The "Opportunity Knocks" Supplemental Merit Scholarships Project:

A Randomized Evaluation of Merit Aid for Students Receiving Need-Based Aid

Prepared by Joshua Angrist, Tony Chambers, Philip Oreopoulos and Tyler Williams for the

Higher Education Quality Council of Ontario

Disclaimer:

The opinions expressed in this research document are those of the authors and do not necessarily represent the views or official polices of the Higher Education Quality Council of Ontario or other agencies or organizations that may have provided support, financial or otherwise, for this project

Cite this publication in the following format:

Angrist, J., Chambers, T., Oreopoulous, P. and Williams, T. *The "Opportunity Knocks" Supplemental Merit Scholarships Project: A Randomized Evaluation of Merit Aid for Students Receiving Need-Based Aid.* . Toronto: Higher Education Quality Council of Ontario.

Published by:

The Higher Education Quality Council of Ontario

1 Yonge Street, Suite 2402

Toronto, ON Canada

M5E 1E5

Phone: (416) 212-3893

Fax: (416) 212-3899

Web: www.heqco.ca

E-mail: info@heqco.ca

Table of Contents

Acknowledgements

The Opportunity Knocks (OK) Project is the result of collaboration among a large number of individuals. We are indebted to the many university administrators who helped make this project happen – especially to our project manager at the University of Toronto, Meera Rai, for her incredible enthusiasm and dedication. We very much appreciate the encouragement and interest from Ken Norrie, Fiona Deller, Richard Wiggers and others at the Higher Education Quality Council of Ontario (HEQCO) and are also grateful for their support for randomized demonstration projects. Special thanks also go to OK's student advisors and to Florian Hoffmann and Matt Boire for research assistance.

Executive Summary: Introduction to Opportunity Knocks

Helping individuals obtain a college or university education, regardless of background, remains a key priority for provincial and federal governments in Canada. More and more postsecondary education (PSE) students, however, struggle academically. While PSE enrolment has increased, completion rates have fallen. Within Canadian universities today, about 70 per cent of entering students eventually graduate, and some schools face completion rates of as low as 50 per cent. Average grades have also fallen or been inflated. Administrators' efforts to reverse these trends by offering additional support services such as advising, time management workshops and remedial education have been generally unsuccessful.

Another explanation for worsening academic performance is declining study time. Recent evidence shows a substantial fall in average study times among postsecondary students over the last four decades. Greater financial constraints on today's students and an increased need to work part time may prevent them from spending more time on school. On the other hand, poor-performing students may simply see less need to achieve better grade performance because they perceive obtaining an undergraduate degree as the primary benefit from postsecondary education. Or it is possible that more PSE students are myopic. Students invest time and effort for uncertain returns that are not realized until many years later. This uncertainty may lead students to focus more on immediate gratification and present opportunities and spend less time on school work. Many stakeholders are interested in how to motivate students to overcome these difficulties and perform better in school.

The goal of the Opportunity Knocks (OK) Project was to effectively learn more about the potential for merit-based scholarships to provide both additional financial support and more motivation for improved academic performance. OK was a randomized field experiment that involved first year and second year students receiving financial aid in 2008/09 at the University of Toronto Scarborough (UTSC). Students on financial aid were chosen because monetary incentives should be more meaningful to them. The campus includes a diverse student body, most of whom commute from home. About ten thousand full-time students attend each year. All first year and second year students on financial aid were invited to participate in OK. Those selected by lottery into the treatment group were offered merit scholarships for obtaining course grades above 70 per cent, as well as regular peer advising services.

More specifically, for each one-semester course (with a full course load being 5 courses worth

2.5 credits), students received $100 for obtaining a grade average of 70 per cent and $20 for each percentage point above 70 per cent. For example, a student receiving a grade of 76 per cent would have received $220. If a student received 76 per cent in all of her or his 10 courses over the school year (5 each semester), she or he would have received a total of $2,200 (equal to 10 × $220).

OK participants selected for treatment were also assigned a peer advisor of the same gender and were offered opportunities to engage in e-mail correspondence with that advisor to discuss academic matters, as well as issues arising from campus life. Peer advisors were enthusiastic, paid upper year students or recent graduates with successful academic achievement. Each peer advisor was assigned to 50 students who had been selected for the OK treatment program. Advisors were the key front-line service and information providers for OK participants. They proactively sent e-mails to advisees approximately once every two to three weeks, whether or not a response was acknowledged. These e-mails offered advice on upcoming academic events and workshops and on how to approach particular periods in the academic calendar such as midterms and finals. Advisors also provided information about the Opportunity Knocks scholarships, including payment schedules and reminders of how scholarships were calculated.

Goals, Attitudes and Aspirations of Participants

A majority of all participants (57.2 per cent) stated that they were very concerned about having sufficient funds to complete their university degree, reinforcing the fact that these students face serious financial constraints regarding educational opportunities. For most students (76.7 per cent), UTSC was their first-choice school, and many of our participants intended to complete a graduate degree (43.3 per cent). Thus, it seems that our study population has ambitious academic goals but perhaps lack the funds or academic performance to achieve these goals.

As part of the participation sign-up process, we asked two questions concerning understanding of the merit award formula, which gave students $100 for each course with a final grade of 70 per cent or higher and an additional $20 for each percentage point attained above 70 per cent in each course. In the first question, students were asked to calculate the award amount from a grade in a single class, and in the second, they were asked to calculate the total award amount from grades in five different classes. Virtually all participants (84 per cent) answered the first, simpler scholarship calculation question correctly, although only two-thirds answered the second, more difficult question correctly. We sent follow up information about the scholarship formula to participants who responded incorrectly to at

least one of the two questions. In the following program analysis, we look at academic effects for the entire sample and for the subgroup that answered the first question correctly to see if those who understood the merit scholarship better also had a stronger program response.

To determine whether students thought the OK Project's treatment program would help them increase their grades, we asked students what grade they expected to receive if they were selected to be part of the treatment group and what grade they expected to receive if they were not selected. Interestingly, fewer than half (37.3 per cent) of the participants thought they would have a higher grade if they were selected for the program, but the average difference between the expected GPA in the treatment group and the expected GPA in the control group was significant (expectations in the treatment group were 0.195 GPA points higher, on average, than those in the control group – about 3 percentage points higher on a percentage scale), suggesting that participants, on average, expected the program to be somewhat helpful. Figure 1 shows smoothed kernel densities of the grade distributions for expected grade averages (as percentages) for first year and second year men and women. Of those who did report differences in expected grades as a result of being selected into the treatment program, most said their grades would increase by 5 to 10 percentage points. The expected effect was more concentrated among students who believed their grades would fall on the lower end of the distribution.

Impacts on Advisor Use

Many participants conveyed their enthusiasm for the advisor component of the program. Summary Table 1 shows that a majority of participants sent e-mails to their advisors throughout the year, and the percentage of students who made use of the advisors highest. Women were also more likely to contact their advisors over the course of the whole year. Even among students who did not communicate with an advisor, most conveyed appreciation for the information provided in the regular e-mails they received. The table also shows that program officers received communication from virtually all participants, providing support that all students were aware of, and responsive to, the program.

Impacts on Achievement

Summary Table 2 summarizes the main impacts Opportunity Knocks had on academic performance for the full sample. Control group students attempted about two credits (roughly four half courses) each semester on average, and the Treatment Effect rows show that the OK Project treatment program had a positive effect on credits attempted in both semesters for most subgroups. However, these effects are

statistically different from zero (denoted by asterisks) in only one case (and then only at the 10 per cent level). In contrast to previous literature, which has found that academic programs have larger effects on women than on men, we find the strongest effect on credits attempted by first year men in the winter semester-specific results shown in Table 6. We also find smaller, marginally significant effects in the winter semester for all men, all first years and the whole sample – and in the fall semester for second year men (significant at the 10 per cent level; see Table 6). It seems that men, particularly first year men, did respond to the OK program by attempting more credits, although the responses were small. The fact that we find stronger effects on enrolment during the winter semester might suggest that students understood their earnings potential in the program better after receiving payment for their fall courses or that the peer academic counselling gradually encouraged them to be more ambitious. According to this interpretation, we would expect the positive effects on winter credits attempted to persist if the OK program were implemented for longer than one year.

A natural benchmark for program effects can be created by comparing the amount of scholarships actually earned in the experimental group with the earnings that students in the control group would have been entitled to had they been in the program. A large program effect should be reflected in larger-than-expected earnings, where expected earnings are measured using hypothetical earnings based on the grade distribution in the control sample. The results suggest that the OK program had no effect on would-be scholarship earnings for first year men and women. However, the effects on second year men and women are uniformly positive and sometimes statistically significant. In particular, the effect of the program on winter earnings for second year men is about $171 (significant at the 5 per cent level), and the effect on full year earnings for all second year students is about $185 (significant at the 5 per cent level). Overall, it seems that the OK program had a small positive effect on scholarship earnings for second years, particularly second year men, but no effect on earnings for first years.

The effects on grades are weaker than the effects on earnings. This is because the grade effects appear concentrated in the 70 per cent range but not elsewhere. Average grades for second year men responded positively to the program, but the 2.551 percentage point increase in these average grades is significant only at the 10 per cent level. Most other effects for second years are slightly positive or very close to zero.

Summary Table 2 also shows that students in the treatment program achieved a 70 per cent grade in a higher number of courses than did students in the control group. Second year men in the

treatment group achieved, on average, a 70 per cent grade or higher in 0.598 more winter classes than did the control group (statistically significant at the 5 per cent level; see semester-specific results presented in Table 9). This translated into an increase of almost one whole class above 69 per cent, on average, over the whole year (statistically significant at the 5 per cent level). The OK program had similar, smaller effects on the number of courses in which all second years obtained 69 per cent.

None of the estimated effects are particularly large, though the earnings decomposition analysis provides reasonably strong evidence that second year men responded to the financial incentives in the OK program. The larger incentive ($100) at the 70 per cent level seems to have been more effective than the smaller, incremental incentive ($20) for grades above 70 per cent. Bigger effects in the winter semester are consistent with the results for credits attempted, which indicates that students may have understood the program better after one semester of experience and/or that the academic counselling portion of the OK program took some time to bear fruit. In both cases, a more permanent implementation of the OK program might have larger, sustained, positive effects.

The OK program had very small effects on achievement in the year after the scholarship offer ended. This is understandable given that the program had small effects during the study period itself. The program did decrease dropout rates somewhat for men during the study period, with the largest effect (a 4.7 percentage point decrease) occurring for second year men in fall 2008.

Larger Effects for Students That Understood Program Better

Summary Table 3 presents similar estimated effects – in this case on participants who responded correctly to a pre-treatment question regarding how the merit scholarships would be calculated. Within the restricted sample, the scholarship earnings effect of the OK program on all second years is positive for both semesters and for the full year. Full-year earnings are, on average, about $259 higher (significant at the 1 per cent level) for second years in the treatment group, which is about one-quarter of a standard deviation. Second year men had a positive, marginally significant grade response to the OK program in the winter, and all second years had a positive, marginally significant response to the program in the winter and over the full year (semester-specific results are shown in Table 12). These effects are about 1.3 to 2.7 percentage points, or about one-fifth to one-quarter of a standard deviation. From these results, it seems that second years who understood the scholarship formula did, in fact, respond to the OK program more strongly, though first years did not, and the second year effects are still small in magnitude.

The OK program increased the number of courses above 69 per cent noticeably for second year women, second year men and all second years in both semesters. The increase for second year women and men over the full year were 0.773 courses and 0.976 courses, respectively (both significant at the 5 per cent level). This analysis suggests that the OK program made the biggest positive impact on grades near the 70 per cent threshold. These positive effects appear to be a response to the strong incentive ($100) to get a grade of 70 per cent or more in each class. In addition, the effects are stronger for the population that demonstrated better understanding of the award formula by answering the first test question on scholarship earnings correctly. The smaller, $20 incentive for each grade point earned above 70 per cent seems to have had some positive impact on grades in this restricted sample, but the effect is less pronounced.

Policy Implications

The OK program was popular with participants: sign-up rates and program engagement were high, and many program participants were enthusiastic about the experience. At the same time, consistent with an emerging body of evidence evaluating interventions of this type, overall program effects on achievement were modest. Treated students earned more in OK scholarship money than we would have expected based on the control grades distribution, but this appears to have been largely due to an effect on the number of courses in which students earned a grade of at least 69 per cent, especially among second year students. This localized response did not translate into a substantial boost in overall achievement, but it was noticeably stronger in a large subsample of students who appear to have understood the award scheme well.

The past decade has seen a growing number of pilot studies of pay-for-performance schemes for students at various levels. The picture that emerges from this work is of mostly modest overall effects. The overall achievement impact of the OK program was small, though stronger effects emerge for subgroups. It may be that different incentive schemes, such as larger amounts or lower thresholds, would lead to larger effects. Another study, for example, found more significant impacts on low-income and low-performing students when they were offered scholarships for course completion and service use. Overall, however, our interpretation of the research is that merit-based aid is an expensive approach, which generates only modest effects on retention and performance. One explanation for this is that under-performing students may not know how to study effectively: more time spent, using ineffective study methods, will not help improve grades. Another explanation is simply that lack of

academic preparation better explains under-performance than lack of effort. Unfortunately, availability of advisory and other support services has not helped extensively. Other potential avenues for improving retention and performance, or alternative approaches to teaching, are therefore needed at the high school and postsecondary levels.

