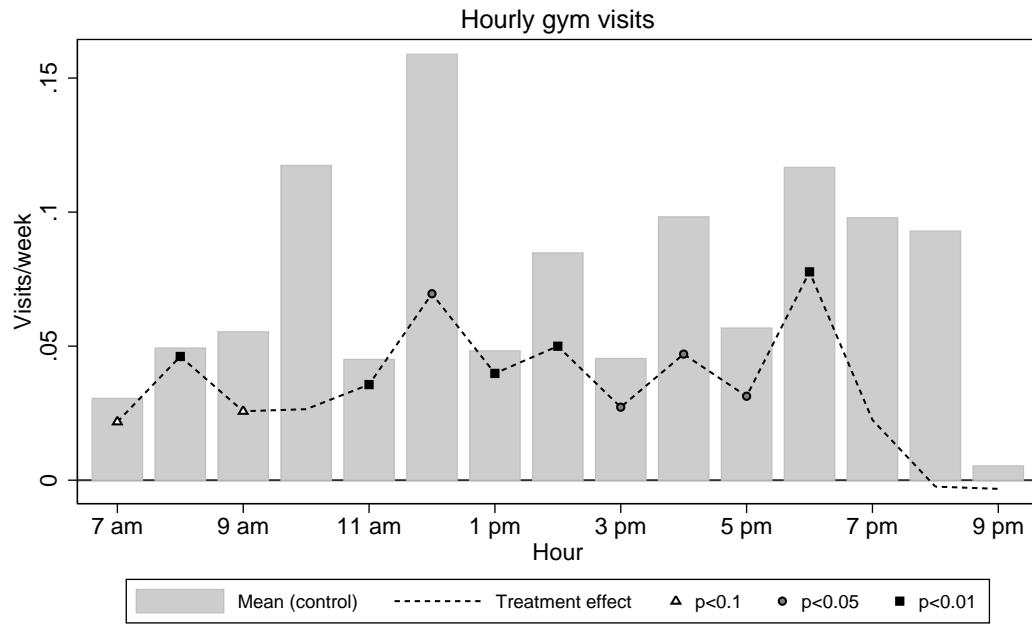
Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. Outcomes correspond to self-reported hours spent on the respective activities on a typical day on the weekend.

Figure A.1: Effect of incentives on gym usage by week beyond the first year



Notes: The figure depicts weekly mean visits of the treated and the control group, and their difference. p-values stem from a t-test on the equality of means. Calendar weeks 38 to 51 correspond to the first semester and calendar weeks 8 to 21 to the second semester. The respective exam periods are from calendar week 3 to 7 and 25 to 29. In the middle of each semester, a two-week break without classes takes place. No incentives are provided in these breaks. The incentives are only provided in the first semester and the second semesters.

Figure A.2: Effect of incentives on hourly gym visits (2013)



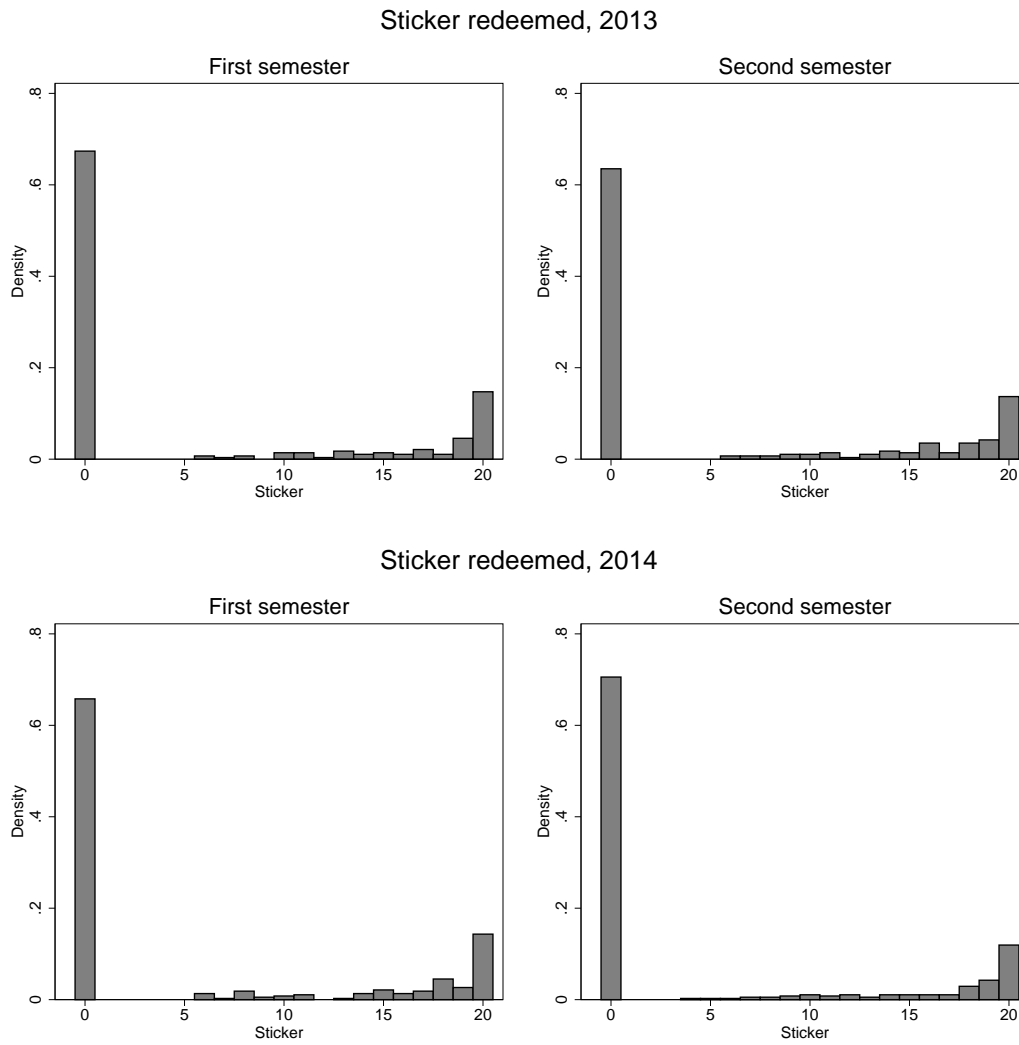
Notes: The figure depicts weekly mean visits of the control group the treatment effects of the incentives at different hours of the day. p-values stem from a t-test on the equality of means.

Table A.9: Effect of incentives on different types of sports (2013)

	On-campus (scanner)							Overall (self-reported)						
	(1) Team	(2) Individual	(3) Gym/fitness	(4) Cardio	(5) Group fitness	(6) Other	(7) None	(8) Team	(9) Individual	(10) Gym/fitness	(11) Cardio	(12) Group fitness	(13) Other	(14) None
Cash incentives	-0.01 (0.03) [0.69]	-0.03 (0.02) [0.14]	0.08** (0.04) [0.04]	0.01 (0.02) [0.68]	0.06* (0.03) [0.04]	-0.00 (0.01) [0.74]	-0.10*** (0.04) [0.01]	0.02 (0.04) [0.66]	0.02 (0.04) [0.68]	0.02 (0.06) [0.70]	-0.01 (0.05) [0.83]	-0.01 (0.04) [0.72]	0.02 (0.03) [0.59]	-0.05 (0.04) [0.17]
Constant	0.11*** (0.02)	0.07*** (0.01)	0.29*** (0.03)	0.04*** (0.01)	0.13*** (0.02)	0.01** (0.01)	0.36*** (0.03)	0.11*** (0.03)	0.11*** (0.03)	0.33*** (0.04)	0.18*** (0.03)	0.10*** (0.03)	0.05*** (0.02)	0.12*** (0.03)
Observations	567	567	567	567	567	567	567	277	277	277	277	277	277	277
						<i>First semester</i>								
Cash incentives	-0.01 (0.02) [0.74]	0.01 (0.02) [0.69]	0.08** (0.04) [0.04]	0.02 (0.02) [0.25]	0.01 (0.03) [0.79]	0.01 (0.01) [0.69]	-0.12*** (0.04) [0.00]	-0.00 (0.04) [0.93]	0.02 (0.03) [0.62]	-0.02 (0.05) [0.69]	0.00 (0.04) [0.94]	0.01 (0.04) [0.76]	0.01 (0.03) [0.81]	-0.01 (0.03) [0.68]
Constant	0.06*** (0.01)	0.04*** (0.01)	0.25*** (0.03)	0.04*** (0.01)	0.13*** (0.02)	0.01 (0.01)	0.48*** (0.03)	0.12*** (0.03)	0.09*** (0.02)	0.36*** (0.04)	0.19*** (0.03)	0.11*** (0.03)	0.06*** (0.02)	0.07*** (0.02)
Observations	567	567	567	567	567	567	567	315	315	315	315	315	315	315

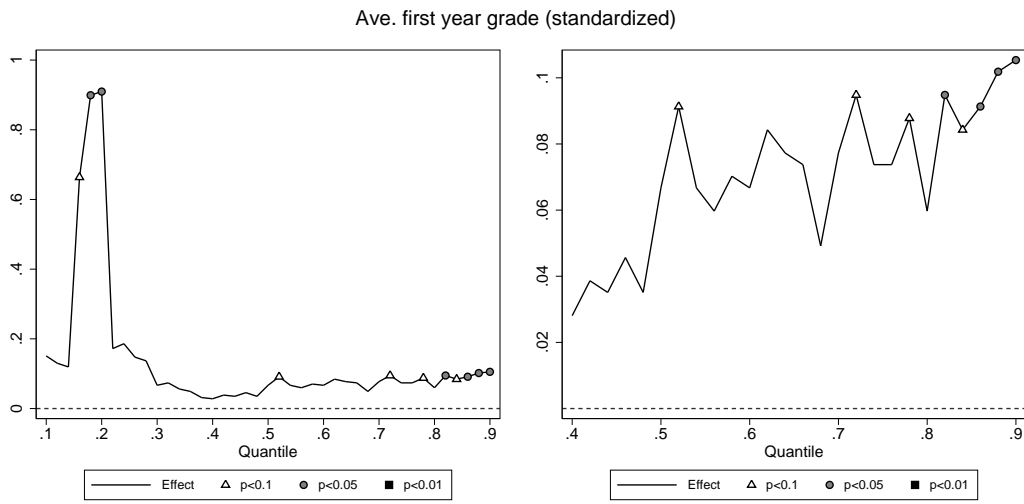
Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. The outcomes indicate whether or not the most frequent activity belongs to the respective category. See Appendix A.7 how activities are grouped. “Scanner” are outcome measures based on the scanner data at the gym entrance and thus only consider on-campus activities. “Self-reported” are outcome measures based on the survey data and correspond to activities overall.

Figure A.3: Redeemed sticker



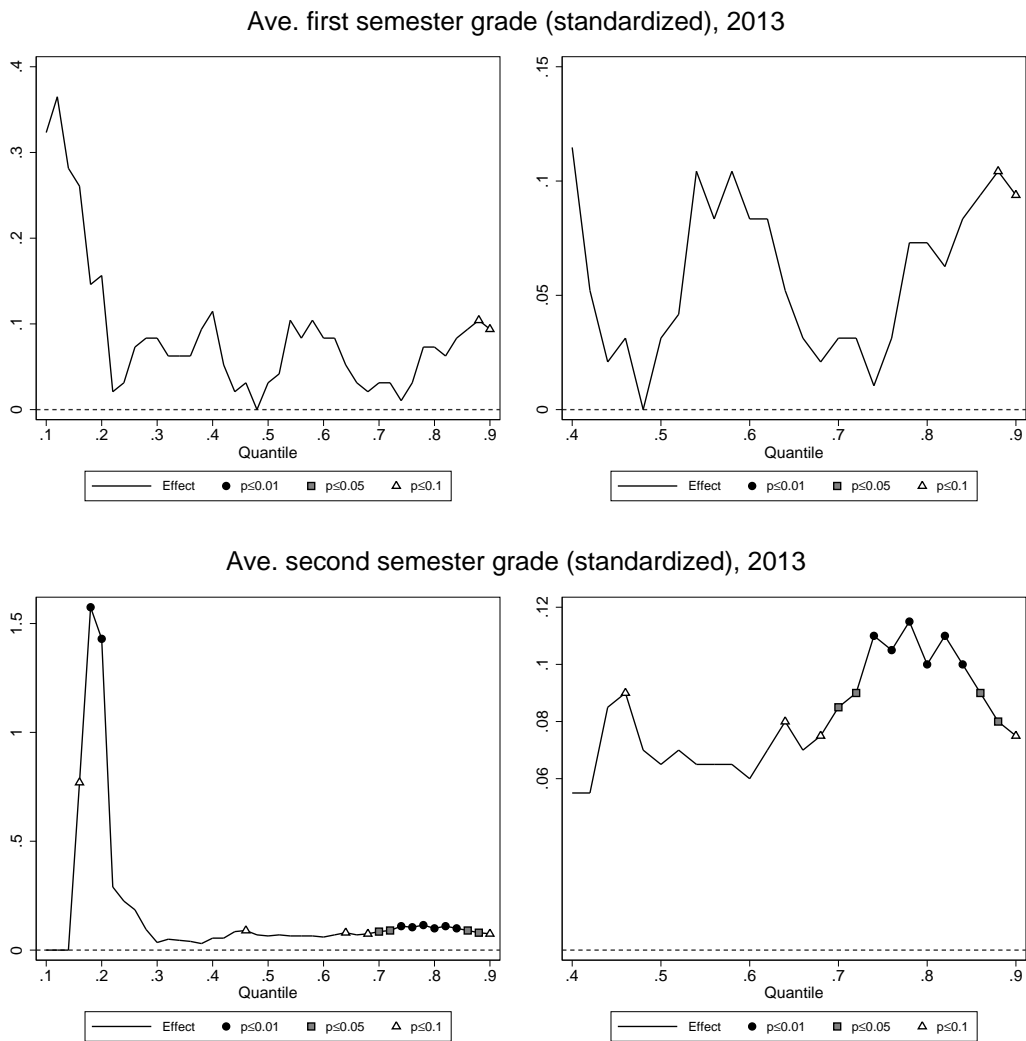
Notes: Histograms of redeemed stickers by cohort and semester.

Figure A.4: Quantile reduced form effects on average first-year grade (2013)



Notes: The graphs depict the estimated coefficients at the corresponding quantiles. The outcome is the average grade over the first year course work weighted by the number of credits for each course. The right panel shows a zoom on the quantiles above 0.4.

Figure A.5: Quantile reduced form effects on average first- and second-semester grades



Notes: The graphs depict the estimated coefficients at the corresponding quantiles. The outcome is the average grade over the first-semester (upper panel) and second-semester (lower panel) course work weighted by the number of credits for each course.



## A.2 Main results for 2014 cohort

Table A.10: Effect of incentives on educational outcomes (2014)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
<i>First semester</i>							
Cash incentives			-0.03 (0.03) [0.23]	0.03 (0.07) [0.59]	0.07 (0.07) [0.31]	0.05 (0.07) [0.48]	-0.02 (0.07) [0.84]
Observations			754	754	754	754	754
<i>Second semester</i>							
Cash incentives	0.01 (0.04) [0.88]	-0.00 (0.07) [0.96]	0.02 (0.03) [0.58]	-0.02 (0.07) [0.76]	-0.04 (0.07) [0.60]	-0.01 (0.07) [0.87]	-0.03 (0.07) [0.73]
Observations	754	754	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. Results not including individuals with missing grades are reported in Table A.4.

Table A.11: IV estimates of effect of gym visits per week on grades (2014)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
	<i>First semester</i>						
Visits 1st sem.			-0.14	0.14	0.30	0.22	-0.07
			(0.11)	(0.30)	(0.30)	(0.30)	(0.31)
Observations			754	754	754	754	754
	<i>Second semester</i>						
Visits total	0.04	-0.02	0.10	-0.09	-0.18	-0.06	-0.12
	(0.17)	(0.35)	(0.18)	(0.36)	(0.37)	(0.35)	(0.36)
Observations	754	754	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. The treatment variables are standardized to have mean zero and standard deviation one within cohort and randomization sample.

Table A.12: Effect of incentives on health and health behavior (2014)

	Health			Health behavior		
	(1) Good health	(2) Pain	(3) BMI	(4) Healthy nutrition	(5) Smokes	(6) Drinks regularly
<i>First semester</i>						
Cash incentives	-0.00 (0.04) [0.95]	0.02 (0.03) [0.57]	0.17 (0.26) [0.51]	0.04 (0.06) [0.51]	-0.04 (0.04) [0.36]	0.02 (0.05) [0.69]
Constant	0.84*** (0.03)	0.06*** (0.02)	21.63*** (0.17)	0.46*** (0.04)	0.14*** (0.03)	0.20*** (0.03)
Observations	298	297	297	299	300	297
<i>Second semester</i>						
Cash incentives	-0.03 (0.05) [0.62]	0.05 (0.03) [0.18]	0.47 (0.33) [0.16]	-0.05 (0.07) [0.49]	0.00 (0.05) [1.00]	0.01 (0.06) [0.86]
Constant	0.84*** (0.04)	0.04** (0.02)	21.50*** (0.21)	0.53*** (0.05)	0.12*** (0.03)	0.27*** (0.04)
Observations	208	210	208	210	208	209

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Good health” indicates whether or not a student reports to be in good or very good health. “Pain” indicates whether a student often felt physical pain during the semester. It is based on a five-point scale going from always to never. The indicator takes one if students choose 1 or 2 and zero otherwise. “BMI” is calculated as follows:  $BMI_i = \frac{weight_i}{(height_i/100)^2}$  where  $weight_i$  is student  $i$ 's self-reported weight in kilogrammes and self-reported  $height_i$  the height in centimeter. “Healthy nutrition” indicates whether or not a student cares strongly or very strongly about a healthy diet. “Smokes” indicates whether or not a student smokes at the time of the survey. “Drinks regularly” indicates whether or not a student reports to drink either beer, wine, liquor, or mixed drinks regularly.

Table A.13: Effect of incentives on characteristics of sport and exercise (2014)

	Self-reported activity			Sports type		Time of day		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Overall	On-campus	High intensity	Weights/gym	Other	Morning	Afternoon	Evening
<i>First semester</i>								
Cash incentives	0.00 (0.17) [0.99]	0.25 (0.16) [0.12]	-0.02 (0.06) [0.79]	0.08 (0.06) [0.13]	0.12*** (0.05) [0.01]	0.04 (0.03) [0.13]	0.11*** (0.03) [0.00]	0.05* (0.03) [0.07]
Constant	2.34*** (0.13)	1.40*** (0.12)	0.54*** (0.04)	0.35*** (0.04)	0.27*** (0.03)	0.15*** (0.02)	0.25*** (0.02)	0.23*** (0.02)
Observations	299	298	266	754	754	754	754	754
<i>Second semester</i>								
Cash incentives	0.48** (0.20) [0.02]	0.48** (0.20) [0.02]	-0.01 (0.07) [0.91]	0.05 (0.05) [0.38]	0.07* (0.04) [0.06]	0.04* (0.03) [0.11]	0.05* (0.03) [0.09]	0.02 (0.02) [0.33]
Constant	2.61*** (0.16)	1.54*** (0.15)	0.55*** (0.05)	0.31*** (0.04)	0.21*** (0.02)	0.15*** (0.02)	0.21*** (0.02)	0.16*** (0.02)
Observations	211	210	197	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Overall days of sport” corresponds to the number of days a student does sports or exercise in a normal week during the semester. “On-campus days of sport” corresponds to the number of days a students does sports or exercise on-campus in a normal week during the semester. “High intensity” indicates whether or not a student reports to generally be exhausted with strong perspiration and breathlessness. “Weights”/gym” corresponds to weekly gym visits in this category. Students declared the category when scanning their student ID at the gym entrance. “Other” corresponds to activities other than weights/gym. “Time of day” corresponds to mean weekly gym visits in the morning (7 am to noon), afternoon (noon to 5 pm), and evening (5 pm to 10 pm).