Chapter One The Opportunity Knocks Program

1.1 Introduction

Helping individuals obtain a college or university education, regardless of background, remains a key priority for provincial and federal governments in Canada. Providing need-based financial aid, tax deferral incentives, regulated tuition and employment opportunities are major programs designed to increase enrolment in postsecondary institutions, especially among students with disadvantaged backgrounds. Enrolment, however, is not the end objective of these policies. Higher education is promoted because of its potential to generate significant individual, social and economic gains from the experience.[1] What happens to students after enrolment is, therefore, crucial in determining the extent to which students actually benefit from the experience.

Recent evidence shows higher ratios of PSE students struggling academically than in the past. While PSE enrolment has increased, completion rates have fallen. Among Canadian universities today, about 70 per cent of entering students graduate, and some schools face completion rates as low as 50 per cent (Grayson & Grayson, 2003). Average grades have also fallen or been inflated in several fields (Anglin & Meng, 2000). With the view that achievement problems reflect a weak academic background, administrators have responded by offering additional support services such as advising, time management workshops and remedial education, but without much success in reversing these trends.

Reduced study time or effort may also explain falling academic performance. Babcock and Marks (2010) provide some evidence of this by documenting substantial declines in study time by postsecondary students over the last four decades. One reason why today's students might be studying less is that they are poorer and/or face heavier financial obligations. That is, today's students may have

[1] Bob Rae, for example, emphasizes in Ontario: A Leader in Learning education policy reforms: "Strong and vibrant higher education is essential for the creation of an educated and skilled workforce, which in turn is needed to build and maintain a competitive and prosperous society. From a social perspective, accessible and affordable higher education provides opportunities for all capable and interested members of society to seek knowledge, broaden their capabilities, and develop their own potential. From an economic perspective, higher education institutions provide a fertile environment for the creation of the human and knowledge capital that is essential for sustained economic growth" (Rae, 2005 http://www.edu.gov.on.ca/eng/document/reports/postsec.pdf).

to work more hours in part-time jobs to overcome greater financial constraints, leaving less time for study and preventing them from realizing long-term and lifetime gains from improving grade performance. Reducing or removing these financial constraints by providing additional financial assistance could produce greater social and monetary gains over these students' lifetimes than the cost of the initial assistance, since they will likely earn more money in the future if they get better grades and learn more in school. On the other hand, under-performing students may simply see no need to achieve better grade performance because they perceive the acquisition of an undergraduate degree to be the primary benefit of postsecondary schooling. A student's advantage in obtaining an undergraduate degree may have more to do with avoiding categorization by employers as someone with only a high school degree than with the learning that occurs during PSE. Many employers, for example, require a job applicant to have an undergraduate degree just to apply for a position, but they may ignore or focus less on additional transcript information or other characteristics, such as leadership abilities or work or research experience. Students not intending to continue school after graduation from their first postsecondary credential may choose to study just enough to obtain their bachelor's degree but not enough to obtain stellar grades, even if they are capable of doing that well.

Other theories attempting to explain lower academic performance imply that students may have myopia. Education involves immediate costs for future gains, with students investing time and effort for returns that are uncertain and may not be realized until many years later. This uncertainty may lead students to focus more on immediate gratification and present opportunities and to spend less time on school work. A final theory suggests that some students may not have the background or experience to actually know the value of academic performance regardless of their capacity to do well. If they knew the present and future value of such high performance, perhaps they would apply themselves differently.

1.2 The Student Achievement and Retention Project Results and Protocol

The project presented here builds on our previous study, the Student Achievement and Retention (STAR) project (Angrist, Lang, & Oreopoulos, 2009). Project STAR began as an effort to explore the potential of student financial incentives that would foster immediate and continued grade improvement. The program was based on the possibility that offering more immediate incentives (money) to poorly performing students for obtaining grades of approximately 70 per cent or more could offset corresponding immediate costs (such as study time and effort to obtain these grades) and help students

discover their academic potential. Project STAR also examined the impact of additional student services – student advising and supplemental course instruction – as separate and complementary programs offered to students along with financial incentives for good grades. At the beginning of the school year (fall 2005), all first year students entering the University of Toronto Mississauga (whether on financial aid or not) – except those with a high school grade point average (GPA) in the upper quartile – were selected for one of three program groups or a comparison group. Students selected into the Student Fellowship Program (SFP) group were offered merit scholarships of $1,000, $2,500, and $5,000 for obtaining year-end grade average thresholds in the 70 to 85 per cent range (the exact thresholds depended on each student's high school GPA quartile). Students selected for the Student Support Program (SSP) group were offered additional student services not offered before, including peer advising by upper year students and Supplemental Instruction that provided critical thinking strategies for performing well in a particular course. Finally, the Student Fellowship and Support Program (SFSP) group was offered both programs. Project STAR produced a number of findings, which offered suggestions for ways to modify the project's design and population to deliver larger gains. Among these findings were the following:

1. Offering merit scholarships improved first semester performance, but this gain dissipated for some students by the end of the second semester.
2. The merit scholarships increased use of outside-class academic support services.
3. The largest gains occurred for students who were also offered peer advising.
4. Use of peer advising increased when students were matched to a peer advisor of the same gender, but peer advising alone had no effect on academic outcomes.
5. The effects were much larger for women than for men.
6. Specifically, women who were offered both merit scholarships and peer advising showed significant increases in GPA, credits earned and chances of ending up in academic good standing, both in the first year and the second year, even though the program was offered only in the first year.

The positive, long-lasting impacts on females suggest a potential for merit scholarships to provide additional financial aid while helping foster motivation and long-term achievement for some students. The results help validate the introduction of merit scholarships as an additional type of financial aid. Understanding which alternative designs might produce larger and/or more cost-effective gains is key. The finding that the gains were not uniform across groups suggests that context matters in

determining whether merit scholarships can improve achievement and retention.

Our new experiment built on lessons from Project STAR to produce a program aimed at generating cost-effective improvements to student achievement while providing students with an additional type of financial aid. Experience from Project STAR suggested that some students may have thought the grade targets were out of reach. Students at risk of dropping out improved their GPAs as a result of the program, but some may not have improved enough to qualify for the scholarship. We concluded that lower award targets could attract more sustained interest from at-risk students because the targets would be seen as more obtainable. A simpler target scheme could also be easier for students to remember and understand. The greater impacts from the merit scholarship program that included peer advising also suggested that reaching out regularly to scholarship-eligible students by offering assistance and sound academic advice may increase chances for academic success. These experiences were incorporated into the Opportunity Knocks program design.

1.3 The Opportunity Knocks Supplemental Merit Scholarship Project

The Opportunity Knocks (OK) Project was designed to evaluate the effectiveness of supplemental merit scholarships offered to first and second year university students receiving financial aid. The study participants were first and second year students at the University of Toronto at Scarborough (UTSC). The student population at UTSC is generally representative of undergraduates in Ontario and Canada except for the fact that the school offers cooperative education opportunities beginning roughly in the third or fourth year of a program. About half of first year students at UTSC end their year with grades lower than 70 per cent, and 10 per cent of all students are placed on academic probation. Six-year completion rates (graduated within six years since beginning program) are about 75 per cent.[2]

The OK Project targeted all first and second year students at UTSC who were receiving financial aid, as well as those who had received financial aid in the previous year. Of the eligible students who signed up (65.3 per cent), 400 were randomly selected for the treatment group, which offered merit scholarships to students who obtained a 70 per cent grade average or higher in courses taken during the 2008/09 school year. In addition, treatment group students were assigned upper year, same-gender peer advisors who were available to meet in person or by e-mail. Peer advisors initiated frequent contacts in order to offer help and guidance. Those not assigned to the treatment group formed the comparison

[2] UTSC is a satellite to the University of Toronto main campus, with somewhat lower admission standards and lower completion rates than the main university campus. The students and curriculum at UTSC are much like those at University of Toronto Mississauga, where we implemented Project STAR.

group for our study; they received neither the merit scholarship nor the extra peer advising service. In response to the evidence discussed previously in this report, we based the overall structure of this intervention on the combined "incentives and services" model that was most successful among those piloted in Project STAR – but with a number of important modifications that are described in more detail in Chapter 2. The outcomes of primary interest are academic achievement, academic standing, credits earned and retention.

The goals of this study were to

1. assess students' changing education aspirations, motivations, financial concerns and objectives for PSE;
2. demonstrate whether merit scholarships may effectively improve initial and long-term PSE achievement, increase retention and encourage intellectual development;
3. evaluate the overall cost effectiveness of a merit scholarship program in providing financial aid and improving PSE achievement;
4. reinforce the positive results from Project STAR and confirm the hypothesis that peer advising contributes to the effectiveness of offering a merit scholarship program; and
5. survey students' own impressions concerning the effectiveness of a merit scholarship program before and after participation.

As noted, many students who enter university leave without a degree. Poor academic achievement is a strong predictor of student completion, and it is also related to detrimental career outcomes. The main purpose of the OK Project was to provide additional evidence about whether enriched supplemental academic support, coupled with merit-based scholarships, can improve early performance and motivation among first and second year undergraduate students, and in doing so, potentially lower attrition and improve other subsequent education, social and labour market outcomes.

1.4 Related Research on Merit Scholarships at the Postsecondary Level

Policies that offer financial incentives for academic achievement are not new. Such merit scholarships have a long history in the postsecondary context, but traditional programs, such as the Canadian Excellence Awards, have focused on a small number of very high achievers.[3] However, a recent development in the scholarship field is an attempt to use financial awards and incentives to motivate

[3] An American awards program, The National Merit program, awards roughly 8,200 scholarships to students selected from 1.4 million PSAT takers.

good but not spectacular students. Examples include state tuition waivers for students who maintain a B average such as Georgia's HOPE Program. As Dynarski (2005) notes, programs of this sort are relevant for many students. For example, nearly 60 per cent of high school graduates in Georgia qualify for a HOPE scholarship (if they go to college). In addition to providing more financial resources for college, a second goal of the HOPE program is to promote academic achievement (Seligman, Milford, O'Looney, & Ledbetter, 2004). The belief is that a promise of a scholarship may increase the time devoted to school work and lead students to acquire better study habits.

The evidence for the impact of postsecondary merit scholarships is more limited than that for the impact of academic support services, though interest in this topic is growing as scholarship programs have expanded. In Canada, the only two studies on merit scholarships at the postsecondary level are Project STAR (previously discussed) and the Foundations for Success (FFS) experiment (MacDonald, Bernstein, & Price, 2009).

Foundations for Success offered at-risk students at three community colleges in Ontario $750 per semester (over two semesters) for obtaining a GPA of 2.0 or higher and completing 12 hours of student service activities. Treated program students were also assigned case managers to recommend which particular services to use. Grade point averages rose about 0.1 points for the incentive group compared to the control group, and retention between year 1 and year 2 rose from 62.6 per cent to 67.2 per cent, with larger effects among students with lower family incomes, ESL students and students with weaker high school grades. Offering additional student services on their own, without the financial incentives, had no effects, as had been the case with the Project STAR results.

In the United States, MDRC (a not-for-profit policy evaluation corporation) implemented the Opening Doors experiment in Louisiana in the 2005 school year. The subject sample was parents entering community colleges. Treated participants were offered $1,000 for each of two semesters if they enrolled in college at least half time and maintained a grade of C or better (a GPA of 2.0 or more). Outcome results before Hurricane Katrina disrupted operations indicated that more treated students followed through with registering for college and that treated students were significantly more likely to register for the second semester (Richburg-Hayes et al., 2009). MDRC is following up on these results with the Performance-Based Scholarship (PBS) Demonstration to evaluate whether similar findings arise in different environments and with different variations of this setup.

A number of quasi-experimental evaluations suggest that programs like Georgia HOPE do boost college attendance and completion (Cornwell, Mustard, & Sridhar, 2006; Dynarski, 2002, 2005).

An evaluation of gradual increases in tuition payments in response to delayed completion undertaken by Garibaldi, Giavazzi, Ichino, & Rettore (2007) found substantial improvements in on-time completion for Italian college women, while a study on the impacts from the Gates Millennium Scholar Program found effects on retention for minority students (DesJardins & McCall, 2008).

It is also worth mentioning other recent evidence concerning incentives for academic performance among primary and secondary students. The amounts of the financial incentives for these students are typically much smaller than those used in PSE programs. Bettinger (2008), for example, evaluated a program that provided up to $100 to students in third, fourth, fifth and sixth grades for successful completion of their standardized testing and found modest effects on mathematics scores. Fryer (2010) found little effect from paying money to fourth and seventh graders at economically disadvantaged schools to perform well on tests, but he did find significant test score effects after paying students to read. To explain these results, the author hypothesized that children may not know how to improve their test scores, so it is better to incentivize activities that lead to learning, such as reading. Ashworth et al. (2001) explored the impact of a pilot project of the Education Maintenance Allowance (EMA), which provided stipends to students from low-income backgrounds for continuing high school beyond the minimum length required. About 9,000 students offered the EMA in five pilot areas were matched to 5,500 students with similar initial characteristics in five other selected control areas. The EMA raised full-time education participation among students older than 16 years of age by 5.9 percentage points, with no positive or negative spillover effects on ineligibles in the pilot areas.

Kremer, Miguel and Thornton (2009) report results from a randomized evaluation of a merit scholarship program for adolescent girls in Kenya. Girls who scored well on academic exams had their school fees paid the following year and received a cash grant for school supplies. Eligible girls showed significant gains in exam scores, and this persisted following the competition. Ineligible boys in schools with eligible girls also showed sizable average test gains (suggesting some kind of spillover), as did girls with low pre-test scores, who were unlikely to win. Angrist and Lavy (2002) evaluated a demonstration program that provided substantial cash incentives to high school students in Israel based on previous test scores and low socioeconomic status, as well as another program that randomized across schools. The mean pass rate of the standardized matriculation exam 32 per cent among the 10 treated schools compared to 24.5 per cent for the 10 non-treated schools. No differences in pass rates were found for the experiment done at the individual level.