Table A.14: Effect of incentives on general time use (2014)

	Study related			Work & household		Leisure			
	(1) Class	(2) Studying	(3) Commute	(4) Work	(5) Chores	(6) Sport	(7) Digital Media	(8) Friends	(9) Other leisure
<i>First semester</i>									
Cash incentives	0.10 (0.18) [0.56]	-0.42** (0.21) [0.04]	-0.13 (0.11) [0.21]	0.23* (0.12) [0.06]	-0.01 (0.10) [0.92]	0.07 (0.08) [0.41]	-0.02 (0.14) [0.90]	-0.04 (0.13) [0.75]	0.06 (0.13) [0.68]
Constant	4.39*** (0.14)	2.56*** (0.16)	1.03*** (0.08)	0.29*** (0.06)	1.20*** (0.08)	0.90*** (0.06)	1.54*** (0.10)	1.41*** (0.10)	0.71*** (0.08)
Observations	300	298	300	291	300	298	297	297	284
<i>Second semester</i>									
Cash incentives	-0.06 (0.21) [0.77]	0.13 (0.21) [0.56]	-0.31** (0.13) [0.02]	-0.08 (0.17) [0.65]	-0.23 (0.14) [0.10]	-0.00 (0.13) [0.97]	-0.10 (0.19) [0.65]	-0.25 (0.17) [0.14]	-0.28* (0.14) [0.07]
Constant	3.78*** (0.16)	2.37*** (0.15)	1.06*** (0.11)	0.50*** (0.13)	1.28*** (0.13)	1.10*** (0.12)	1.74*** (0.15)	1.70*** (0.14)	0.96*** (0.12)
Observations	212	212	210	206	210	210	210	211	204

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. Outcomes correspond to self-reported hours spent on the respective activities on a regular weekday.

Table A.15: Effect of incentives on gym usage by student sex (2014)

	Semester			Time of day		
	(1) Visits total	(2) 1st sem	(3) 2nd sem	(4) Morning	(5) Afternoon	(6) Evening
<i>Male</i>						
Cash incentives	0.12 (0.07) [0.11]	0.15* (0.09) [0.08]	0.09 (0.08) [0.27]	0.05 (0.03) [0.13]	0.05 (0.04) [0.20]	0.02 (0.03) [0.54]
Constant	0.61*** (0.05)	0.66*** (0.06)	0.56*** (0.06)	0.16*** (0.02)	0.26*** (0.03)	0.19*** (0.02)
F-statistic	2.48	2.98	1.20	2.38	1.74	0.39
Observations	452	452	452	452	452	452
<i>Female</i>						
Cash incentives	0.22*** (0.08) [0.00]	0.28*** (0.08) [0.00]	0.16* (0.08) [0.07]	0.03 (0.03) [0.36]	0.12*** (0.04) [0.00]	0.07* (0.04) [0.07]
Constant	0.52*** (0.05)	0.58*** (0.05)	0.46*** (0.06)	0.14*** (0.02)	0.19*** (0.02)	0.19*** (0.02)
F-statistic	8.07	11.42	3.40	0.85	10.78	3.47
Observations	302	302	302	302	302	302

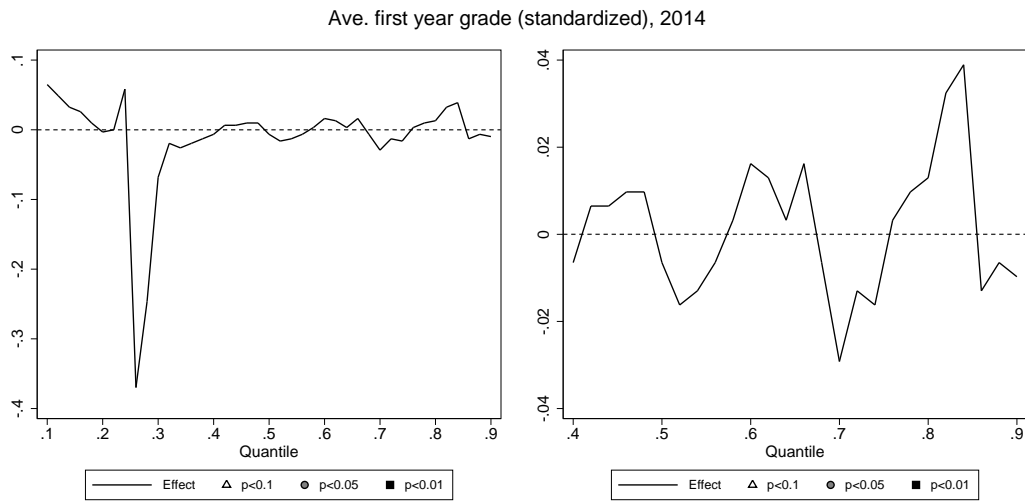
Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Visits total” corresponds to average visits per week during the 20 treatment weeks in the first year. “1st sem” corresponds to average visits per week during the 10 treatment weeks in the first semester. “2nd sem” corresponds to average visits per week during the 10 treatment weeks in the second semester.

Table A.16: Effect of incentives on educational outcomes by student sex (2014)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
<i>Male</i>							
<i>First semester</i>							
Cash incentives			-0.04 (0.03) [0.36]	-0.00 (0.10) [0.97]	0.06 (0.10) [0.50]	0.00 (0.10) [1.00]	-0.07 (0.10) [0.44]
Observations			452	452	452	452	452
<i>Second semester</i>							
Cash incentives	0.02 (0.05) [0.70]	-0.02 (0.09) [0.84]	-0.01 (0.04) [0.91]	-0.02 (0.09) [0.78]	-0.05 (0.09) [0.58]	-0.02 (0.09) [0.81]	-0.04 (0.09) [0.65]
Observations	452	452	452	452	452	452	452
<i>Female</i>							
<i>First semester</i>							
Cash incentives			-0.03 (0.04) [0.50]	0.09 (0.11) [0.40]	0.08 (0.11) [0.44]	0.13 (0.11) [0.23]	0.07 (0.11) [0.51]
Observations			302	302	302	302	302
<i>Second semester</i>							
Cash incentives	-0.01 (0.06) [0.90]	0.02 (0.11) [0.86]	0.07 (0.06) [0.27]	-0.02 (0.12) [0.87]	-0.02 (0.11) [0.89]	0.00 (0.11) [1.00]	-0.00 (0.12) [1.00]
Observations	302	302	302	302	302	302	302

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization.

Figure A.6: Quantile reduced form effects on average first year grade (2014)



Notes: The graphs depict the estimated coefficients at the corresponding quantiles. The outcome is the average grade over the first year course work weighted by the number of credits for each course.

### A.3 Exploration of cohort differences

In this section we explore the drivers of the differences in the effects of the incentives on participation in on-campus activities between the two cohorts. In the first cohort, the incentives increased participation on average by 0.26 visits per week; in the second cohort the effect was only 0.16 visits per week (Table 2). The differences in effect size between the two cohorts were even more pronounced in the second semester.

We analyze whether these differences are driven by observable characteristics that vary over the course of the study year. It might be, for example, that differences in weather or overall gym usage explain part of the treatment effect difference. Students who would have otherwise reacted to the incentives might not do so if the weather is too good or they perceive the facilities as over crowded. We first calculate the daily treatment effects. Then, we regress the daily treatment effect on a cohort dummy (Table A.17, Column (1)). The coefficient gives us the average difference in daily treatment effect between the two cohorts. This difference is -0.01 and is highly significant. In Column (2) we add variables capturing weather conditions (hours of sunshine, amount of precipitation, minimum temperature, and wind speed), a dummy indicating whether lectures were held on a given day, the share of the day a gym staff was present at the gym entrance, and month and weekday indicators. While these variables have strong explanatory power, evident by the increase in the adjusted  $R^2$  from 0.03 to 0.048, they leave the cohort dummy unaffected. In Column (3) we additionally include the overall number of gym visits on a given day by type of activity (all students who are not treated in the experiment). This further increases the adjusted  $R^2$  but again does not affect the overall cohort difference. These results suggest that observed time varying factors cannot explain the difference between the cohorts.

This finding suggests that the differences between cohorts must be explained by one or more of the following reasons. First, the environment at the university might have changed. For instance, tuition fees for foreign students were raised substantially in Fall 2014, the first semester of the 2014 cohort.<sup>31</sup> This increase might decrease the relative value of the incentives. It might further force foreign students to work more outside of the university, reducing the time they have available to respond to the incentives. Second, students that react to the incentives (compliers) in the first and second cohort might have distinct attributes and thus differ in how their educational outcomes are affected. Unfortunately, the literature provides little guidance on how to characterize complier populations in applications with continuous treatments. Third, the different effect sizes may be due to random variation, which would suggest that one or both of the results are outliers on the distribution of the estimated effects.

---

<sup>31</sup>The semester fees were raised from CHF 2,216 to CHF 3,326.



Table A.17: Explaining variation in daily treatment effects

	(1)		(2)		(3)	
	Raw		Set 1		Set 2	
Second cohort (0/1)	-0.01***	(0.00)	-0.01***	(0.00)	-0.01***	(0.00)
Weather						
Sunshine duration			-0.00	(0.00)	0.00	(0.00)
Precipitation			-0.03*	(0.02)	-0.04**	(0.02)
Air temperature min			-0.05	(0.03)	-0.05	(0.03)
Wind speed			0.05	(0.05)	0.04	(0.05)
Lecture period (0/1)			0.03***	(0.00)	0.03***	(0.00)
Control of ID scanner usage			-0.01***	(0.01)	-0.02**	(0.01)
Month						
January (0/1)						
February (0/1)			-0.01**	(0.00)	-0.01**	(0.00)
March (0/1)			0.01*	(0.01)	0.01	(0.01)
April (0/1)			0.00	(0.01)	0.01	(0.01)
May (0/1)			0.01	(0.01)	0.01**	(0.01)
September (0/1)			-0.02***	(0.01)	-0.02***	(0.01)
October (0/1)			0.02***	(0.01)	0.02***	(0.01)
November (0/1)			0.02***	(0.01)	0.02***	(0.01)
December (0/1)			0.00	(0.01)	0.00	(0.01)
Weekday						
Monday (0/1)						
Tuesday (0/1)			0.01***	(0.00)	-0.00	(0.01)
Wednesday (0/1)			0.01***	(0.00)	-0.00	(0.00)
Thursday (0/1)			0.02***	(0.00)	0.00	(0.01)
Friday (0/1)			0.00	(0.00)	-0.00	(0.01)
Saturday (0/1)			-0.01**	(0.00)	-0.01**	(0.01)
General facility usage						
Team (N visits)					0.04***	(0.01)
Individual (N visits)					0.02	(0.01)
Gym/fitness (N visits)					-0.00	(0.00)
Cardio (N visits)					0.01	(0.01)
Group fitness (N visits)					-0.02***	(0.01)
Other (N visits)					0.03	(0.02)
Constant	0.03***	(0.00)	0.00	(0.00)	0.01	(0.01)
Adjusted $R^2$	0.03		0.48		0.51	
Observations	407		407		407	

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. The dependent variable is the daily treatment effect. The sample contains all days on which the sports facilities were open between the first day of the first semester and the last day of the second semester. “Sunshine duration” corresponds to total number of minutes of sunshine. “Precipitation” is measured in millimeter of precipitation on one square meter. “Air temperature min” corresponds to the daily minimum temperature in degrees celsius. “Wind speed” corresponds to the maximum hourly mean of the day measured as meters per second. All weather indicators are rescaled with a factor of 1/100. Weather data are provided by the Federal Office of Meteorology and Climatology, MeteoSwiss. “Lecture period” is a dummy indicating whether lectures were held on a given

## A.4 Results with covariates

All of the following tables correspond one-to-one to the respective tables in the main text. However, the specifications in the appendix include the following set of controls: students' sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities. Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002).