In short, it seems that monetary incentives at the postsecondary level for grades and other

outcomes have had positive effects on grades and/or completion in most scenarios, though often the effects are not very large or are limited to subgroups of the study populations. Programs that have combined academic support services and merit scholarships have consistently had positive effects on grades, especially in Project STAR, and financial incentives tested at the high school and earlier levels show mixed results, though many have been quite effective. However, most programs are problematic because they have provided big incentives at grade thresholds. Some students may start so far below these thresholds that they think they are unreachable, while others may easily move above the threshold and then stop working hard. The OK Project hoped to find larger academic effects by providing incentives to improve in each class at every grade level above 69 per cent. (We describe the design in detail below.) It also focused on students with demonstrated financial need, since they should care more about financial incentives.

1.5 Overview of This Report

This report is organized into eight remaining chapters, with supporting appendices. Chapter 2 provides essential background information about the implementation details of the OK program. It describes the design of the merit scholarship and academic support components of the program as well as their motivating theory. Chapter 3 describes the contextual environment of the campus where the project was conducted, including a discussion of the characteristics and circumstances of the study's sample. Sign-up and participation rates are described in Chapter 4. This chapter indicates the extent to which students accessed student services and whether interest in the programs was related to background characteristics. The quantitative analyses of the OK Project's impacts are found in Chapters 5 through 7. Chapters 2 through 7 use quantitative data collected for this report to test whether the OK Project had measurable effects on various academic outcomes. Chapter 8 provides a more qualitative analysis by reporting and discussing conversations held with students after they had participated in the OK Project for one year. Chapter 9 synthesizes the results of the analyses across all the domains. It discusses the results as a whole and addresses their policy implications. It also mentions important avenues of further research. The report's appendices provide supplementary tables and materials used during student recruitment and program implementation.

Chapter Two Implementation Details

2.1 Introduction

This chapter describes the operational details of the OK program. As mentioned in Chapter 1, the treatment comprised merit scholarship eligibility, supplemented by academic advising support, with the design of each component based on the successful portions of Project STAR. However, the OK program differed from STAR in five major ways:

1. The incentive scheme was linear, with small rewards at a number of grade levels instead of large rewards at only a few grade targets.
2. Only students on financial aid were eligible.
3. The scholarship formula was the same for everyone, rather than being based on high school grade average.
4. Academic support services included peer counselling, as in STAR, but no review sessions.
5. There was only one treatment group, which received merit scholarships and academic support services, instead of three treatment groups receiving only merit scholarships, only academic support services or both together.

The following sections describe the OK treatment program in more detail.

2.2 The Incentive Component

The first component of the OK treatment program consisted of merit scholarships. There are several options for designing such scholarships, but, as mentioned in Chapter 1, we wanted our design to be easily understood and to provide clear incentives for a broad range of students. For each one-semester course (with a full course load being five courses worth 2.5 credits), students received $100 for obtaining a grade average of 70 per cent and $20 for each percentage point above a grade of 70 per cent. For example, a student receiving a grade of 76 per cent would have received $220 ($100 for attaining the 70 per cent threshold and an additional $120 for achieving 6 percentage points above the minimum at $20 for each point). If this student received 76 per cent in all of her 10 courses over the school year, she would have received a total of $2,200 (equal to 10 × $220).

We felt that providing incentives in each course at every grade level above a fairly low threshold would sharpen the merit scholarship incentives compared to the ones offered in Project STAR and would address some students' concerns that the Project STAR targets seemed unattainable. Offering awards for each course also reduces incentives to take fewer courses (and, conversely, increases incentives to take more courses). Furthermore, the linear incentive structure creates incentives to perform better at every grade level. In Project STAR, there was no incentive to improve

for students with grade averages above 84 per cent and no incremental gain from improving grades between the $1,000, $2,500, and $5,000 targets. In addition, the smaller incremental gains for each course in the OK Project could cause students to focus more on individual tests throughout the term.

2.3 The Advisor Component

In our earlier work with Project STAR, we found significant effects on academic achievement (for women in particular) as a result of offering both merit scholarships and student services, while offering merit scholarships alone led to only short-lived effects. Most participating students in the combined program of Project STAR utilized the student advising services but not the supplemental instruction. Interaction with an advisor was greater in cases where advisors were of the same gender as the participant. Based on the possibility that student advising served as an important complement for generating these effects, we therefore incorporated advising into the Opportunity Knocks treatment program.

With few exceptions, OK participants selected for treatment were assigned a peer advisor of the same gender and were offered opportunities to engage in e-mail correspondence with their advisor to discuss academic matters, as well as issues arising from campus life. Peer advisors were paid, upper year students or recent graduates with successful records of academic achievement. Candidates were interviewed after being recommended through UTSC's Academic Advising and Career Centre. All candidates had previous experience through a UTSC summer program, in which they offered advice to incoming first year students, and they possessed broad-based understanding of campus student support services. In all, we hired eight extremely enthusiastic and knowledgeable students (three males, five females) from diverse backgrounds, all of them excelling in multiple areas of their studies. Advisors participated in a one-day training session on how to provide information and guidance specific to the Opportunity Knocks program. They could contact the OK project manager and the director of the Academic Advising and Career Centre at any time for advice and met with them periodically throughout the year. Advisors were also trained to identify circumstances that called for professional intervention and to make appropriate referrals.

Each peer advisor was assigned 50 students selected for the OK treatment program. Advisors were the key front-line service and information providers for OK participants. They proactively sent e-mails to advisees approximately once every two to three weeks, whether or not a response was acknowledged. In these e-mails, they offered advice about upcoming academic events and workshops

and about how to approach particular periods in the academic calendar, such as midterms and finals. Advisors also provided information about the Opportunity Knocks scholarships, including reminders about payment schedules and about how scholarships were calculated. In addition, advisors responded to advisees' e-mails and could meet with advisees in designated public areas. To gauge the levels of response received, advisors completed monthly logs regarding all correspondence with students. Recruiting materials and sample correspondence with advisees (anonymized) are provided in Appendices A, C, and D (all appendices available upon request).

2.4 Program Implementation

The target population for the OK treatment program included full-time students beginning first and second year who were receiving or eligible for financial aid and who were enrolled in at least 2.5 credits (half of a full load) at the start of the fall 2008 semester. These criteria were checked using administrative data in August and/or early September 2008. All first-year students who had applied for financial aid as of August 12 and all second-year students who had either applied for aid that year or received aid in the previous year (2,063 students) were mailed a brochure about the Opportunity Knocks Project treatment program on August 20, noting that eligible participants would receive a $20 university bookstore gift certificate for signing up and that participants would then be randomly assigned to the treatment program. The layout of this brochure and other copies of recruitment materials are also provided in Appendix A (available upon request).

Signing up involved visiting the OK website, www.utoronto.ca/opportunityknocks. This website provided information about the program, frequently asked questions (and answers) and contact information. Students were also directed to the survey page to participate. They were required to log in, using their university identification number and password. After logging in, they were asked to consent to allow their student survey and administrative data to be used for purposes of researching the impacts of the program. Consenting students were then asked a small set of questions about their family background, financial aid concerns, education aspirations, understanding of how the scholarships would be calculated and grade expectations within two scenarios: whether they were selected into the treatment program or not. Participants were notified by e-mail and surface mail as to whether they had been selected into the scholarship program or not, and those who had been selected where told how they could access their $20 university bookstore gift certificate.

A number of additional methods were used to help recruit participants. Posters about the

program were displayed around campus. On August 26, e-mails were sent to the same students to whom the brochures had been mailed, and these e-mails were also sent to additional students, who had since applied for financial aid (2,129 students in total – including both those who had previously applied and the new applicants). The e-mails briefly described the program and also invited students to visit the OK website. A short presentation about the program was given during first year orientation on September 2 to all attending students. Information sessions, where students or parents could attend, were held in the evenings on September 8 and September 15. Among students who were sent an e-mail or brochure, those who had not signed up by September 16 (933 students) were sent a "last chance" e-mail to do so by September 18. In total, 1,429 signed up. Among those who signed up, we kept those who had applied for financial aid for the 2008 school year, had a recorded high school grade average, had an e-mail address and were enrolled in at least 1.5 credits on September 18. This provided us with the program sample of 1,271. Students were randomized that day into the treatment or the control group. We wanted to make sure that we could test the effects of the program specifically for women, men, first year students and second year students, so we assigned 100 students to the treatment group at random from each gender-year cell. For example, 246 male first year students signed up. We chose 100 of these students at random for the treatment group and assigned the other 146 to the comparison group. We followed the same procedure for female first year students, female second year students and male second year students.[4] This selection process yielded 400 students in the treatment group (100 in each of the four gender-year combinations), leaving 871 students in the comparison group. Twelve participants later dropped all their courses or de-registered for the fall 2008 semester, and although some of these students returned for the winter 2009 semester, a total of 68 participants had left school by the winter term. However, these dropouts occurred at similar rates in both the treatment and the comparison groups, so we doubt that the departure of these students affected the comparability of the two groups.

Participating OK students were notified by e-mail about their selected group on September 23. Treated students were asked to respond to acknowledge they had received the acceptance notification, but unfortunately not all of these replies were recorded. Treated participants were also sent a letter by surface mail indicating that they had been selected for the merit scholarship program. Included in the letter was a wallet-sized card to help remind them of the program and how scholarships would be

[4] In fact, we stratified our treatment group selection on high school grade quartile within each gender-year cell as well, taking 25 participants from each quartile for the treatment group.

calculated. Assigned advisors also sent participants welcoming e-mails and congratulated them on being selected for the program. On October 16, we sent e-mails to treated students who had incorrectly responded to at least one of the two sign-up survey questions about how the scholarships were to be calculated and reminded them of these formulas with examples. Throughout the year, peer advisors remained proactive in sending encouraging words, offers of assistance and reminders of the OK program incentives as upcoming tests and major assignment deadlines, midterms and final exams approached during the academic term.

Scholarship amounts were calculated at the end of each semester. Students could log in to the OK website to see their expected merit scholarship award. (E-mails were sent, notifying students of the website in early January, and advisors were instructed to let their advisees know about the website.) Letters were also mailed in early January, notifying students of the actual amount of their fall semester merit scholarships. Awards were credited to students' university accounts in early February, and if there was no outstanding balance (which occurred in some cases), the amount for each student was deposited directly into the student's bank account or it was mailed as a cheque. Winter semester awards were not processed until July. Letters and e-mails were sent, letting students know the amount of their awards and congratulating them on completing the Opportunity Knocks treatment program. All scholarship amounts were tax free.

Chapter Three OK Participant and Location Description

3.1 Overview of Location and Other Available Academic Services

The University of Toronto Scarborough (UTSC) enrolls approximately 10,000 students, with about 2,500 incoming first year students, most of them from neighbouring high schools. It is primarily a commuter school. Roughly 90 per cent of students in our sample live at home with parent(s). The university is also ethnically diverse. Fifty-eight per cent of our treatment sample listed a non-English mother tongue. Many students are from Asia, South Asia and the Middle East. The programs offered are largely bachelor of arts and science programs, with an option to take cooperative education.

The proportion of students who were in first year in 1996 at UTSC and who graduated within six years since entry (i.e., by 2002) was 70.3 per cent. UTSC students who attempt 4.0 full-year credits and who have a GPA less than 1.50 (between a D+ and a C- average) are placed on academic probation. Students on probation must raise their Cumulative GPA above 1.50 by the end of the following school session (fall, winter or summer term) or attain a sectional GPA of more than 1.70 to avoid a one-year

suspension. Students are usually required to have a 1.85 Cumulative GPA (between a C- and a C average) to graduate. Twenty-three per cent of our control group were placed on academic probation or withdrew by the start of the second year of their program.

As with all Canadian universities, UTSC makes available to all students a range of academic and student services. Student orientation occurs prior to the start of first year classes, with multiple workshops offered about campus life, class scheduling, time management and where to go for academic counselling. Academic advising by staff at the registrar's office is offered, on occasion, on a walk-in basis and group basis. Financial counselling is also offered. The university provides assistance to students who have physical or learning disabilities by, for example, providing alternative test and examination arrangements, note taking services, volunteer services and evaluation of special needs. The university's Academic Advising and Career Centre lists job postings, provides resumé critiques and career counselling, and arranges employer information sessions and job fairs. The Health and Wellness Centre can provide first aid, pregnancy tests, annual physicals, treatment for minor illnesses and over-the-counter medications. Personal counselling (including crisis counselling) is also available at the Health and Wellness Centre to help with family problems, relationships, anxiety, depression, stress management, sexuality, bereavement and eating disorders.

3.2 Overview of Socioeconomic Background Characteristics of OK Participants

The participants in the OK treatment program are all full-time students who receive need-based financial aid. Tables 1, 2 and 3 provide more detailed information about participants' demographic and academic backgrounds, goals, attitudes and aspirations.[5] This information is a combination of data from administrative records and from our pre-program survey. In each table, the "Control Mean" columns give average and standard deviation statistics for the entire control group as well as for gender-year subgroups. Since students were randomly assigned to the control and treatment groups, these control group averages are generally representative of the whole sample averages as well. The "Treatment Effect" column reports the difference between the treatment and control group averages, though these differences are statistically significant in only a few cases.[6] In this section and the

[5] More detailed survey responses can be found in Appendix B (available upon request).

[6] These coefficients are obtained from a regression of the characteristic of interest on a variable equal to one if the student was in the control group and zero otherwise, with controls for each gender-year-high school grade average quartile cell. We have to control for the gender-year-grade average quartile cells because we selected our sample randomly within these cells, rather than randomly from the whole population. The fact that the coefficients on the treatment dummy are rarely statistically significant suggests that our randomization was effective in generating treatment and control groups whose

following section, we use the control means to describe the background characteristics of the entire participant population.

Looking at Table 1, we see that the program participants are mostly college students of traditional age (first years are just over 18 years old, on average, and second years are just over 19 years old, on average). These students had average grades around 82 per cent in high school, which translates into an A-minus average at most schools. Second years, who have already received one year of university grades, continued to receive B-minus grades, on average, in college (first year average GPA is 2.648), although this corresponds to a lower percentage grade at UTSC (a grade between 70 and 72 per cent earns a B minus at UTSC).

Notably, many students represent the first generation in their family to be born in Canada; only 5.9 per cent of participants' mothers and only 5.4 per cent of participants' fathers were born in Canada. As we might expect, given these statistics, only 41.6 per cent of participants speak English as their first language (though presumably all participants speak English quite well). Both of these numbers are quite a bit lower than the average for Canadian university students (Association of Canadian Community Colleges and Human Resources and Social Development Canada, 2007). Still, the study participants come from reasonably educated families. About half of participants' parents graduated from PSE (43.9 per cent of mothers graduated and 53.2 per cent of fathers graduated), while nearly 80 per cent of their parents graduated from high school, which is similar to the average parental education levels for all Canadian university students.

Moving to Table 2, we see that 87.7 per cent of the participants live with their parents – as we would expect, given the financial constraints on these individuals and the low rates of on-campus living among the entire UTSC student body. Participants' financial need is also evident in that 43.6 per cent of the participants plan to work during the school year (some in a co-op work program during the winter semester) and in that participants expect about $6,500 in loans and grants on average (though this information is incomplete, since, at the time of our survey, many respondents were unsure of how much financial aid they would receive).

3.3 Goals, Attitudes and Aspirations of Participants

Table 2 also describes the goals, attitudes and aspirations of our participants. A majority of participants

---

members were about the same, on average, in terms of measurable characteristics. Thus, we can confidently attribute any difference in treatment and control outcomes to the OK program, rather than to differences in the treatment and control populations.

(57.2 per cent) stated that they were very concerned about having sufficient funds to complete their university degree, reinforcing the fact that these students face serious financial constraints regarding educational opportunities. For most students (76.7 per cent), UTSC was their first choice school, and many of our participants intend to eventually complete a graduate degree (43.3 per cent). Thus, it seems that our study population has ambitious academic goals but perhaps lacks the funds or academic performance to achieve these goals.

Table 3 examines whether participants understood the OK merit scholarship formula and how much of an effect they thought the program would have on their academic performance. We asked two questions concerning their understanding of the merit award formula, which rewarded students $100 for each course with a final grade of 70 per cent or higher and an additional $20 for each percentage point above 70 per cent attained in each class. The first question asked students to calculate the award amount from a grade in a single class; the second asked them to calculate the total award amount from grades in five different classes. We had hoped that the award formula was easy to understand. Also, if participants didn't understand the merit award formula, it would be difficult to interpret the academic impacts of the OK program. Fortunately, 83.9 per cent of participants answered the first, simpler scholarship calculation question correctly (although only two-thirds answered the more difficult question correctly). We sent follow up information about the merit scholarship formula to any participant who responded incorrectly to at least one of the two questions. In the program analysis, we look at academic effects for the entire sample and for the subgroup that answered the first question correctly to see whether those who understood the merit scholarship better also had a stronger program response.

To determine whether students thought the OK program would help them increase their grades, we asked students what grade they expected to receive if they were selected for the treatment group and what grade they expected if they were not selected. (To help solicit correct expectations, we offered a $500 gift certificate for the student who came closest to his or her actual grade average.) Interestingly, fewer than half (37.3 per cent) of the participants thought that they would have a higher grade if they were selected for the program, but the average difference between the expected GPA in the treatment group and the expected GPA in the control group was significant (0.195 GPA points, or about 3 percentage points on a percentage scale), suggesting that participants on average expected the program to be somewhat helpful. Figure 1 shows smoothed kernel densities of the grade distributions for expected grade averages (in per cent) for first and second year men and women. Of those who did

report differences in expected grades as a result of being selected for the treatment program, most said their grades would increase by 5 to 10 percentage points.

Chapter Four Impacts on Sign-Up and Advisor Use

4.1 Sign-Up Impacts from Offering OK

Since we chose our treatment and control groups after students signed up to participate in our study, we were interested in which types of students signed up for the program. In particular, when reflecting on the academic effects of the OK program, we should remember that these effects are primarily relevant to the types of individuals who signed up to participate. The effect of the program on other types of individuals, such as those who did not sign up, might be different. For example, if the OK merit awards were made automatic for all financially needy individuals at UTSC, we might expect a smaller average effect than we measure here, since students who signed up for the program likely did so because it would provide strong motivation for them to get higher grades. Likewise, those who didn't sign up may have done so because the program wouldn't have given them much additional motivation. We are also interested in who signed up for the OK program simply so that we know what types of students are interested in merit awards and academic support services of this kind.

The first column of Table 4 displays average characteristics for those who signed up for the OK program. The second column shows the same averages for those who didn’t sign up and the third column reports the differences between the two. For example, the first row tells us that the sign-up population was 65 per cent women, while the population that didn't sign up was only 58 per cent women (7 percentage points less). This means that women signed up at higher rates than men. Those who signed up also had slightly higher average grades in high school, – by about 3 percentage points. There are no differences between the two populations in terms of year in school or having English as their first language, but students who signed up attempted about 0.2 to 0.3 more credits each semester on average than those who didn't sign up. The fourth column regresses each characteristic on a variable equal to one if the student signed up and zero if she or he did not, controlling for each combination of gender, year and high school grade average quartile. A positive coefficient in this regression suggests that, within each gender-year-grade quartile cell, those who signed up had a higher value for that characteristic on average. Because we randomized participants within each of these cells, we are most interested in sign-up differences that remain within the cells. The fourth column regression coefficients confirm that, even within cells, those who signed up took more credits and had higher high school

grade averages (these coefficients are positive and statistically significant).[7] We do not look at whether women or first year students were more likely to sign up in this framework, since they are part of the cell controls, and differences in their sign-up rates are less important for the interpretation of our results, since we can easily estimate different effects of the OK program for these groups due to our randomization design.

The raw differences between those who signed up and those who did not in gender ratio, high school grade average and attempted credits are not surprising, since women, those with higher high school grade averages and those who take more credits generally earn higher grades in PSE and therefore stand to benefit more from the program. Since the merit scholarship awards are given for each course, those who take more credits are also likely to earn higher awards.[8] This information is useful in that it gives us some idea of who might sign up for a merit scholarship and academic support program at another university. In addition, we must keep in mind the differences in background characteristics within gender-year-grade quartile cells when we weigh the academic effects of the OK program in subsequent chapters. These academic effects are representative of this sign-up population and can be extrapolated only to other populations that are similar. That is, the overall effects that we measure are relevant only to similar populations that have higher high school grade averages and attempt more credits than the average across all UTSC students.

4.2 Advisor Contacts in the OK Program

We included the peer academic advisor component of the OK program because students gave good feedback on this portion of Project STAR and because the combination of academic support services and merit scholarships increased the size and duration of program effects in Project STAR. In the OK program, we collected information on participant-advisor communication and other program-related communication to help assess students' use of the services and level of program engagement.

Table 5 reports the fraction of students across various gender-year subgroups that made various forms of program-related contact. Although only 38 per cent of all treatment group participants sent an e-mail to their advisor in the fall, this number increased to 50 per cent during the winter, and 70 per cent of students sent at least one e-mail over the course of the whole year. As we might expect, a higher

[7] Statistical tests (not shown) find that those who take more credits or have higher high school GPAs do not respond much differently to the OK program. This finding suggests that our results should generalize to a broader population.

[8] One caveat regarding this discussion is that those who take more credits may have lower grades because of their larger workload.

percentage of first years made use of the advisors. Women were also more likely to contact the advisors over the course of the whole year. These results suggest that most students found the advisors helpful at some point. The table also shows that 72.5 per cent of all treatment group students checked their merit scholarship earnings on the program website, and women checked a bit more often than men. Taking into account this checking and any other program-related contact, we find that 87 per cent of treatment students actively engaged with the OK program in some way. This result suggests that the vast majority of participants were aware of, and interested in, the program.

Chapter Five Impacts on Course Enrolment

As discussed earlier, the OK program awarded students money for each class in which they received a 70 per cent grade or higher. We hoped that this formula would encourage students to attempt more credits and therefore progress toward degree completion more quickly. Table 6 presents the effects of the OK program on course enrolment for the fall 2008 and winter 2009 semesters and for the full 2008/09 academic year. The last column of the table shows the estimated effect of the OK program for the entire sample population, and columns to the left display effects for various gender-year subgroups of the participants. The Control Mean rows display the average number of credits enrolled in by the control group so that we can judge whether the effects of the program are large or small relative to how many credits are usually attempted. The Treatment Effect rows display coefficients from linear regressions of number of credits attempted on a dummy variable equal to one if the student was in the treatment group and zero if in the control group. These coefficients give us the effect of the OK program on the average number of credits attempted. For example, the Fall Treatment Effect row for first year women has a coefficient of -0.039, which implies that the OK program reduced the average number of credits attempted by first year women by 0.039 (although this effect is not statistically different from zero). In each regression, we also include controls for gender, students' year in school, high school grade average and grade quartile, whether the student's first language is English or not, parental education and whether they answered the first test question on the earnings formula correctly.

We can see from the Control Mean rows of Table 6 that control group students attempted about two credits each semester on average. The Treatment Effect rows show that the OK program had a positive effect on credits attempted in both semesters for most subgroups. However, these effects are only statistically different from zero (denoted by asterisks) in some cases. In contrast to previous literature that finds larger effects of academic programs on women than men, we find the strongest

effect on credits attempted by first year men in the winter semester. However, this increase of 0.135 credits (significant at the 5 per cent level) is still only about one-quarter of a standard deviation. We also find smaller, marginally significant effects in the winter semester for all men, all first years and the whole sample – and in the fall semester for second year men (significant at the 10 per cent level). Most other effects are positive but insignificant. Since we do not find significant effects across both semesters for any single subgroup, the results for the full academic year are predictably weaker, though the estimates remain positive in most cases.[9]

In summary, it seems that men, particularly first year men, did respond to the OK program by attempting more credits, although the differences are small. The fact that we find stronger effects on enrolment during the winter semester might suggest that students understood their earnings potential in the program better after receiving payment for their fall courses or that the peer academic counselling gradually encouraged them to be more ambitious. Under this interpretation, we would expect the positive effects on winter credits attempted to persist if the OK program were implemented for longer than one year.

Since treatment students attempted about the same number of credits as control students, we do not have to worry about differences in workload when calculating program effects on scholarship earnings, average grades and other outcomes. If we find a positive effect of the program on grades, for example, we might worry that students worked the same amount but decreased their course load in order to improve their grades. However, we find little effect on course loads, so we do not need to worry about this alternative channel for program effects on grades.

Chapter Six Impacts on Academic Achievement during the Program Implementation

6.1 Impacts on First and Second Semester Earnings and Grades

The remaining tables (Tables 7 through 19) are organized similarly to Table 6, with the Control Mean rows presenting averages and standard deviations within the control group for various gender-year subgroups to give a sense of scale and with the Treatment Effect rows displaying the measured response of the treatment group average to the OK program. All control variables in the regressions for the Treatment Effect columns in the rest of the tables in the report are the same as in Table 6.

A natural benchmark for program effects is a comparison of the amount actually earned in the

[9] Full-year effects will not necessarily be equal to the sum of the winter and fall semester effects, since some students dropped out of UTSC during the experiment.

control group with the earnings controls they would have been entitled to had they been in the program.[10] A large program effect should be reflected in larger-than-expected earnings, where expected (i.e., hypothetical) earnings are measured using the grade distribution in the control sample. In this vein, Tables 7 and 8 analyze program effects on total scholarship earnings from the program and average grades across all courses, respectively, where we use the earnings that control group students would have received, given their grades in each class. We should perhaps expect the largest effect of the program to come from merit scholarship earnings, since earnings represent the actual financial incentive to improve grades. Nevertheless, the academic counselling component of the program may affect average grades more than merit scholarship earnings.

The results in Table 7 suggest that the OK program had no effect or a very small negative effect on earnings for first year men and women. However, the effects on second year men and women are uniformly positive and sometimes statistically significant. In particular, the effect of the program on winter earnings for second year men is about $171 (significant at the 5 per cent level), and the effect on full-year earnings for all second years is about $185 (significant at the 5 per cent level).[11] These responses are only about one-quarter of a standard deviation of earnings, however. The measured effects for the whole sample are positive in both semesters and over the full year, but they are not statistically significant. Overall, it seems that the OK program had a small positive effect on earnings for second years, particularly second year men, but no effect on earnings for first years.

Table 8, which analyzes the effects of the OK program on average grades, reinforces the results in Table 7. The similarity in the results is not surprising, since higher grades result in higher earnings. However, although generally of the same sign, the effects on grades are a bit weaker than the effects on earnings. Average grades for second year men in the winter responded positively to the program, but the increase in their average grades of 2.551 percentage points is only significant at the 10 per cent level. Most other effects for second years are slightly positive or very close to zero. As for earnings, the effect of the OK program on first years' grades is near zero and even slightly negative during the winter semester (though always far from statistically significant), and the effects on the entire sample suggest that the OK program had very little effect on average grades overall.

---

[10] Ashenfelter and Plant (1990) use these “hypothetical payments” to measure the labour supply effects of exposure to a negative income tax.