Table A.18: Effect of incentives on gym usage (2013, with covariates)

	Semester			Time of day		
	(1) Visits total	(2) 1st sem	(3) 2nd sem	(4) Morning	(5) Afternoon	(6) Evening
Cash incentives	0.26*** (0.05) [0.00]	0.30*** (0.05) [0.00]	0.23*** (0.07) [0.00]	0.08*** (0.03) [0.00]	0.11*** (0.03) [0.00]	0.06** (0.03) [0.00]
Constant	0.43 (0.35)	0.58* (0.31)	0.29 (0.49)	-0.06 (0.18)	0.25* (0.13)	0.24 (0.17)
F-statistic						
Observations	566	566	566	566	566	566

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). "Visits total" corresponds to average visits per week during the 20 treatment weeks in the first year. "1st sem" corresponds to average visits per week during the 10 treatment weeks in the first semester. "2nd sem" corresponds to average visits per week during the 10 treatment weeks in the second semester. Covariates include students' sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.

Table A.19: Effect of incentives on educational outcomes (2013, with covariates)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
<i>First semester</i>							
Cash incentives			-0.04 (0.03) [0.12]	0.11 (0.08) [0.19]	0.12 (0.08) [0.14]	0.08 (0.08) [0.32]	0.14* (0.08) [0.09]
Observations			566	566	566	566	566
<i>Second semester</i>							
Cash incentives	-0.00 (0.04) [0.98]	0.12 (0.08) [0.15]	-0.04 (0.04) [0.26]	0.12 (0.08) [0.15]	0.14* (0.08) [0.09]	0.12 (0.08) [0.15]	0.15* (0.08) [0.07]
Observations	566	566	566	566	566	566	566

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. Covariates include students’ sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.

Table A.20: Effect of incentives on health and health behavior (2013, with covariates)

	Health			Health behavior		
	(1) Good health	(2) Pain	(3) BMI	(4) Healthy nutrition	(5) Smokes	(6) Drinks regularly
<i>First semester</i>						
Cash incentives	-0.07 (0.05) [0.18]	0.06** (0.03) [0.05]	0.31 (0.44) [0.53]	0.06 (0.06) [0.34]	-0.02 (0.04) [0.67]	0.02 (0.05) [0.66]
Constant	0.41 (0.27)	0.21 (0.17)	19.35*** (2.30)	0.14 (0.45)	-0.18 (0.26)	-0.15 (0.31)
Observations	278	276	273	277	277	277
<i>Second semester</i>						
Cash incentives	-0.05 (0.04) [0.32]	0.05* (0.03) [0.06]	0.15 (0.27) [0.57]	0.00 (0.06) [0.94]	0.04 (0.04) [0.26]	0.03 (0.05) [0.52]
Constant	0.97** (0.42)	0.18 (0.27)	20.08*** (1.91)	-0.10 (0.41)	-0.05 (0.34)	-0.10 (0.39)
Observations	313	313	314	315	313	313

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Good health” indicates whether or not a student reports to be in good or very good health. “Pain” indicates whether a student often felt physical pain during the semester. It is based on a five-point scale going from always to never. The indicator takes one if students choose 1 or 2 and zero otherwise. “BMI” is calculated as follows:  $BMI_i = \frac{weight_i}{(height_i/100)^2}$  where  $weight_i$  is student  $i$ 's self-reported weight in kilogrammes and self-reported  $height_i$  the height in centimeter. “Healthy nutrition” indicates whether or not a student cares strongly or very strongly about a healthy diet. “Smokes” indicates whether or not a student smokes at the time of the survey. “Drinks regularly” indicates whether or not a student reports to drink either beer, wine, liquor, or mixed drinks regularly. Covariates include students' sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, not speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.

Table A.21: Effect of incentives on well-being (2013, with covariates)

	Felt often...				Satisfied with...		
	(1) Pressure	(2) Energy	(3) Sad	(4) Calm	(5) Health	(6) Sleep	(7) Leisure
<i>First semester</i>							
Cash incentives	-0.04 (0.06) [0.54]	0.10* (0.06) [0.09]	0.04 (0.05) [0.47]	-0.03 (0.06) [0.55]	-0.08 (0.06) [0.16]	-0.02 (0.06) [0.79]	0.09 (0.06) [0.14]
Constant	0.90** (0.39)	0.31 (0.51)	0.49* (0.30)	0.31 (0.45)	0.99** (0.40)	0.07 (0.45)	-0.14 (0.38)
Observations	277	277	278	278	278	278	278
<i>Second semester</i>							
Cash incentives	-0.01 (0.06) [0.88]	-0.02 (0.06) [0.71]	0.02 (0.05) [0.77]	-0.01 (0.06) [0.89]	-0.06 (0.05) [0.25]	-0.04 (0.06) [0.50]	-0.01 (0.06) [0.81]
Constant	0.10 (0.40)	0.94*** (0.36)	0.68* (0.37)	0.74** (0.35)	0.42 (0.34)	-0.19 (0.40)	0.52 (0.40)
Observations	315	314	314	315	315	314	315

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Pressure” indicates whether a student often felt pressure during the semester. “Energy” indicates whether a student often felt energetic during the semester. “Sad” indicates whether a student often felt sad during the semester. “Calm” indicates whether a student often felt calm during the semester. All four outcomes are based on a five-point scale going from ‘always’ to ‘never’. The respective indicator takes one if students answer 1 or 2 and zero otherwise. “Health”, “Sleep”, and “Leisure” indicate whether a student is satisfied in the respective domain. All three outcomes are based on a ten point scale ranging from ‘not at all satisfied’ to ‘completely satisfied’. The indicators take the value one if a student answers 8 or above. Covariates include students’ sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.

Table A.22: Effect of incentives on characteristics of sports and exercise (2013, with covariates)

	Self-reported activity			Sports type		Time of day		
	(1) Overall	(2) On-campus	(3) High intensity	(4) Weights/gym	(5) Other	(6) Morning	(7) Afternoon	(8) Evening
<i>First semester</i>								
Cash incentives	0.09 (0.18) [0.60]	0.30** (0.15) [0.04]	0.02 (0.07) [0.73]	0.24*** (0.06) [0.00]	0.05 (0.04) [0.25]	0.09*** (0.03) [0.00]	0.12*** (0.03) [0.00]	0.09*** (0.03) [0.00]
Constant	3.82*** (1.31)	1.14 (0.93)	0.53 (0.46)	0.15 (0.34)	0.43 (0.26)	0.02 (0.18)	0.27 (0.16)	0.29 (0.18)
Observations	278	278	251	566	566	566	566	566
<i>Second semester</i>								
Cash incentives	-0.05 (0.18) [0.77]	0.21 (0.14) [0.12]	-0.14** (0.06) [0.01]	0.22*** (0.06) [0.00]	0.01 (0.04) [0.81]	0.08** (0.03) [0.01]	0.11*** (0.03) [0.00]	0.04 (0.03) [0.17]
Constant	3.18*** (1.20)	1.30 (1.07)	0.70 (0.45)	-0.14 (0.50)	0.43** (0.21)	-0.14 (0.21)	0.23 (0.16)	0.20 (0.22)
Observations	315	315	296	566	566	566	566	566

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Overall days of sport” corresponds to the number of days a student does sports or exercise in a normal week during the semester. “On-campus days of sport” corresponds to the number of days a students does sports or exercise on-campus in a normal week during the semester. “High intensity” indicates whether or not a student reports to generally be exhausted with strong perspiration and breathlessness. “Weights”/gym” corresponds to weekly gym visits in this category. Students declared the category when scanning their student ID at the gym entrance. “Other” corresponds to activities other than weights/gym. “Time of day” corresponds to mean weekly gym visits in the morning (7 am to noon), afternoon (noon to 5 pm), and evening (5 pm to 10 pm). Covariates include students’ sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.

Table A.23: Effect of incentives on different types of sports (2013, with covariates)

	On-campus (scanner)							Overall (self-reported)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Team	Individual	Gym/fitness	Cardio	Group fitness	Other	None	Team	Individual	Gym/fitness	Cardio	Group fitness	Other	None
Cash incentives	-0.02 (0.03)	-0.04** (0.02)	0.09** (0.04)	0.01 (0.02)	0.05* (0.03)	-0.01 (0.01)	-0.08** (0.03)	0.03 (0.04)	0.01 (0.04)	0.01 (0.06)	0.00 (0.04)	-0.01 (0.03)	0.01 (0.03)	-0.05 (0.04)
Constant	0.39** (0.18)	-0.12 (0.17)	0.22 (0.26)	0.10 (0.09)	-0.15 (0.20)	0.02 (0.03)	0.55*** (0.21)	0.38 (0.28)	0.21 (0.23)	0.44 (0.32)	-0.08 (0.38)	-0.27 (0.34)	0.10 (0.14)	0.23 (0.17)
Observations	566	566	566	566	566	566	566	277	277	277	277	277	277	277
Cash incentives	-0.01 (0.02)	0.01 (0.02)	0.09** (0.04)	0.02 (0.02)	-0.01 (0.03)	0.01 (0.01)	-0.10*** (0.04)	0.00 (0.03)	0.02 (0.03)	-0.03 (0.05)	0.01 (0.04)	0.01 (0.03)	0.01 (0.03)	-0.01 (0.03)
Constant	0.14 (0.14)	0.16* (0.10)	0.16 (0.24)	-0.06 (0.09)	0.32** (0.15)	-0.01 (0.03)	0.28 (0.27)	0.23 (0.27)	0.51* (0.29)	0.78** (0.32)	-0.52 (0.39)	-0.07 (0.26)	-0.06 (0.21)	0.13 (0.24)
Observations	566	566	566	566	566	566	566	315	315	315	315	315	315	315

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). The outcomes indicate whether or not the most frequent activity belongs to the respective category. See Appendix A.7 how activities are grouped. "Scanner" are outcome measures based on the scanner data at the gym entrance and thus only consider on-campus activities. "Self-reported" are outcome measures based on the survey data and correspond to activities overall. Covariates include students' sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.

Table A.24: Effect of incentives on general time use (2013, with covariates)

	Study related			Work & household		Leisure			
	(1) Class	(2) Studying	(3) Commute	(4) Work	(5) Chores	(6) Sport	(7) Digital Media	(8) Friends	(9) Other leisure
<i>First semester</i>									
Cash incentives	0.27 (0.17) [0.09]	0.26 (0.18) [0.15]	0.25*** (0.09) [0.00]	0.19* (0.10) [0.05]	-0.10 (0.10) [0.30]	0.03 (0.09) [0.72]	-0.08 (0.13) [0.52]	-0.17* (0.10) [0.07]	-0.12 (0.10) [0.20]
Constant	2.01 (1.59)	-1.99 (1.22)	0.53 (0.83)	-1.73 (2.43)	0.33 (0.91)	0.31 (0.94)	1.47** (0.68)	-0.01 (0.52)	1.12* (0.59)
Observations	278	278	275	273	278	277	276	276	266
<i>Second semester</i>									
Cash incentives	0.28* (0.15) [0.06]	0.05 (0.19) [0.78]	0.29*** (0.09) [0.00]	0.10 (0.11) [0.39]	0.03 (0.09) [0.80]	0.17* (0.09) [0.07]	0.10 (0.10) [0.32]	-0.02 (0.11) [0.90]	-0.04 (0.12) [0.73]
Constant	2.11 (1.47)	5.41*** (1.68)	2.28*** (0.64)	0.74 (0.72)	1.47** (0.65)	1.68*** (0.55)	3.36*** (0.78)	1.14 (0.95)	2.99*** (1.13)
Observations	315	311	308	308	315	309	311	308	296

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). Outcomes correspond to self-reported hours spent on the respective activities on a regular weekday. Covariates include students' sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, e) other activities.



Table A.25: IV estimates of effect of gym visits per week on educational outcomes (2013, with covariates)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
<i>First semester</i>							
Visits 1st sem.			-0.12 (0.08)	0.31 (0.23)	0.35 (0.23)	0.23 (0.23)	0.41* (0.23)
Observations			566	566	566	566	566
<i>Second semester</i>							
Visits total	-0.00 (0.12)	0.36 (0.24)	-0.12 (0.11)	0.36 (0.24)	0.42* (0.24)	0.34 (0.24)	0.45* (0.24)
Observations	566	566	566	566	566	566	566

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. The treatment variables are standardized to have mean zero and standard deviation one within cohort and randomization sample. Covariates include students’ sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, e) other activities.

## A.5 Results with randomization block fixed effects

Table A.26: Effect of incentives on gym usage (2013, with block FE)

	Semester			Time of day		
	(1) Visits total	(2) 1st sem	(3) 2nd sem	(4) Morning	(5) Afternoon	(6) Evening
Cash incentives	0.26*** (0.06) [0.00]	0.28*** (0.07) [0.00]	0.23*** (0.07) [0.00]	0.08*** (0.03) [0.00]	0.12*** (0.03) [0.00]	0.06** (0.03) [0.02]
Constant	0.21 (0.14)	0.28 (0.18)	0.15 (0.13)	0.00 (0.02)	0.08 (0.07)	0.13 (0.08)
F-statistic						
Observations	567	567	567	567	567	567

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Visits total” corresponds to average visits per week during the 20 treatment weeks in the first year. “1st sem” corresponds to average visits per week during the 10 treatment weeks in the first semester. “2nd sem” corresponds to average visits per week during the 10 treatment weeks in the second semester. All regressions include randomization block fixed effects.

Table A.27: Effect of incentives on educational outcomes (2013, with block FE)

	(1) Passed first year	(2) Overall grade	(3) Missing grade	(4) Avg semester grade	(5) Business grade	(6) Economics grade	(7) Law grade
	<i>First semester</i>						
Cash incentives			-0.05* (0.03) [0.08]	0.13 (0.08) [0.15]	0.15* (0.08) [0.08]	0.10 (0.08) [0.23]	0.15* (0.08) [0.08]
Observations			567	567	567	567	567
	<i>Second semester</i>						
Cash incentives	0.00 (0.04) [1.00]	0.14* (0.08) [0.09]	-0.04 (0.04) [0.26]	0.14 (0.08) [0.10]	0.16* (0.08) [0.07]	0.13 (0.08) [0.13]	0.17** (0.08) [0.04]
Observations	567	567	567	567	567	567	567

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. All regressions include randomization block fixed effects.

Table A.28: Effect of incentives on health and health behavior (2013, with block FE)

	Health			Health behavior		
	(1) Good health	(2) Pain	(3) BMI	(4) Healthy nutrition	(5) Smokes	(6) Drinks regularly
<i>First semester</i>						
Cash incentives	-0.02 (0.05) [0.67]	0.04 (0.03) [0.12]	0.18 (0.42) [0.65]	0.08 (0.06) [0.20]	-0.01 (0.04) [0.73]	0.01 (0.05) [0.80]
Constant	0.41* (0.22)	-0.02 (0.01)	22.71*** (0.98)	0.17 (0.17)	0.01 (0.02)	0.19 (0.18)
Observations	278	276	273	277	277	277
<i>Second semester</i>						
Cash incentives	-0.04 (0.04) [0.36]	0.06** (0.03) [0.05]	0.03 (0.27) [0.93]	0.00 (0.06) [0.96]	0.03 (0.04) [0.40]	0.02 (0.05) [0.71]
Constant	1.02*** (0.03)	-0.03 (0.03)	23.82*** (0.24)	0.50 (0.36)	0.49 (0.35)	0.49 (0.37)
Observations	313	313	314	315	313	313

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Good health” indicates whether or not a student reports to be in good or very good health. “Pain” indicates whether a student often felt physical pain during the semester. It is based on a five-point scale going from always to never. The indicator takes one if students choose 1 or 2 and zero otherwise. “BMI” is calculated as follows:  $BMI_i = \frac{weight_i}{(height_i/100)^2}$  where  $weight_i$  is student  $i$ 's self-reported weight in kilograms and self-reported  $height_i$  the height in centimeter. “Healthy nutrition” indicates whether or not a student cares strongly or very strongly about a healthy diet. “Smokes” indicates whether or not a student smokes at the time of the survey. “Drinks regularly” indicates whether or not a student reports to drink either beer, wine, liquor, or mixed drinks regularly. All regressions include randomization block fixed effects.

Table A.29: Effect of incentives on well-being (2013, with block FE)

	Felt often...				Satisfied with...		
	(1) Pressure	(2) Energy	(3) Sad	(4) Calm	(5) Health	(6) Sleep	(7) Leisure
<i>First semester</i>							
Cash incentives	-0.02 (0.06) [0.72]	0.08 (0.05) [0.14]	0.04 (0.05) [0.45]	-0.03 (0.06) [0.58]	-0.07 (0.05) [0.20]	-0.01 (0.06) [0.92]	0.11* (0.06) [0.06]
Constant	0.41* (0.23)	0.17 (0.19)	-0.02 (0.02)	0.61*** (0.22)	0.43* (0.22)	0.20 (0.19)	0.16 (0.20)
Observations	277	277	278	278	278	278	278
<i>Second semester</i>							
Cash incentives	0.01 (0.06) [0.90]	-0.02 (0.05) [0.72]	0.02 (0.05) [0.67]	-0.01 (0.05) [0.80]	-0.05 (0.05) [0.39]	-0.07 (0.06) [0.18]	-0.01 (0.06) [0.81]
Constant	-0.00 (0.03)	0.01 (0.03)	-0.01 (0.03)	1.01*** (0.03)	0.52 (0.35)	0.54 (0.34)	0.51 (0.36)
Observations	315	314	314	315	315	314	315

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Pressure” indicates whether a student often felt pressure during the semester. “Energy” indicates whether a student often felt energetic during the semester. “Sad” indicates whether a student often felt sad during the semester. “Calm” indicates whether a student often felt calm during the semester. All four outcomes are based on a five-point scale going from ‘always’ to ‘never’. The respective indicator takes one if students answer 1 or 2 and zero otherwise. “Health”, “Sleep”, and “Leisure” indicate whether a student is satisfied in the respective domain. All three outcomes are based on a ten point scale ranging from ‘not at all satisfied’ to ‘completely satisfied’. The indicators take the value one if a student answers 8 or above. All regressions include randomization block fixed effects.

Table A.30: Effect of incentives on characteristics of sport and exercise (2013, with block FE)

	Self-reported activity			Sports type		Time of day		
	(1) Overall	(2) On-campus	(3) High intensity	(4) Weights/gym	(5) Other	(6) Morning	(7) Afternoon	(8) Evening
<i>First semester</i>								
Cash incentives	-0.03 (0.17) [0.83]	0.31** (0.15) [0.04]	0.05 (0.06) [0.47]	0.21*** (0.07) [0.00]	0.07 (0.05) [0.10]	0.09*** (0.03) [0.00]	0.12*** (0.03) [0.00]	0.08** (0.03) [0.01]
Constant	1.81*** (0.45)	1.08** (0.47)	0.49* (0.25)	0.25 (0.18)	0.03 (0.06)	-0.03 (0.02)	0.09 (0.08)	0.21 (0.14)
Observations	278	278	251	567	567	567	567	567
<i>Second semester</i>								
Cash incentives	-0.20 (0.16) [0.20]	0.17 (0.15) [0.26]	-0.13** (0.06) [0.02]	0.20*** (0.06) [0.01]	0.04 (0.04) [0.41]	0.07** (0.03) [0.03]	0.12*** (0.03) [0.00]	0.04 (0.03) [0.14]
Constant	2.10 (1.38)	0.41 (0.31)	0.00 (.)	0.11 (0.13)	0.04 (0.06)	0.03 (0.03)	0.07 (0.08)	0.05 (0.04)
Observations	315	315	296	567	567	567	567	567

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Overall days of sport” corresponds to the number of days a student does sports or exercise in a normal week during the semester. “On-campus days of sport” corresponds to the number of days a student does sports or exercise on-campus in a normal week during the semester. “High intensity” indicates whether or not a student reports to generally be exhausted with strong perspiration and breathlessness. “Weights”/gym” corresponds to weekly gym visits in this category. Students declared the category when scanning their student ID at the gym entrance. “Other” corresponds to activities other than weights/gym. “Time of day” corresponds to mean weekly gym visits in the morning (7 am to noon), afternoon (noon to 5 pm), and evening (5 pm to 10 pm). All regressions include randomization block fixed effects.