[11] If we restrict the fall and winter samples to be the same as the full-year sample, then the treatment effects for the full year are the sum of the fall and winter treatment effects. Estimated effects for the full year may not be the sum (or average) of the two-semester effects in this and all other tables in the report, since the sample for the full year is generally different from the sample for either semester (the full-year sample excludes students who dropped out of school during either semester).

It may seem puzzling that the OK program could have a larger effect on scholarship earnings than on grades. We suspect that this is because earnings more accurately capture students' financial motivation from the scholarships. For example, students in the OK program have the strongest incentive to increase their grades around the 70 per cent threshold, since they earn $100 for crossing above 70 per cent and only $20 for each percentage point above 70. If students increased their grades only near the 70 per cent mark, then looking at differences in grades alone may capture less variation than looking at program earnings, which assign a big jump at the 70 per cent level. To test whether this feature of the earnings function can explain the different results for grades and earnings, we can decompose the effect on earnings into an effect on the number of courses where students earned more than 69 per cent (each course over 69 per cent earns $100) and an effect on the total percentage points earned over 69 per cent across all classes (each percentage point earns $20). The stronger effects on earnings above suggest that we will find stronger effects on the first component (which impacts earnings more heavily) than on the second component (which impacts grades and earnings in equal proportions).

Table 9 presents program effects on the first component. The dependent variable is equal to the number of courses in which a student achieved a 70 per cent grade or higher. Again, the coefficients presented in the Treatment Effect row are from regressions of this dependent variable on a variable equal to one if the student was in the treatment group and zero if he or she was in the control group. Since we found stronger effects on earnings than on grades, we expect to find that treatment students achieved a 70 per cent grade in a higher number of courses. The results in Table 9 validate this claim. Second year men in the treatment group achieved a 70 per cent grade in 0.598 more winter classes than in the control group, on average (statistically significant at the 5 per cent level). This translated into an increase of almost one whole class above 69 per cent on average over the whole year (statistically significant at the 5 per cent level). The OK program had similar, smaller effects on the number of courses over 69 per cent for all second years.

Table 10 reports the effects of the OK program for the second component of the earnings formula: the total number of percentage points earned above 70 per cent across all classes. The coefficients in the Treatment Effect column now estimate the difference between the treatment and the control group in the total number of percentage points above 70 per cent across all classes. We expect to find smaller effects on this second earnings component, given that there is little overall effect on grades. The OK program increased the percentage points above 70 per cent by 4.080 points for second

year women in the fall semester and by 5.557 points for second year men in the winter semester (about 1 point per class in both cases), but these effects are only significant at the 10 per cent level. These translate to a full-year effect of 6.353 percentage points for all second years over the full year (also significant at the 10 per cent level). However, there seems to have been a great deal of variation in the treatment responses, since these effects are still about one-sixth to one-fifth of a standard deviation but are only marginally statistically significant. In addition, the OK program actually decreased the total percentage points above 70 by 2.356 points for all first years in the winter semester. These weak results match up well with the results for average grades in Table 8.

Taken together, the results in Tables 9 and 10 help to reconcile the larger effects for earnings in Table 7. It seems that most of the program effect on earnings was due to an increase in the number of courses with grades above 69 per cent. Grades did not shift up as much above the 70 per cent threshold – despite the $20 incentive per percentage point – so the program had very little effect on grades overall. Although the OK program increased grades and earnings slightly for second years (particularly for second year men during the winter semester), the program had very little effect when we look at all participants together. None of the estimated effects are particularly large, though the earnings decomposition analysis provides reasonably strong evidence that second year men responded to the financial incentives in the OK program. The larger incentive ($100) at the 70 per cent level seems to have been more effective than the smaller incentive ($20) above 70 per cent. Lastly, bigger effects in the winter semester are consistent with the results for credits attempted, which stresses that students may have understood the program better after one semester of experience and/or that the academic counselling portion of the OK program took some time to bear fruit. In both cases, a more permanent implementation of the OK program might have larger, sustained, positive effects.

These strong results for the winter semester also suggest that we might find larger program effects if we focus on the participants who understood the program from the beginning. One way to identify these participants is to restrict the sample to those who responded correctly to the first test question about how earnings would be calculated. Tables 11 and 12 present the same effects as in Tables 7 and 8 on merit scholarship earnings (hypothetical for the control group), qualification for the scholarships and average grades, respectively, but restricted to this subsample. As a reminder, this question asked students to calculate earnings from a grade in a single class. If students answered the question incorrectly, we might wonder if they understood the OK program well enough for it to affect their academic effort and performance. According to this reasoning, the sample that answered the

question correctly should show a larger response to the OK program.

Table 11 indeed repeats the positive, statistically significant effect of the OK program on winter semester earnings for second year men and finds a positive, statistically significant effect of the program on fall earnings for second year women for the restricted sample. As in the full sample, effects for first years are generally slightly negative but far from statistically significant. Most striking is that, within the restricted sample, the effect of the OK program on all second years is positive for both semesters and for the full year. Full-year earnings are about $259 higher (significant at the 1 per cent level), on average, for second years in the treatment group, which is about one-quarter of a standard deviation.

Table 12 presents results for average grades, which are similar to the effects on scholarship earnings. Second year men had a positive, marginally significant grades response to the OK program in the winter semester, and all second years had a positive, marginally significant response to the program in the winter semester and over the full academic year. These effects are about 1.3 to 2.7 percentage points, or about one-fifth to one-quarter of a standard deviation. From these results, it seems that second years who understood the scholarship formula did, in fact, respond to the OK program more strongly, though first years did not, and the second year effects are still small in magnitude.

We again find that the effects on average grades are weaker than the effects on scholarship earnings. To explain this difference, we break the earnings effect into its two components for the restricted sample in Tables 13 and 14. In Table 13, as expected, we find that the OK program increased the number of courses above 69 per cent noticeably for second year women, second year men and all second years in both semesters. The increase for second years over the full year was 0.816 courses (significant at the 1 per cent level). There are very few negative effects estimated for this restricted sample, and none are statistically significant. In fact, the number of courses with grades over 69 per cent in the treatment group is higher for the entire (restricted) sample by 0.161 courses in the fall semester and by 0.371 courses over the full year (significant at the 10 per cent and 5 per cent levels, respectively).

Table 14 analyzes the program effects on the total percentage points above 70 per cent, and, predictably, the effects are not as strong as those for the number of courses above 69 per cent. There are a variety of marginally significant effects between 3 and 9 percentage points for second year women, second year men and all second years. The only effect significant at the 5 per cent level is an increase of about 9 percentage points above 70 per cent for all second years over the full year.

This analysis strengthens the conclusion that the OK program made the biggest positive impact on grades near the 70 per cent threshold. These positive effects appear to be a response to the strong incentive ($100) to get a grade above 69 per cent in each class. Also, the effects are stronger for the population that demonstrated better understanding of the award formula by answering the first test question on merit scholarship earnings correctly. The smaller $20 incentive for each grade point earned above 70 per cent seems to have had some positive impact on grades in this restricted sample, but the effect is less pronounced.

6.2 Impacts on Earned Course Credits

We turn now to Table 15, which presents the effects of the OK program on credits earned each semester and over the full year in courses eligible for scholarship earnings.[12] We are especially interested in whether the OK program increased the number of credits earned by treatment students, since degree completion and retention are a significant concern at UTSC and many other universities. As we might expect, the responses of credits earned to the OK program are similar to the responses of average grades. Most program effects are very close to zero, but the OK program had positive, insignificant effects on credits earned for second year men, a marginally significant effect of 0.152 credits earned for first year men in the winter semester and a marginally significant effect of 0.105 credits earned for all men in winter. These increases are only about one-fourth to one-seventh of a standard deviation. The effects of the program on average credits earned for the entire sample are slightly positive, but not statistically significant. Given that the OK program primarily affected grades at the 70 per cent threshold in the analysis above, these results are not surprising. The scholarships probably seemed too unattainable to motivate students with grades near the 50 per cent threshold.

6.3 Impacts on Probation and Academic Standing

In addition to looking at credits earned as a measure of completion and retention, we can examine whether the OK program had effects on the likelihood of being on academic probation or in good academic standing at the end of the 2008/09 academic year. A student is placed on academic probation at UTSC if their GPA for the academic year is below a certain threshold (the threshold may depend on their academic program, but it is generally around 1.6, or about a C minus). A student is considered to be in good academic standing if their yearly GPA is above a threshold (often the probation threshold)

[12] This excludes pass-fail courses.

and they have attempted a minimum number of credits (usually at least 1.5 credits per semester). Remaining on academic probation for too long results in suspension, meaning that these categories give some idea of whether students are progressing toward their degrees, as well as whether they are achieving good grades.[13]

Table 16 shows the effects of the OK program on academic standing (good standing or probation). We find essentially zero effect of the program on either measure of academic standing for all subgroups and for the entire sample. Since the main effects of the program appear to be for second year men, particularly in the winter, it is not surprising that we find no effects here on these full-year academic measures. Another possible reason for the small effects on probation for second years is that very few second years (only about 1 per cent of the control group) are on probation. This is likely because many students on probation do not stay in school for a second year.

6.4 Distributional Impacts on Grades

Figures 2 through 7 display the smoothed distributions of the outcome variables between our treatment and control groups for first and second year men and women. The figures help show whether the program affected students at particular points in the distribution but not others. For example, the program may have impacted a high fraction of low-performing students but not high-performing students, leading to a shift in the lower tail of the grade distribution but not the upper tail. Such shift may not be easily revealed from examining only mean differences.

Generally, the grade distributions are similar between treatment and comparison groups. The grade distributions are not significantly different. Hypothetical earnings are marginally significant, corresponding to the mean differences results, which also suggest marginal impacts on the number of courses with grades over 70 per cent.

Chapter Seven Impacts on Second Year Achievement and Retention

Beyond effects on student achievement during the study period, the OK program could also have increased the likelihood of students staying in school, since it offered financial assistance, provided peer counselling and encouraged good grades. In addition, the program may have helped students improve their study skills, which should have led to better long-run academic performance. To measure any such effects, we first look at Table 17, which analyzes whether the OK program affected the

[13] Some students may not fall into either category due to the credit requirements and to program-specific rules.

likelihood of students dropping out of school. The dependent variable in the regressions in Table 17 is equal to one if a student failed to enroll in any courses or dropped all of her or his courses during the relevant time period; otherwise, it is equal to zero. The full-year variable is equal to one if the student dropped out of either semester or both semesters. Thus, the coefficients in the Treatment Effect rows estimate the effect of the OK program on the likelihood of dropping out of school.

Despite the fact that first year men showed no response in academic achievement to the OK program, we find that the program reduces their likelihood of dropout by 2.5 percentage points (significant at the 10 per cent level) during the winter semester and over the full year. More in line with what we have found so far, second year men in the OK program are 4.7 percentage points less likely to drop out in the fall semester (significant at the 5 per cent level). These effects are enough to generate decreases in dropout rates of 2.1 percentage points for all men in the fall semester (significant at the 5 per cent level) and 1.6 percentage points for all first years in the winter semester and over the full year (significant at the 10 per cent level). There is no evidence that the OK program increased dropout rates for any subgroup.

Tables 18 and 19 describe the effects of the OK program on achievement in the semester following the study period (fall 2009). These results will help us determine whether any improvement in study skills or work ethic (due to the program's financial incentives and peer counselling) led to better academic performance down the road. In each table, the dependent variables are listed in the left-most column, next to the Control Mean and Treatment Effect rows, and “first year” and “second year” refer to the students’ standing in the 2008/09 academic year. All variables are for fall 2009, and although we present results on scholarship earnings in Table 18, students did not receive any actual earnings (we calculated hypothetical earnings from their fall 2009 grades).

In Table 18, we see that the OK program had a marginally statistically significant effect on GPA for first year women (0.146 GPA points) and for all first years (0.118 GPA points), as well as a marginally significant effect on first years' average grades (1.333 percentage points).[14] Table 19 shows that the program had very little effect on credits attempted or earned in fall 2009. Nearly all of the effects on academic achievement in Tables 18 and 19 are positive, though not statistically significant. The effects of the program on the likelihood of dropping out are mixed and very close to zero for all subgroups. Given that we did not find very large effects of the OK program during the study period, it

[14] “First year students” continues to refer to students who were in their first year during the OK program study period, meaning that they were actually second years during the fall 2009 semester.

is understandable that we find mixed and mostly statistically insignificant effects of the program during fall 2009.

Chapter Eight Follow-Up Communication with OK Participants

8.1 Focus Group Advisors

The OK project manager, along with two of the principal investigators (Chambers and Oreopoulos) conducted a discussion with all eight advisors at the end of the fall semester (December 8, 2008) to acquire a better understanding of advisors' impressions of their students and the impacts of the program. We asked about their typical experience from sending regular e-mails and about the extent to which students seemed to respond. In general, advisors felt that participants appreciated this regular contact, even though they did not always respond. One advisor, for example, noted that "Students will respond within an hour or a week depending on how busy they are. I remember sending my first e-mail in the early October and got one the beginning of December saying, 'I am sorry I have not responded I have been so busy, but things are going well and I will let you know if something comes up.' Sometimes if you haven't sent anything out in a while, you'll get an e-mail saying this is what is going on or someone is trying to sell me their notes, how do I deal with that? So yah, that's how it has been for me the past few months." No advisor said an advisee asked to stop receiving e-mails.