Table A.31: Effect of incentives on different types of sports (2013, with block FE)

	On-campus (scanner)							Overall (self-reported)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	
	Team	Individual	Gym/fitness	Cardio	Group fitness	Other	None	Team	Individual	Gym/fitness	Cardio	Group fitness	Other	None	
Cash incentives	-0.01 (0.03) [0.66]	-0.03 (0.02) [0.15]	0.08** (0.04) [0.04]	0.01 (0.02) [0.69]	0.05** (0.03) [0.04]	-0.00 (0.01) [0.73]	-0.10*** (0.04) [0.01]	0.02 (0.04) [0.61]	0.01 (0.04) [0.68]	0.01 (0.06) [0.84]	-0.01 (0.05) [0.92]	-0.00 (0.03) [0.94]	0.01 (0.03) [0.66]	0.01 (0.03) [0.12]	-0.05 (0.04) [0.22]
Constant	0.12 (0.11)	0.13 (0.11)	0.40** (0.16)	-0.01 (0.01)	-0.03* (0.02)	0.00 (0.01)	0.39** (0.16)	-0.01 (0.02)	0.39* (0.23)	0.19 (0.18)	0.20 (0.19)	0.00 (0.01)	-0.01 (0.01)	0.22 (0.19)	0.22 (0.19)
Observations	567	567	567	567	567	567	567	277	277	277	277	277	277	277	277
							<i>First semester</i>								
Cash incentives	-0.01 (0.02) [0.74]	0.01 (0.02) [0.65]	0.08** (0.04) [0.03]	0.02 (0.02) [0.28]	0.00 (0.03) [0.90]	0.01 (0.01) [0.70]	-0.12*** (0.04) [0.00]	-0.01 (0.04) [0.79]	0.01 (0.03) [0.72]	-0.02 (0.05) [0.68]	0.00 (0.04) [0.95]	0.02 (0.03) [0.56]	0.01 (0.03) [0.83]	0.01 (0.03) [0.72]	-0.01 (0.03) [0.72]
Constant	0.12 (0.11)	-0.01 (0.01)	0.18 (0.13)	-0.01 (0.01)	-0.00 (0.01)	-0.00 (0.00)	0.73*** (0.15)	0.00 (0.02)	-0.01 (0.02)	0.51 (0.35)	-0.00 (0.02)	-0.01 (0.02)	-0.00 (0.01)	0.50 (0.37)	0.50 (0.37)
Observations	567	567	567	567	567	567	<i>Second semester</i>	315	315	315	315	315	315	315	315

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). The outcomes indicate whether or not the most frequent activity belongs to the respective category. See Appendix A.7 how activities are grouped. “Scanner” are outcome measures based on the scanner data at the gym entrance and thus only consider on-campus activities. “Self-reported” are outcome measures based on the survey data and correspond to activities overall. All regressions include randomization block fixed effects.

Table A.32: Effect of incentives on general time use (2013, with block FE)

	Study related			Work & household		Leisure			
	(1) Class	(2) Studying	(3) Commute	(4) Work	(5) Chores	(6) Sport	(7) Digital Media	(8) Friends	(9) Other leisure
<i>First semester</i>									
Cash incentives	0.27*	0.30	0.24***	0.17	-0.11	0.02	-0.07	-0.15	-0.16*
	(0.16)	(0.19)	(0.09)	(0.11)	(0.09)	(0.09)	(0.12)	(0.09)	(0.10)
	[0.09]	[0.09]	[0.00]	[0.09]	[0.27]	[0.85]	[0.59]	[0.14]	[0.09]
Constant	4.69***	1.78***	1.00**	0.13	1.44***	0.49***	2.33***	1.56***	0.46**
	(0.95)	(0.43)	(0.43)	(0.12)	(0.27)	(0.15)	(0.67)	(0.30)	(0.19)
Observations	278	278	275	273	278	277	276	276	266
<i>Second semester</i>									
Cash incentives	0.27*	0.07	0.29***	0.10	0.03	0.12	0.08	-0.01	-0.10
	(0.16)	(0.20)	(0.10)	(0.11)	(0.09)	(0.08)	(0.10)	(0.11)	(0.13)
	[0.08]	[0.71]	[0.00]	[0.38]	[0.73]	[0.20]	[0.43]	[0.91]	[0.47]
Constant	3.61**	2.21*	2.11*	4.45	3.73**	1.19**	1.71***	1.26**	3.30**
	(1.53)	(1.24)	(1.16)	(3.22)	(1.62)	(0.50)	(0.16)	(0.54)	(1.67)
Observations	315	311	308	308	315	309	311	308	296

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). Outcomes correspond to self-reported hours spent on the respective activities on a regular weekday. All regressions include randomization block fixed effects.

Table A.33: IV estimates of effect of gym visits per week on educational outcomes (2013, with block FE)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
<i>First semester</i>							
Observations			567	567	567	567	567
<i>Second semester</i>							
Visits total	0.00	0.42*	-0.13	0.41*	0.47*	0.39	0.50**
	(0.12)	(0.24)	(0.11)	(0.24)	(0.25)	(0.24)	(0.25)
Observations	567	567	567	567	567	567	567

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. The treatment variables are standardized to have mean zero and standard deviation one within cohort and randomization sample. All regressions include randomization block fixed effects.

## A.6 Course requirements first-year

Table A.34: Course requirements in first year

Type	Name	Credits
<i>First Semester (Fall)</i>		
Core Studies	Business Administration I	5.5
	Economics I	5.5
	Law I	5.5
Core Elective	Mathematics I or Law B I	3.5
Leadership Skills	Introduction to Academic Writing	3
Critical Thinking	History or Philosophy	2
Cultural Awareness	Foreign Language Level I	4
<i>Second Semester (Spring)</i>		
Core Studies	Business Administration II	5.5
	Economics II	5.5
	Law II	5.5
Core Elective	Mathematics II or Law B II	3.5
Leadership Skills	Integrative Project	4
Critical Thinking	Psychology or Sociology	2
Cultural Awareness	Foreign Language Level I	4
First-year Paper		5

## A.7 Course and sports offers in main university gym

**Team sports:** basketball, soccer, handball, lacrosse, field hockey, field hockey (university league), volleyball

**Individual sports, racket sports, martial arts:** badminton, Brazilian Jjiujitsu, capoeira, karate-do, kick boxing, tennis, table tennis

**Weights and fitness:** weights, gym

**Cardio:** cardio (indoor), indoor cycling, indoor rowing, group jogging, individual jogging

**Group fitness and dance:** ballet, bodypump, body toning, boot camp, cheerleading, contemporary dance, CrossFit, dance aerobic, dancing for women, boxing fitness, fitness cocktail,

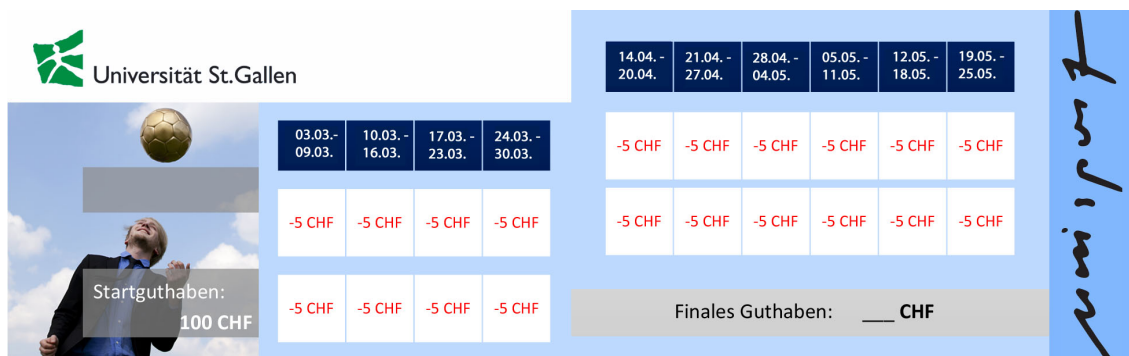


fitness-mix, functional training, high intensity conditioning, hip hop, jazz dance, conditioning, conditioning circuit, M.A.X., pilates, back fitness, step, step n tone, stretch/strength, TRX, tae bo, dance: street style, classic dance, yoga, zumba

**Others:** conconi test, frisbee, MBSR, massage, parkour/trikking/slackline, training consultation

## A.8 Example sticker card

Figure A.7: Example sticker card of in the second semester of 2014

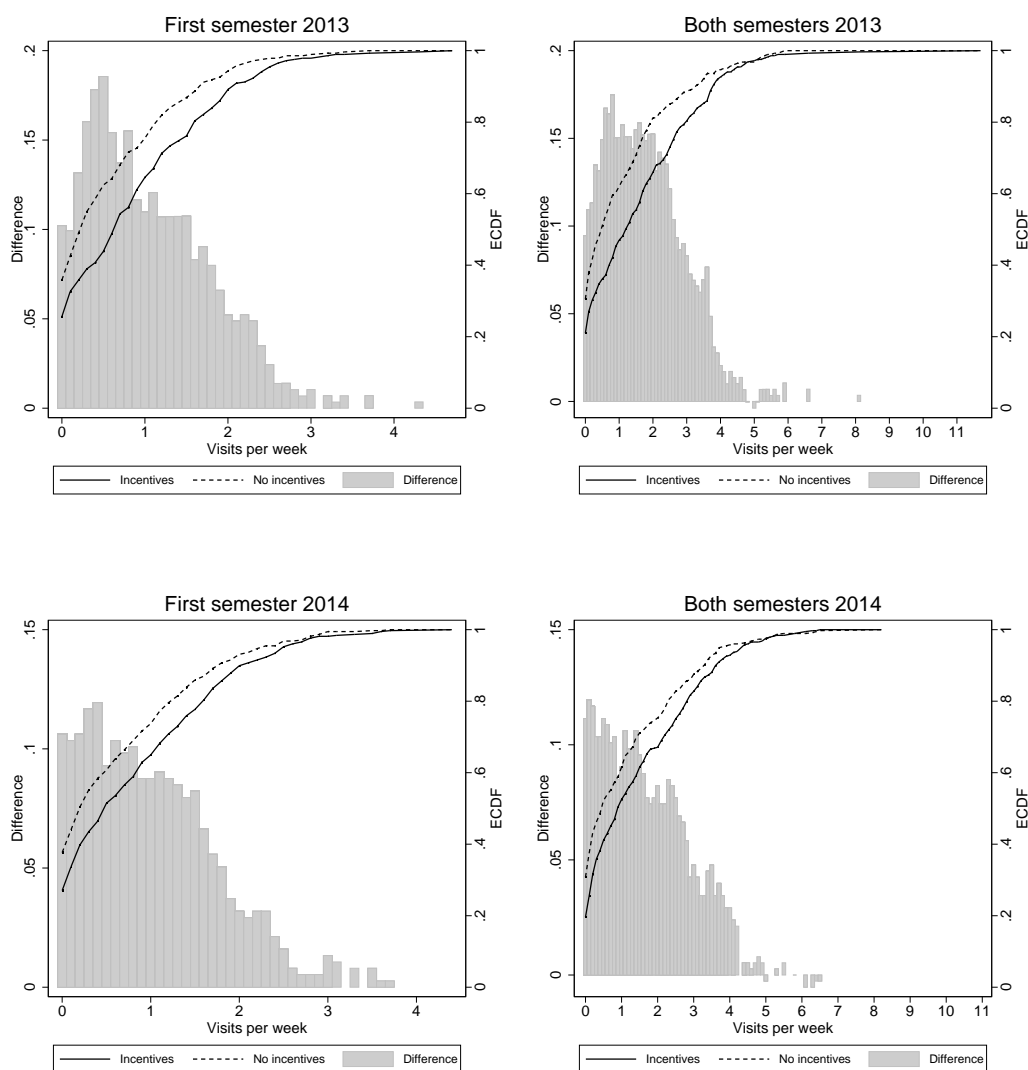


Notes: The card is folded in the middle. It contains two spots for stickers each week over ten weeks. The German words “Startguthaben” and “Finales Guthaben” translate to “Initial endowment” and “final endowment”.

## A.9 Test of monotonicity implication

In general, the monotonicity assumption is not testable, but in the presence of multivalued treatments, it has a testable implication. Angrist and Imbens (1995) show that the cumulative distribution functions of the instrument (here gym visits per week) given the instrument status (here cash incentives) should not cross. Figure A.8 plots the estimated cumulative distributions functions (ECDF) for the gym visits per week in the first semester and gym visits per week in the entire first year. For both cohorts, the ECDF of those students that received the cash incentives is below the ECDF of those students that did not receive the incentives. This finding reassures that violations of monotonicity are not a concern in our setting.

Figure A.8: Estimated cumulative distribution function of gym visits



The Effect of Physical Activity on Student  
Performance in College: An Experimental Evaluation  
- Online Appendix -

Hans Fricke, Michael Lechner, and Andreas Steinmayr

May 16, 2017

## B Online Appendix

### B.1 Results with covariates for the 2014 cohort

All of the following tables correspond one-to-one to the respective tables in the main text. However, the specifications in the appendix include the following set of controls: students' sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities. Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002).

Table B.1: Effect of incentives on gym usage (2014, with covariates)

	Semester			Time of day		
	(1) Visits total	(2) 1st sem	(3) 2nd sem	(4) Morning	(5) Afternoon	(6) Evening
Cash incentives	0.18*** (0.05) [0.00]	0.23*** (0.05) [0.00]	0.13** (0.05) [0.02]	0.05** (0.02) [0.03]	0.08*** (0.02) [0.00]	0.04* (0.02) [0.06]
Constant	0.11 (0.27)	0.05 (0.25)	0.17 (0.34)	0.08 (0.13)	0.02 (0.13)	0.02 (0.11)
F-statistic						
Observations	754	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). "Visits total" corresponds to average visits per week during the 20 treatment weeks in the first year. "1st sem" corresponds to average visits per week during the 10 treatment weeks in the first semester. "2nd sem" corresponds to average visits per week during the 10 treatment weeks in the second semester. Covariates include students' sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.

Table B.2: IV estimates of effect of gym visits per week on educational outcomes (2014, with covariates)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
<i>First semester</i>							
Visits 1st sem.			-0.13 (0.09)	0.15 (0.26)	0.28 (0.26)	0.25 (0.26)	-0.04 (0.26)
Observations			754	754	754	754	754
<i>Second semester</i>							
Visits total	0.02 (0.15)	-0.00 (0.30)	0.10 (0.15)	-0.08 (0.30)	-0.13 (0.31)	-0.05 (0.30)	-0.09 (0.30)
Observations	754	754	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. The treatment variables are standardized to have mean zero and standard deviation one within cohort and randomization sample. Covariates include students’ sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, e) other activities.

Table B.3: Effect of incentives on educational outcomes (2014, with covariates)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
	<i>First semester</i>						
Cash incentives			-0.04 (0.03) [0.18]	0.04 (0.07) [0.56]	0.08 (0.07) [0.27]	0.07 (0.07) [0.34]	-0.01 (0.07) [0.89]
Observations			754	754	754	754	754
	<i>Second semester</i>						
Cash incentives	0.00 (0.04) [0.91]	-0.00 (0.07) [1.00]	0.02 (0.03) [0.48]	-0.02 (0.07) [0.77]	-0.03 (0.07) [0.68]	-0.01 (0.07) [0.87]	-0.02 (0.07) [0.74]
Observations	754	754	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. Covariates include students’ sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.

## B.2 Results with randomization block fixed effects for the 2014 cohort

Table B.4: Effect of incentives on health and health behavior (2014, with covariates)

	Health			Health behavior		
	(1) Good health	(2) Pain	(3) BMI	(4) Healthy nutrition	(5) Smokes	(6) Drinks regularly
	<i>First semester</i>					
Cash incentives	0.02 (0.04) [0.69]	0.01 (0.03) [0.61]	0.21 (0.25) [0.38]	0.05 (0.06) [0.41]	-0.04 (0.04) [0.38]	0.03 (0.05) [0.62]
Constant	1.16*** (0.36)	0.17 (0.17)	18.50*** (1.68)	0.47 (0.42)	-0.30 (0.32)	0.12 (0.30)
Observations	298	297	297	299	300	297
	<i>Second semester</i>					
Cash incentives	-0.01 (0.06) [0.90]	0.06* (0.03) [0.07]	0.53* (0.32) [0.08]	-0.07 (0.07) [0.28]	-0.01 (0.05) [0.91]	0.01 (0.06) [0.90]
Constant	0.27 (0.33)	0.12 (0.18)	22.80*** (2.16)	0.31 (0.47)	-0.41 (0.36)	-0.06 (0.60)
Observations	208	210	208	210	208	209

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Good health” indicates whether or not a student reports to be in good or very good health. “Pain” indicates whether a student often felt physical pain during the semester. It is based on a five-point scale going from always to never. The indicator takes one if students choose 1 or 2 and zero otherwise. “BMI” is calculated as follows:  $BMI_i = \frac{weight_i}{(height_i/100)^2}$  where  $weight_i$  is student  $i$ 's self-reported weight in kilogrammes and self-reported  $height_i$  the height in centimeter. “Healthy nutrition” indicates whether or not a student cares strongly or very strongly about a healthy diet. “Smokes” indicates whether or not a student smokes at the time of the survey. “Drinks regularly” indicates whether or not a student reports to drink either beer, wine, liquor, or mixed drinks regularly. Covariates include students' sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, not speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.

Table B.5: Effect of incentives on characteristics of sport and exercise (2014, with covariates)

	Self-reported activity			Sports type		Time of day		
	(1) Overall	(2) On-campus	(3) High intensity	(4) Weights/gym	(5) Other	(6) Morning	(7) Afternoon	(8) Evening
<i>First semester</i>								
Cash incentives	0.17 (0.16) [0.31]	0.35** (0.13) [0.02]	-0.02 (0.06) [0.75]	0.13*** (0.05) [0.00]	0.10** (0.04) [0.01]	0.05** (0.02) [0.02]	0.12*** (0.03) [0.00]	0.06** (0.03) [0.03]
Constant	2.10** (1.03)	0.70 (0.82)	0.32 (0.45)	0.08 (0.24)	-0.03 (0.20)	0.09 (0.13)	0.02 (0.15)	-0.06 (0.15)
Observations	299	298	266	754	754	754	754	754
<i>Second semester</i>								
Cash incentives	0.44** (0.21) [0.03]	0.50*** (0.18) [0.01]	-0.04 (0.07) [0.60]	0.07 (0.05) [0.18]	0.06* (0.04) [0.07]	0.05* (0.03) [0.05]	0.05** (0.03) [0.05]	0.02 (0.02) [0.28]
Constant	2.27* (1.28)	0.09 (1.26)	0.73 (0.60)	0.12 (0.30)	0.05 (0.20)	0.06 (0.16)	0.01 (0.15)	0.10 (0.12)
Observations	211	210	197	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Overall days of sport” corresponds to the number of days a student does sports or exercise in a normal week during the semester. “On-campus days of sport” corresponds to the number of days a students does sports or exercise on-campus in a normal week during the semester. “High intensity” indicates whether or not a student reports to generally be exhausted with strong perspiration and breathlessness. “Weights”/gym” corresponds to weekly gym visits in this category. Students declared the category when scanning their student ID at the gym entrance. “Other” corresponds to activities other than weights/gym. “Time of day” corresponds to mean weekly gym visits in the morning (7 am to noon), afternoon (noon to 5 pm), and evening (5 pm to 10 pm). Covariates include students’ sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, and e) other activities.