Strategies for sending out e-mails varied. A sample of (anonymized) advisor e-mails appears in Appendix D (available upon request). In general, advisors sent a variety of messages, some discussing upcoming workshops, others with short words of encouragement and others to remind advisees about the scholarship program. One advisor described her experience this way: "I generally send out e-mails saying, what's going on, having a good week and kind of outline a few seminars or social things that are going on campus and pretty much end saying, e-mail me back if you have any questions. I tried to mix it up by saying, short e-mail, saying hi and want to make sure everyone is cool, and if you want a list of academic and social stuff, e-mail me back and I will make it available. It is all available to them on the Internet and stuff, but I did get some responses saying I missed that e-mail this week, could you send me the list? In that case, I tried that and found that I wasn't getting as many replies. I also tried making the e-mails sound more being individualized being like, hi just wanted to chat. But making it more personal, I got more responses being like hey thanks for checking up and thinking about me and taking the time to e-mail me. So those work, those different avenues work, but a decent amount of feedback and based on the timelines does make a difference." Few advisees requested face-to-face

meetings.

We asked advisors what questions advisees most commonly asked. Many were related to administrative matters, such as switching or dropping a course, study tips and referrals to workshops or services (especially writing-related services). Advisees thought that most questions were related to academic concerns and "less about life things." Asked whether they thought students understood how the scholarship program worked, most agreed that they did (e.g., "They do understand the process – high marks mean higher dollars") but said they often received questions about timing of payments. Similar to findings from our previous experiment (Project STAR), students told advisors they were experiencing stress and frustration because of grades falling below levels they were attaining in high school.

8.2 One-on-One Interviews with Treated Students

We e-mailed 50 randomly selected OK students in the treatment group at the end of the school year (May 13, 2009) to ask them questions about their overall impressions of the program. We offered $25 movie gift certificates for responding. There were 10 questions, which students could answer simply by replying to the e-mail invitation. In all, 30 students responded. More detailed responses are listed in Appendix E (available upon request).

We asked whether students often received e-mails from their OK peer advisor, whether they felt these e-mails were helpful or annoying and whether they e-mailed their advisor specific questions or concerns. The feedback was very positive and consistent with our feedback from advisors. Everyone said they had received advisor e-mails about once or twice a month. All but one of the 30 respondents said the advisor e-mails were helpful. One noted, "They were helpful, not annoying. I think the advisor made good decisions between sending us important reminders and information without being redundant. It was especially important to receive the e-mails about the scholarship money quickly after marks were sent in." Another said, "I find it very useful that someone was actually helping me through school."

Many students commented that the information provided from the e-mails was helpful enough and did not feel that further contact was necessary. One student said, "I never felt like I needed to inquire about anything more then what was already in the e-mails she had send me. The program and its details seemed straightforward enough to me." Some students who did reach out to their advisors had mostly administrative questions ("Around the start of the year, I was a little confused on how the

OK payment was issued, but thankfully, my peer adviser clarified all matters"). A few who did reach out for academic advice felt the feedback was quite worthwhile. One made this comment:

> I did ask for help from my advisor on specific concerns . . . on more than one occasion. One such issue of concern was the marks I was getting in some of my courses. My marks were too low and very de-motivating. I began to question myself, and was seriously considering a program switch. My peer advisor reassured me of how normal my situation was for first-year students, and was kind enough to tell his story of how he was once in a similar position. He basically outlined the steps he took to get from being a stressed out student with low self-esteem (due to poor grades, etc.), to an involved, well-rounded student making the most of his University experience. I found his story to be very motivating and it reassured me that I will in fact get through this school journey. Thanks to some basic study tips and other bits of advice from my advisor, I am more confident and I don't let little misfortunes bring me down.

We also asked about whether students felt the scholarships led to increased motivation. Of the 30 respondents, 27 said the scholarships motivated them. Some thought the program was very effective (e.g. "Every time I began to lose interest in a particular course, I would remind myself that I just need to well . . . keep with it; the rewards will be tremendous. A scholarship is one such reward . . . and it sure is helpful, as it lifts a lot of the financial burdens I'm faced with when it comes to paying tuition & other fees"). Others thought the program was partially effective (e.g., "This scholarship did affect my motivation to study at some point, but I was motivated most when the marks of my first semester were low"). Respondents often cited concern about tuition and other university-related costs as reasons behind their interest in the program.

While virtually all respondents felt the program motivated them, only half felt the program led them to study more. Several students felt their opportunity for more study time was limited (e.g., "[The program made me study more] but not much. I usually follow my schedule between work and school. So the amount of time I could have spent on study is somehow limited"). Others felt the program helped them focus ("As someone who gets sidetracked easily, I kept it in mind that staying focused would pay off in more than one way, and so yes, it did affect the amount of time I devoted to studying"). Another said, "I think what's great about the program is that when you feel like you're beginning to procrastinate, you think about the outcome of this program and want to get back to studying." A second year student felt worse participating in the program after seeing his marks fall compared to first year: "I did abnormally poor this year compared to my usual standards and it just so happened to coincide with Opportunity Knocks. The money reminder just kind of made me feel 'worse'

about myself."

Interestingly, 26 of the 30 respondents felt the program improved their academic performance. Many of them felt the program helped them substantially. The comments lined up with our prior beliefs about how the program might affect performance. Some appreciated the opportunity to earn scholarships for "good" but not necessarily "outstanding" grades: "Most definitely. Personally, I don't find that U of T offers as many scholarship opportunities as other university, I think it was rewarding to know that my academic performance was acknowledged and rewarded." Some felt they increased performance out of financial concerns: "[E]specially now with the economic downfall, it is extremely difficult to muster up the finances to help pay for tuition without relying on OSAP. I kind of looked at Opportunity Knocks as my employer who gives me more money the better I performed in my studies." One student felt the program would have a long-lasting effect: "Definitely! The program had significantly improved my grades! And I cannot wait to see what I can accomplish next year."

We also asked these students whether they felt the program is worth continuing, more for the advisor program or for the merit scholarship program. All but one respondent felt strongly that it would be worth continuing in some form. Most felt both aspects of the program were worthwhile, but some singled out the scholarship program as more important. Asked whether they would change anything about the program, many said no, although a few mentioned they would prefer to have had more face-to-face and informal interaction with advisors. Virtually everyone expressed gratitude from being selected for the program. For example, one student closed their e-mail by saying,

> The OK Program has been an essential part of my student experience, and in many ways crucial to my academic performance. I think that having a peer advisor as opposed to just the regular counselors working in the University is very important. With all the stress that universities cause their students – financially or otherwise, it's really nice to know there is a program like Opportunity Knocks to help students every step of the way.

Chapter Nine Summary and Implications of the Estimated Impacts of the OK Program

9.1 Summary of Impact Estimates

Impacts on Program Sign-Up

A total of 1,271 students were qualified for, and agreed to participate in, Opportunity Knocks. This amounts to about 60 per cent of those originally contacted (in fact, two-thirds signed up, but some of these did not meet program eligibility criteria). Participants were self-selected and do not constitute a random sample of UTSC first year students, but overall, the participant sample seems fairly

representative of UTSC. Although those who signed up in advance of random assignment tended to be somewhat better students as measured by their high school GPAs and courses attempted, these differences are not dramatic and these variables are not strongly correlated with the treatment effects.

Impacts on Student Service Take-Up

Especially encouraging is the fact that among the 400 treated students, program interest and engagement seem high. Roughly 73 per cent of treated students checked their scholarship amounts online, while 87 per cent engaged in some kind of interaction with program staff or the program website. It seems especially noteworthy that 70 per cent of treated students e-mailed their advisors. This level of engagement compares favourably with that in Project STAR.

Impacts on Course Enrolment

OK had only isolated effects on student course load as measured by credits attempted. Across gender-year subgroups, the largest full-year effect on credits attempted is for men, a marginally significant increase – of about 0.16. This is less than a one-fifth of a standard deviation of the distribution of credits attempted among men. An important implication of these modest effects is that there should be little selection bias in estimates of effects on grades, since the nature of student schedules was essentially unchanged in the treatment group.

Impacts on First Year Achievement and Retention

A natural benchmark for program effects is a comparison of the amount actually earned in the treatment group with the earnings control group participants would have been entitled to had they been in the program. (Ashenfelter and Plant [1990] use similar “hypothetical payments” outcomes to measure the labour supply effects of exposure to a negative income tax.) A large program effect should be reflected in larger-than expected earnings, where expected earnings are measured using the grade distribution in the control sample.

The OK program appears to have had modest effects on program earnings, with the largest full-year effect appearing for second year students and men (especially, second year men). These effects are on the order of 15 per cent of the control mean, though most are only marginally significant. The full-sample effect on hypothetical earnings is a statistical zero.

The OK program did not increase student achievement as measured by average grades: the full-

year effect is -0.037 percentage points, with no strongly significant effects in any gender-year subgroups or other subsamples. Likewise, the program had little effect on credits earned.

The juxtaposition of zero effects on average grades and credits earned with modest earnings effects raises the question of how these different results can be reconciled. An analysis focusing on the components of the OK award formula shows that the OK program boosted the number of courses in which students earned a grade of 70 per cent or better. Like the earnings results, this effect was concentrated among second year students, especially men, where the increase in courses above 69 per cent amounted to almost a full course. Thus, OK-treated students appear to have been motivated to clear the key program hurdle. At the same time, their effort to obtain the award cutoff mark was not enough to boost overall GPAs.

It is also noteworthy that the OK program appears to have had somewhat stronger effects on the subpopulation of students who responded correctly to a question designed to gauge student understanding of program rules (about 84 per cent of the treated sample answered this question correctly). For example, among second years who answered the question correctly, treatment students earned a grade above 69 per cent in 0.8 more courses than control students, a highly precisely estimated effect. The overall program effect on courses with a grade above 69 per cent is also significantly different from zero in the subsample that understood the earnings formula, but the overall effect on average grades of 0.41 percentage points is not.

Impacts on Second Year Achievement and Retention

The OK program had very small effects on achievement in fall 2009. This is understandable, given that the program had small effects during the study period itself. The program did decrease dropout rates somewhat for men during the study period, with the largest effect (a 4.7 percentage point decrease) occurring for second year men in fall 2008.

9.2 Interpretation and Implications

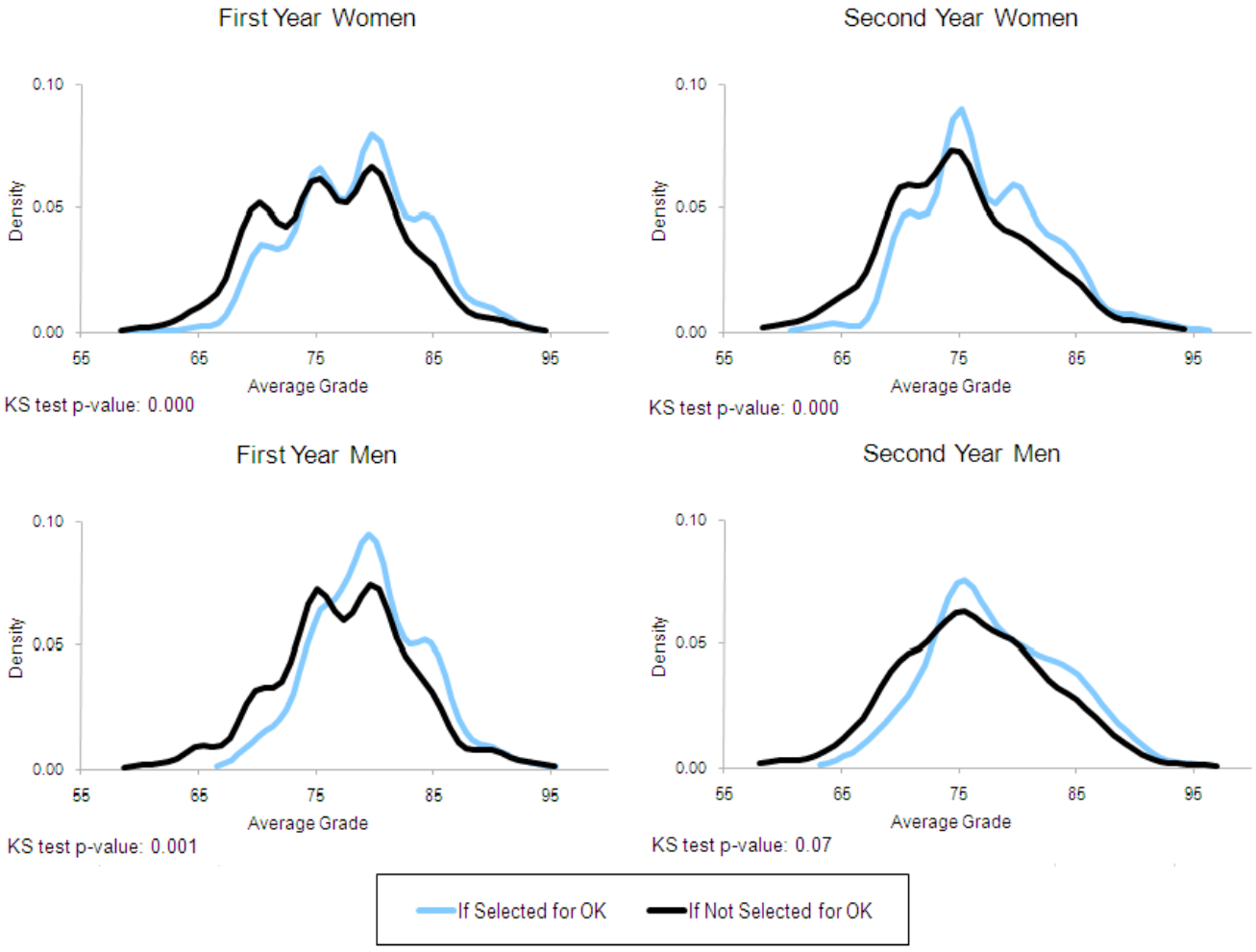
The OK program was popular with participants: sign-up rates and program engagement were high, and many program participants were enthusiastic about the experience. At the same time, consistent with an emerging body of evidence evaluating interventions of this type, overall program effects on achievement were modest. Treated students earned more in OK scholarship money than we would have expected based on the control grades distribution, but this appears to be largely due to an effect on the

number of courses in which students earned a grade of at least 70 per cent, especially among second year students. This localized response did not translate into a substantial boost in overall achievement, but it was noticeably stronger in a large subsample of students who appear to have understood the award scheme well.

The past decade has seen a growing number of pilot studies of pay-for-performance schemes for students at various levels. The picture that emerges from this work is of mostly modest overall effects. As in the randomized evaluations of Angrist, Lang, and Oreopoulos (2009); Angrist and Lavy (2002); and Fryer (2010), the overall achievement impact of the OK program was small, though stronger effects emerge for subgroups, such as second year men. It may be that different incentive schemes, such as larger amounts, lower grade thresholds or incentives for actual behaviour instead of outcomes would lead to larger effects. The Foundations for Success study, for example, found that offering scholarships for course completion and service use had more significant impacts on low-income and low-performing students than did incentivizing specific grade levels alone (MacDonald, Bernstein, & Price, 2009). Overall, however, our interpretation of the research on a variety of different study populations is that merit-based aid is an expensive approach to trying to generate modest effects on retention and performance. We might have been able to generate larger effects with the OK program by increasing the reward amounts at each grade level, but such a program would become very expensive. Likewise, using big rewards at only a few grade targets (as in Project STAR) seems slightly more effective, but students in most sections of the grades distribution gain little motivation from these incentives. One explanation for the muted effectiveness of merit scholarships is that poor-performing students may not know how to effectively study: more time on ineffective study methods will not help improve grades. Another explanation is simply that lack of academic preparation better explains poor PSE performance than lack of effort. Unfortunately, availability of advising and other support services have not helped extensively. Other potential avenues for improving retention and performance may therefore be found at the high school level or through alternative approaches to teaching.

Tables and Figures

Figure 1. Kernel Densities of Predicted Average Grade if Selected or Not Selected for the OK Treatment Program



Note: Before being assigned to the OK treatment program and control group, students were asked what grades they would achieve in each scenario. The figures plot the smoothed kernel densities of these predictions.

Summary Table 1: Fraction of Treated Students Making Program-Related Contact				
	Women		Men	
	First Year	Second Year	First Year	Second Year
Sent at least one e-mail to advisor (fall and winter)	0.790	0.700	0.750	0.560
Any communication to advisor or program officer	0.900	0.890	0.850	0.840
N	100	100	100	100

Notes: Standard deviations are in parentheses.

Summary Table 2: Program Impacts on Full Sample					
Outcome		**Women**		**Men**	
		First Year	**Second Year**	**First Year**	**Second Year**
Number of credits attempted	Control Mean	4.348	4.045	4.298	3.904
	Treatment Effect	-0.011	0.090	0.146	0.229
		(0.083)	(0.113)	(0.082)*	(0.152)
Scholarship earnings (or hypothetical earnings)	Control Mean	1,236.725	1,388.097	1,427.042	1,401.333
	Treatment Effect	-84.446	168.487	4.358	253.315
		(95.008)	(121.041)	(122.493)	(143.231)*
Average course grade	Control Mean	67.855	71.096	69.867	71.453
	Treatment Effect	-0.372	0.484	-0.240	1.167
		(0.961)	(0.933)	(1.210)	(1.091)
Number of courses with grade over 70	Control Mean	4.579	5.215	5.176	5.011
	Treatment Effect	-0.047	0.429	0.122	0.944
		(0.259)	(0.334)	(0.358)	(0.403)**
Did not enrol in subsequent year	Control Mean	0.109	0.094	0.123	0.162
	Treatment Effect	0.018	-0.003	0.035	-0.021
		(0.036)	(0.033)	(0.048)	(0.054)
Average course grade in subsequent fall semester	Control Mean	70.529	73.242	69.991	73.725
	Treatment Effect	1.411	0.345	1.370	-2.152
		(0.908)	(1.165)	(1.495)	(1.453)
	N	449	377	246	199

Notes: The "Control Mean" rows list averages and standard deviations of program earnings within the relevant gender-year subgroup of

the control group. The “Treatment Effect” rows report coefficients from regressions of program earnings on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. For treatment individuals, earnings are equal to actual earnings from the experiment. For control individuals, earnings are equal to what their grades would have earned them if they had been in the treatment group. Full-year courses are included only in the “Winter” and “Full-Year” analyses and are weighted doubly in the earnings calculation. Students are included in the full-year regressions only if they have grades in both semesters. Robust standard errors are in parentheses; standard deviations are in square brackets.
* significant at 10%; ** significant at 5%

Summary Table 3: Program Impacts on Students Demonstrating Better Understanding of Scholarship Calculation

Outcome		Women		Men	
		First Year	Second Year	First Year	Second Year
Number of credits attempted	Control Mean	4.348	4.045	4.298	3.904
	Treatment Effect	-0.011	0.090	0.146	0.229
		(0.083)	(0.113)	(0.082)*	(0.152)
Scholarship earnings (or hypothetical earnings)	Control Mean	1,379.927	1,454.206	1,460.458	1,407.692
	Treatment Effect	-126.753	257.374	50.752	270.677
		(104.168)	(134.807)*	(130.878)	(145.168)*
Average course grade	Control Mean	69.059	71.378	70.093	71.474
	Treatment Effect	-0.674	1.505	0.162	1.144
		(1.033)	(0.966)	(1.269)	(1.117)
Number of courses with grade over 70	Control Mean	5.015	5.360	5.260	5.103
	Treatment Effect	-0.166	0.773	0.230	0.976
		(0.282)	(0.359)**	(0.385)	(0.412)**
	N	359	293	211	162

Notes: The "Control Mean" rows list averages and standard deviations of program earnings within the relevant gender-year subgroup of the control group. The "Treatment Effect" rows report coefficients from regressions of program earnings on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. For treatment individuals, earnings are equal to actual earnings from the experiment. For control individuals, earnings are equal to what their grades would have earned them if they had been in the treatment group. Full-year courses are only included in the "Winter" and "Full-Year" analyses and are weighted doubly in the earnings calculation. Students are included in the full-year regressions only if they have grades in both semesters. Robust standard errors are in parentheses; standard deviations are in square brackets.
* significant at 10%; ** significant at 5%

Table 1. Administrative Characteristics and Parental Characteristics by Gender-Year Categories

	Women				Men				Whole Sample	
	First Year		Second Year		First Year		Second Year			
	Control Mean	Treatment Effect	Control Mean	Treatment Effect	Control Mean	Treatment Effect	Control Mean	Treatment Effect	Control Mean	Treatment Effect
Age	18.23	-0.105	19.19	0.011	18.38	0.014	19.18	0.069	18.67	-0.012
	[0.61]	(0.056)*	[0.51]	(0.056)	[0.81]	(0.104)	[0.46]	(0.070)	[0.76]	(0.036)
N	349	100	277	100	146	100	99	100	871	400
High school	82.82	0.145	82.35	0.302	82.33	-0.344	82.08	-0.387	82.50	-0.024
grade average	[6.56]	(0.238)	[6.19]	(0.217)	[6.44]	(0.310)	[6.73]	(0.338)	[6.44]	(0.134)
N	349	100	277	100	146	100	99	100	871	400
Previous year GPA			2.617	0.023			2.734	-0.047	2.648	-0.005
(2nd years only)			[0.787]	(0.075)			[0.772]	(0.093)	[0.783]	(0.058)
N			276	100			99	100	375	200
1st language	0.404	0.057	0.426	-0.046	0.479	-0.060	0.333	0.097	0.416	0.009
is English	[0.491]	(0.056)	[0.495]	(0.057)	[0.501]	(0.065)	[0.474]	(0.069)	[0.493]	(0.031)
N	349	100	277	100	146	100	99	100	871	400
Mother a high	0.765	0.006	0.827	-0.087	0.747	0.023	0.758	0.042	0.781	-0.009
school graduate	[0.425]	(0.047)	[0.379]	(0.049)*	[0.436]	(0.056)	[0.431]	(0.058)	[0.414]	(0.026)
N	349	100	277	100	146	100	99	100	871	400
Mother a college	0.395	0.065	0.477	-0.016	0.479	0.050	0.424	-0.034	0.439	0.020
graduate	[0.490]	(0.056)	[0.500]	(0.058)	[0.501]	(0.065)	[0.497]	(0.070)	[0.496]	(0.031)
N	349	100	277	100	146	100	99	100	871	400
Father a high	0.742	0.128	0.819	-0.040	0.788	0.002	0.707	0.112	0.770	0.049
school graduate	[0.438]	(0.041)***	[0.385]	(0.048)	[0.410]	(0.053)	[0.457]	(0.059)*	[0.421]	(0.025)**
N	349	100	277	100	146	100	99	100	871	400
Father a college	0.479	0.051	0.581	0.009	0.603	0.047	0.475	0.105	0.532	0.049
graduate	[0.500]	(0.057)	[0.494]	(0.058)	[0.491]	(0.063)	[0.502]	(0.071)	[0.499]	(0.031)
N	349	100	277	100	146	100	99	100	871	400
Mother born in	0.049	0.022	0.087	-0.087	0.034	0.006	0.052	-0.001	0.059	-0.016
Canada	[0.216]	(0.028)	[0.283]	(0.017)***	[0.183]	(0.025)	[0.223]	(0.031)	[0.236]	(0.013)
N	346	99	275	96	145	98	96	100	862	393
Father born in	0.049	0.032	0.066	-0.055	0.048	-0.018	0.042	0.009	0.054	-0.008
Canada	[0.217]	(0.030)	[0.248]	(0.018)***	[0.215]	(0.025)	[0.202]	(0.030)	[0.225]	(0.013)
N	344	98	274	97	145	98	95	100	858	393

Notes: The “Control Mean” columns report averages and standard deviations for variables in the left-most column within the relevant gender-year subgroup of the control group. The “Treatment Effect” columns report coefficients from regressions of each variable in the left-most column on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell). Standard deviations for the control group are in square brackets; robust standard errors are in parentheses. Some respondents did not answer the parents' education questions. They are coded as a separate category (“missing”) and therefore are not coded as high school or college graduates. If data are missing or students failed to respond to other survey questions, these students are excluded from the averages and the number of observations (N).
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2. Student Survey Characteristics and Expectations by Gender-Year Categories

	Women				Men				Whole Sample	
	First Year		Second Year		First Year		Second Year			
	Control Mean	Treatment Dummy	Control Mean	Treatment Dummy	Control Mean	Treatment Dummy	Control Mean	Treatment Dummy	Control Mean	Treatment Dummy
Lives with parents	0.877	0.003	0.881	0.008	0.856	-0.018	0.899	0.001	0.877	-0.001
	[0.329]	(0.037)	[0.325]	(0.037)	[0.352]	(0.047)	[0.303]	(0.043)	[0.328]	(0.020)
N	349	100	277	99	146	99	99	100	871	398
Expects to work	0.442	-0.085	0.465	-0.020	0.336	0.093	0.480	-0.146	0.436	-0.038
during school	[0.497]	(0.055)	[0.500]	(0.059)	[0.474]	(0.064)	[0.502]	(0.070)**	[0.496]	(0.031)
N	346	98	275	97	146	98	98	99	865	392
Expected work	5.794	-0.997	6.402	-0.207	4.402	2.018	6.000	-1.303	5.773	-0.157
hours/week	[7.468]	(0.878)	[8.688]	(1.032)	[6.773]	(1.021)**	[8.188]	(1.196)	[7.865]	(0.512)
N	321	87	256	91	138	85	95	92	810	355
Expected loans (CAD)	4,553	-39.84	5,067	24.70	4,358	284.74	4,967	266.89	4,733	112.08
	[2,394]	(268.63)	[2,072]	(244.79)	[2,267]	(326.00)	[2,066]	(302.19)	[2,250]	(141.23)
N	333	92	270	96	142	92	97	97	842	377
Expected grants (CAD)	1,962	243.84	1,338	201.03	2,147	-388.58	1,431	64.13	1,732	50.73
	[1,995]	(240.59)	[1,168]	(182.21)	[1,990]	(249.15)	[1,204]	(203.13)	[1,717]	(111.27)
N	320	87	257	95	136	93	95	95	808	370
Has very high concern	0.588	-0.017	0.604	0.006	0.521	-0.029	0.500	0.074	0.572	0.004
about funds	[0.493]	(0.056)	[0.490]	(0.057)	[0.501]	(0.064)	[0.503]	(0.071)	[0.495]	(0.031)
N	345	98	275	97	146	98	98	99	864	392
UTSC was first	0.739	-0.090	0.783	0.025	0.740	-0.003	0.859	-0.069	0.767	-0.034
choice school	[0.440]	(0.053)*	[0.413]	(0.047)	[0.440]	(0.057)	[0.350]	(0.054)	[0.423]	(0.027)
N	349	100	277	99	146	99	99	100	871	398
Plans to get a	0.430	0.050	0.408	0.027	0.438	0.037	0.505	0.024	0.433	0.036
graduate degree	[0.496]	(0.057)	[0.492]	(0.058)	[0.498]	(0.065)	[0.503]	(0.071)	[0.496]	(0.031)
N	349	100	277	99	146	99	99	100	871	398

Notes: The “Control Mean” columns report averages and standard deviations for variables in the left-most column, within the relevant gender-year subgroup of the control group. The “Treatment Effect” columns report coefficients from regressions of each variable in the left-most column on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell). Standard deviations for the control group are in square brackets; robust standard errors are in parentheses. If data are missing or students failed to respond to survey questions, these students are excluded from the averages and the number of observations (N).