Table B.6: Effect of incentives on general time use (2014, with covariates)

	Study related			Work & household		Leisure			
	(1) Class	(2) Studying	(3) Commute	(4) Work	(5) Chores	(6) Sport	(7) Digital Media	(8) Friends	(9) Other leisure
<i>First semester</i>									
Cash incentives	0.18 (0.18) [0.32]	-0.53*** (0.20) [0.01]	-0.11 (0.11) [0.30]	0.20* (0.11) [0.08]	-0.01 (0.11) [0.95]	0.09 (0.08) [0.26]	-0.02 (0.14) [0.88]	-0.01 (0.14) [0.95]	-0.01 (0.13) [0.96]
Constant	4.17*** (1.19)	-0.47 (1.40)	0.63 (0.95)	-3.09 (2.58)	-0.98 (0.92)	0.87 (0.76)	2.32** (1.06)	0.49 (1.26)	-1.28 (1.49)
Observations	300	298	300	291	300	298	297	297	284
<i>Second semester</i>									
Cash incentives	-0.03 (0.22) [0.88]	0.11 (0.22) [0.59]	-0.22** (0.11) [0.09]	0.06 (0.15) [0.74]	-0.18 (0.13) [0.22]	0.03 (0.13) [0.83]	-0.13 (0.20) [0.53]	-0.25 (0.16) [0.15]	-0.31** (0.15) [0.04]
Constant	3.57** (1.45)	-0.30 (1.43)	1.33* (0.79)	-1.42 (1.79)	-1.97 (1.44)	1.29 (1.02)	2.44 (1.68)	2.22 (1.41)	1.93 (1.29)
Observations	212	212	210	206	210	210	210	211	204

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). Outcomes correspond to self-reported hours spent on the respective activities on a regular weekday. Covariates include students' sex, age, visits to the gym prior to the start of the intervention period, as well as indicators for having contributed to the student aid fund, being in the law track, Swiss nationality, speaking German as a mother tongue, having been in high school in the canton of a) St. Gallen, b) Zurich, c) in the non-German part of Switzerland, d) outside of Switzerland, and indicators of most frequent type of activity pre-treatment being a) team sports, b) individual sports, c) gym and weights, d) group fitness, e) other activities.

Table B.7: Effect of incentives on gym usage (2014, with block FE)

	(1) Visits total	Semester		Time of day		
		(2) 1st sem	(3) 2nd sem	(4) Morning	(5) Afternoon	(6) Evening
Cash incentives	0.16*** (0.05) [0.00]	0.20*** (0.06) [0.00]	0.11** (0.06) [0.04]	0.04* (0.02) [0.06]	0.08*** (0.03) [0.00]	0.04* (0.02) [0.08]
Constant	0.49*** (0.10)	0.60*** (0.12)	0.39*** (0.10)	0.16*** (0.05)	0.19*** (0.04)	0.14*** (0.05)
F-statistic						
Observations	754	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). "Visits total" corresponds to average visits per week during the 20 treatment weeks in the first year. "1st sem" corresponds to average visits per week during the 10 treatment weeks in the first semester. "2nd sem" corresponds to average visits per week during the 10 treatment weeks in the second semester. All regressions include randomization block fixed effects.

Table B.8: IV estimates of effect of gym visits per week on educational outcomes (2014, with block FE)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
	<i>First semester</i>						
Observations			754	754	754	754	754
	<i>Second semester</i>						
Visits total	0.05 (0.16)	0.02 (0.32)	0.09 (0.17)	-0.06 (0.33)	-0.14 (0.34)	-0.02 (0.32)	-0.08 (0.33)
Observations	754	754	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. The treatment variables are standardized to have mean zero and standard deviation one within cohort and randomization sample. All regressions include randomization block fixed effects.

Table B.9: Effect of incentives on educational outcomes (2014, with block FE)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Passed first year	Overall grade	Missing grade	Avg semester grade	Business grade	Economics grade	Law grade
	<i>First semester</i>						
Cash incentives			-0.04 (0.03) [0.17]	0.04 (0.07) [0.55]	0.08 (0.07) [0.26]	0.06 (0.07) [0.40]	-0.01 (0.07) [0.92]
Observations			754	754	754	754	754
	<i>Second semester</i>						
Cash incentives	0.01 (0.04) [0.77]	0.01 (0.07) [0.94]	0.02 (0.03) [0.59]	-0.01 (0.07) [0.86]	-0.03 (0.07) [0.70]	-0.01 (0.07) [0.95]	-0.02 (0.07) [0.81]
Observations	754	754	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Passed first year” corresponds to a binary indicator for whether or not the student passed the first year curriculum in the first attempt. “Overall grade” corresponds to the average grade over the first year course work weighted by the number of credits for each course. “Avg semester grade” correspond to the average grades over the first and second semester course work weighted by the number of credits of each course. “Business grade”, “Economics grade”, and “Law grade” correspond to the individual grades in the core courses in Business I, Economics I, and Law I in the first semester and Business II, Economics II, and Law II in the second semester. Missing indicates whether or not any of the first- or second-semester grades are missing. All grades are standardized to have mean zero and standard deviation one. For all of the grade outcomes missings are set to the lowest grade before standardization. All regressions include randomization block fixed effects.

Table B.10: Effect of incentives on health and health behavior (2014, with block FE)

	Health			Health behavior		
	(1) Good health	(2) Pain	(3) BMI	(4) Healthy nutrition	(5) Smokes	(6) Drinks regularly
<i>First semester</i>						
Cash incentives	-0.00 (0.04) [0.91]	0.03 (0.03) [0.38]	0.24 (0.25) [0.37]	0.02 (0.06) [0.76]	-0.03 (0.04) [0.47]	0.02 (0.05) [0.63]
Constant	0.77*** (0.11)	-0.02 (0.02)	22.59*** (0.41)	0.58*** (0.13)	0.14 (0.09)	0.16 (0.10)
Observations	298	297	297	299	300	297
<i>Second semester</i>						
Cash incentives	-0.04 (0.05) [0.46]	0.06* (0.03) [0.07]	0.51 (0.31) [0.11]	-0.04 (0.07) [0.62]	0.01 (0.05) [0.86]	0.03 (0.06) [0.68]
Constant	0.86*** (0.12)	0.04 (0.08)	22.58*** (0.52)	0.53*** (0.16)	0.08 (0.11)	0.23* (0.14)
Observations	208	210	208	210	208	209

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). “Good health” indicates whether or not a student reports to be in good or very good health. “Pain” indicates whether a student often felt physical pain during the semester. It is based on a five-point scale going from always to never. The indicator takes one if students choose 1 or 2 and zero otherwise. “BMI” is calculated as follows:  $BMI_i = \frac{weight_i}{(height_i/100)^2}$  where  $weight_i$  is student  $i$ 's self-reported weight in kilograms and self-reported  $height_i$  the height in centimeter. “Healthy nutrition” indicates whether or not a student cares strongly or very strongly about a healthy diet. “Smokes” indicates whether or not a student smokes at the time of the survey. “Drinks regularly” indicates whether or not a student reports to drink either beer, wine, liquor, or mixed drinks regularly. All regressions include randomization block fixed effects.

Table B.11: Effect of incentives on characteristics of sport and exercise (2014, with block FE)

	Self-reported activity			Sports type		Time of day		
	(1) Overall	(2) On-campus	(3) High intensity	(4) Weights/gym	(5) Other	(6) Morning	(7) Afternoon	(8) Evening
<i>First semester</i>								
Cash incentives	0.02 (0.18) [0.90]	0.22 (0.16) [0.16]	-0.03 (0.06) [0.59]	0.08 (0.05) [0.11]	0.12*** (0.04) [0.01]	0.04 (0.03) [0.14]	0.11*** (0.03) [0.00]	0.06* (0.03) [0.05]
Constant	2.11*** (0.28)	1.74*** (0.42)	0.69*** (0.13)	0.40*** (0.11)	0.20** (0.08)	0.17*** (0.05)	0.25*** (0.06)	0.18*** (0.07)
Observations	299	298	266	754	754	754	754	754
<i>Second semester</i>								
Cash incentives	0.55*** (0.21) [0.01]	0.53** (0.21) [0.02]	-0.02 (0.07) [0.83]	0.05 (0.05) [0.37]	0.07* (0.04) [0.06]	0.04* (0.03) [0.09]	0.05* (0.03) [0.08]	0.02 (0.02) [0.30]
Constant	2.09*** (0.35)	1.10*** (0.35)	0.51*** (0.16)	0.27*** (0.10)	0.11** (0.05)	0.14*** (0.05)	0.14*** (0.05)	0.11** (0.04)
Observations	211	210	197	754	754	754	754	754

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters. “Overall days of sport” corresponds to the number of days a student does sports or exercise in a normal week during the semester. “On-campus days of sport” corresponds to the number of days a students does sports or exercise on-campus in a normal week during the semester. “High intensity” indicates whether or not a student reports to generally be exhausted with strong perspiration and breathlessness. “Weights”/gym” corresponds to weekly gym visits in this category. Students declared the category when scanning their student ID at the gym entrance. “Other” corresponds to activities other than weights/gym. “Time of day” corresponds to mean weekly gym visits in the morning (7 am to noon), afternoon (noon to 5 pm), and evening (5 pm to 10 pm). All regressions include randomization block fixed effects.

Table B.12: Effect of incentives on general time use (2014, with block FE)

	Study related			Work & household		Leisure			
	(1) Class	(2) Studying	(3) Commute	(4) Work	(5) Chores	(6) Sport	(7) Digital Media	(8) Friends	(9) Other leisure
<i>First semester</i>									
Cash incentives	0.13 (0.18) [0.50]	-0.54*** (0.20) [0.01]	-0.10 (0.11) [0.38]	0.23* (0.12) [0.05]	0.01 (0.11) [0.93]	0.05 (0.08) [0.56]	-0.05 (0.13) [0.76]	-0.02 (0.13) [0.90]	0.03 (0.13) [0.82]
Constant	4.12*** (0.30)	2.90*** (0.51)	1.09*** (0.23)	0.18 (0.18)	1.14*** (0.24)	1.08*** (0.22)	1.26*** (0.21)	1.37*** (0.35)	1.14*** (0.40)
Observations	300	298	300	291	300	298	297	297	284
<i>Second semester</i>									
Cash incentives	-0.00 (0.21) [0.99]	0.11 (0.22) [0.66]	-0.26** (0.13) [0.04]	-0.07 (0.17) [0.70]	-0.26* (0.15) [0.09]	0.04 (0.13) [0.80]	-0.09 (0.19) [0.67]	-0.23 (0.17) [0.18]	-0.31** (0.15) [0.04]
Constant	3.46*** (0.44)	2.67*** (0.52)	1.07*** (0.20)	0.14 (0.15)	1.24*** (0.25)	0.97*** (0.17)	1.94*** (0.39)	1.51*** (0.30)	1.03*** (0.29)
Observations	212	212	210	206	210	210	210	211	204

Notes: Values in parentheses present robust standard errors. \*, \*\*, and \*\*\* indicate statistical significance of the coefficient on the 10, 5, and 1% level based on these standard errors. Values in squared brackets present p-values from permutation tests with 1,000 replications. Permutations were conducted within randomization clusters following the procedure outlined in Rosenbaum (2002). Outcomes correspond to self-reported hours spent on the respective activities on a regular weekday. All regressions include randomization block fixed effects.