* significant at 10%; ** significant at 5%

Table 3. Questions Related to the OK Program by Gender-Year Categories

	Women				Men				Whole Sample	
	First Year		Second Year		Second Year		First Year			
	Control Mean	Treatment Effect	Control Mean	Treatment Effect	Control Mean	Treatment Effect	Control Mean	Treatment Effect	Control Mean	Treatment Effect
Answered 1st test question on	0.797	0.062	0.845	0.024	0.918	-0.118	0.859	0.061	0.839	0.010
scholarship calculation correctly	[0.403]	(0.040)	[0.363]	(0.040)	[0.276]	(0.046)**	[0.350]	(0.043)	[0.367]	(0.021)
N	349	100	277	100	146	100	99	100	871	400
Answered 2nd test question on	0.616	0.022	0.690	-0.010	0.719	-0.080	0.697	0.002	0.666	-0.014
scholarship calculation correctly	[0.487]	(0.053)	[0.464]	(0.054)	[0.451]	(0.061)	[0.462]	(0.065)	[0.472]	(0.029)
N	349	100	277	100	146	100	99	100	871	400
Expected higher GPA	0.364	0.013	0.393	0.013	0.393	-0.048	0.315	0.051	0.373	0.006
if selected for treatment	[0.482]	(0.057)	[0.489]	(0.059)	[0.490]	(0.064)	[0.467]	(0.068)	[0.484]	(0.031)
N	327	93	262	94	140	92	92	98	821	377
Difference between expected	0.200	-0.014	0.192	0.025	0.220	-0.057	0.145	-0.011	0.195	-0.012
treatment and control GPAs	[0.345]	(0.040)	[0.322]	(0.045)	[0.356]	(0.048)	[0.272]	(0.038)	[0.332]	(0.022)
N	327	93	262	94	140	92	92	98	821	377
Days between invitation to	6.476	-1.130	3.888	1.130	6.911	-0.031	4.697	-0.502	5.524	-0.121
OK program and sign-up	[7.386]	(0.783)	[5.699]	(0.760)	[7.542]	(0.938)	[6.198]	(0.812)	[6.896]	(0.414)
N	349	100	277	100	146	100	99	100	871	400

Notes: The “Control Mean” columns report averages and standard deviations for variables in the left-most column, within the relevant gender-year subgroup of the control group. The “Treatment Effect” columns report coefficients from regressions of each variable in the left-most column on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell). Standard deviations for the control group are in square brackets; robust standard errors are in parentheses. If data are missing or students failed to respond to survey questions, these students are excluded from the averages and the number of observations (N).

** significant at 5%

Table 4. Analysis of Pre-Randomization Sign-Up for the OK Program				
	Means		Differences in Means	
	Signed Up	Didn't Sign Up	No Controls	Strata Controls
Female	0.650	0.580	0.070	
	(0.477)	(0.494)	(0.023)***	
First year student	0.547	0.528	0.019	
	(0.498)	(0.500)	(0.024)	
High school grade average	82.470	79.165	3.305***	0.273
	(6.408)	(5.882)	(0.289)	(0.108)**
First language is English	0.418	0.444	-0.026	-0.018
	(0.493)	(0.497)	(0.024)	(0.025)
Attempted credits (fall)	2.162	1.928	0.234***	0.188
	(0.429)	(0.518)	(0.023)	(0.024)***
Attempted credits (spring)	2.048	1.768	0.280***	0.236
	(0.620)	(0.734)	(0.033)	(0.033)***
Attempted credits (total)	4.210	3.696	0.514***	0.424
	(0.882)	(1.081)	(0.048)	(0.048)***
Observations	1,271	674	1,945	1,945

Notes: The “Means” columns list averages and standard deviations for variables in the left-most column, for all students at UTSC who were invited to participate in the OK program, split by whether or not they actually signed up to participate. The “Differences in Means” columns present coefficients from regressions of each variable in the left-most column on a variable equal to one if an invited student signed up to participate in the experiment. The “Strata Controls” column also controls for each interaction of gender, year in school and high school grade quartile within the gender/year cell. Numbers in parentheses are standard deviations for the “Means” column, and robust standard errors for “Difference in Means” columns. For all columns, the sample excludes students who signed up on their own without receiving an invitation and students who were disqualified from the study because they did not apply for financial aid or signed up for fewer than 1.5 credits in fall 2008. Randomization into treatment and control groups occurs after this sign-up stage.

** significant at 5%; *** significant at 1%

Table 5. Fraction of Treated Students Making Program-Related Contact by Gender and Year

	Women			Men			All First Years	All Second Years	Whole Sample
	First Year	Second Year	All	First Year	Second Year	All			
Sent at least one e-mail to advisor (fall)	0.450 (0.500)	0.390 (0.490)	0.420 (0.495)	0.410 (0.494)	0.270 (0.446)	0.340 (0.475)	0.430 (0.496)	0.330 (0.471)	0.380 (0.486)
Sent at least one e-mail to advisor (winter)	0.520 (0.502)	0.440 (0.499)	0.480 (0.501)	0.660 (0.476)	0.380 (0.488)	0.520 (0.501)	0.590 (0.493)	0.410 (0.493)	0.500 (0.501)
Sent at least one e-mail to advisor (fall and winter)	0.790 (0.409)	0.700 (0.461)	0.745 (0.437)	0.750 (0.435)	0.560 (0.499)	0.655 (0.477)	0.770 (0.422)	0.630 (0.484)	0.700 (0.459)
Record of response to treatment notification	0.050 (0.219)	0.190 (0.394)	0.120 (0.326)	0.100 (0.302)	0.130 (0.338)	0.115 (0.320)	0.075 (0.264)	0.160 (0.368)	0.117 (0.322)
Checked scholarship earnings online	0.760 (0.429)	0.780 (0.416)	0.770 (0.422)	0.650 (0.479)	0.710 (0.456)	0.680 (0.468)	0.705 (0.457)	0.745 (0.437)	0.725 (0.447)
Sent any other e-mail to program website	0.270 (0.446)	0.320 (0.469)	0.295 (0.457)	0.250 (0.435)	0.300 (0.461)	0.275 (0.448)	0.260 (0.440)	0.310 (0.464)	0.285 (0.452)
Made contact in any of the ways above	0.900 (0.302)	0.890 (0.314)	0.895 (0.307)	0.850 (0.359)	0.840 (0.368)	0.845 (0.363)	0.875 (0.332)	0.865 (0.343)	0.870 (0.337)
Observations	100	100	200	100	100	200	200	200	400

Notes: Numbers in the table are fractions of students who made the corresponding form of program-related contact at least once. The sample consists of all treated students, even those who didn't enroll in any courses. However, records of responses to the treatment notification email are incomplete. Thus, these fractions are too low. Standard deviations are in parentheses.

Table 6. Treatment Effects on Credits Attempted

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	2.179	2.144	2.163	2.199	2.051	2.139	2.185	2.119	2.156
		[0.407]	[0.482]	[0.442]	[0.345]	[0.565]	[0.452]	[0.389]	[0.506]	[0.444]
	Treatment Effect	-0.039	0.030	-0.001	0.011	0.123	0.061	-0.010	0.057	0.021
		(0.046)	(0.050)	(0.034)	(0.042)	(0.073)*	(0.041)	(0.031)	(0.040)	(0.025)
	N	449	377	826	246	199	445	695	576	1,271
Winter	Control Mean	2.169	1.902	2.051	2.099	1.854	2.000	2.148	1.889	2.036
		[0.499]	[0.780]	[0.652]	[0.497]	[0.708]	[0.602]	[0.499]	[0.761]	[0.638]
	Treatment Effect	0.028	0.060	0.038	0.135	0.106	0.095	0.065	0.063	0.061
		(0.049)	(0.086)	(0.047)	(0.059)**	(0.104)	(0.055)*	(0.036)*	(0.063)	(0.035)*
	N	449	377	826	246	199	445	695	576	1,271
Full Year	Control Mean	4.348	4.045	4.214	4.298	3.904	4.139	4.333	4.008	4.193
		[0.792]	[1.029]	[0.916]	[0.684]	[1.061]	[0.876]	[0.762]	[1.038]	[0.905]
	Treatment Effect	-0.011	0.090	0.037	0.146	0.229	0.156	0.055	0.120	0.082
		(0.083)	(0.113)	(0.068)	(0.082)*	(0.152)	(0.080)*	(0.057)	(0.087)	(0.051)
	N	449	377	826	246	199	445	695	576	1,271

Notes: The “Control Mean” rows list averages and standard deviations of credits attempted within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of credits attempted on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. Credits are taken from administrative data on number of credits attempted each year and semester. Credit for full-year courses is split half and half between fall and winter. If students left the university, they are coded as having attempted zero credits. Robust standard errors are in parentheses; standard deviations are in square brackets.

* significant at 10%; ** significant at 5%

Table 7. Treatment Effects on (Real and Hypothetical) Program Earnings

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	644.522	694.691	666.774	769.589	744.000	759.502	681.711	707.351	692.729
		[657.021]	[588.855]	[627.701]	[669.865]	[642.009]	[657.794]	[662.657]	[602.447]	[637.245]
	Treatment Effect	-21.466	101.831	39.779	32.466	47.684	10.465	-7.619	73.565	27.281
		(53.105)	(61.275)*	(39.869)	(69.985)	(73.057)	(51.461)	(42.075)	(45.905)	(31.128)
	N	444	374	818	246	195	441	690	569	1,259
Winter	Control Mean	589.123	710.526	640.034	644.225	655.217	648.547	605.289	695.516	642.454
		[608.184]	[598.200]	[606.477]	[600.141]	[682.821]	[632.520]	[605.735]	[621.753]	[613.619]
	Treatment Effect	-59.455	27.241	-18.819	-21.409	170.960	34.529	-53.866	80.651	4.321
		(49.188)	(66.306)	(39.519)	(60.149)	(80.223)**	(49.502)	(37.655)	(51.034)	(30.827)
	N	441	340	781	242	183	425	683	523	1,206
Full Year	Control Mean	1,236.725	1,388.097	1,300.204	1,427.042	1,401.333	1,417.069	1,292.562	1,391.632	1,333.228
		[1,221.834]	[1,090.416]	[1,169.957]	[1,228.497]	[1,265.280]	[1,240.230]	[1,225.593]	[1,137.811]	[1,190.636]
	Treatment Effect	-84.446	168.487	33.302	4.358	253.315	51.633	-67.308	185.272	40.315
		(95.008)	(121.041)	(74.013)	(122.493)	(143.231)*	(95.533)	(74.527)	(91.247)**	(58.186)
	N	441	339	780	242	181	423	683	520	1,203

Notes: The “Control Mean” rows list averages and standard deviations of program earnings within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of program earnings on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. For treatment individuals, earnings are equal to actual earnings from the experiment. For control individuals, earnings are equal to what their grades would have earned them if they had been in the treatment group. Full-year courses are included only in the “Winter” and “Full-Year” analyses and are weighted doubly in the earnings calculation. Students are included in the full-year regressions only if they have grades in both semesters. If we restrict the Fall and Winter samples to be the same as the full-year sample, then the treatment effects for the full year are the sum of the fall and winter treatment effects (this is also true in later tables). Robust standard errors are in parentheses; standard deviations are in square brackets.

* significant at 10%; ** significant at 5%

Table 8. Treatment Effects on Average Grades

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	68.133	71.003	69.406	70.657	72.437	71.359	68.883	71.371	69.952
		[11.593]	[8.400]	[10.391]	[10.915]	[8.391]	[10.016]	[11.443]	[8.410]	[10.319]
	Treatment Effect	0.379	0.444	0.452	0.446	-0.526	-0.503	0.212	0.088	0.068
		(0.946)	(0.950)	(0.661)	(1.178)	(1.063)	(0.828)	(0.742)	(0.693)	(0.514)
	N	444	374	818	246	195	441	690	569	1,259
Winter	Control Mean	67.358	71.150	68.948	68.834	69.970	69.281	67.791	70.830	69.043
		[11.261]	[9.015]	[10.538]	[11.183]	[10.575]	[10.939]	[11.247]	[9.462]	[10.648]
	Treatment Effect	-0.876	-0.089	-0.489	-0.976	2.551	0.104	-0.994	0.768	-0.238
		(1.158)	(1.127)	(0.801)	(1.549)	(1.424)*	(1.029)	(0.900)	(0.904)	(0.634)
	N	441	340	781	242	183	425	683	523	1,206
Full Year	Control Mean	67.855	71.096	69.214	69.867	71.453	70.482	68.445	71.191	69.572
		[10.674]	[7.772]	[9.691]	[10.323]	[8.593]	[9.700]	[10.602]	[7.988]	[9.704]
	Treatment Effect	-0.372	0.484	0.064	-0.240	1.167	-0.155	-0.482	0.641	-0.037
		(0.961)	(0.933)	(0.662)	(1.210)	(1.091)	(0.841)	(0.746)	(0.720)	(0.522)
	N	441	339	780	242	181	423	683	520	1,203

Notes: The "Control Mean" rows list averages and standard deviations of average grades within the relevant gender-year subgroup of the control group. The "Treatment Effect" rows report coefficients from regressions of average grades on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. Average grades are on a 100-point scale. Full-year courses are included only in the "Winter" and "Full-Year" analyses and are weighted doubly in the average grade calculation. Students are included in the full-year regressions only if they have grades in both semesters. Robust standard errors are in parentheses; standard deviations are in square brackets.

* significant at 10%

Table 9. Treatment Effects on Number of Courses with Grade over 69 Per Cent

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	2.354	2.618	2.471	2.699	2.705	2.701	2.456	2.641	2.535
		[1.773]	[1.624]	[1.712]	[1.687]	[1.550]	[1.631]	[1.753]	[1.604]	[1.692]
	Treatment Effect	0.091	0.202	0.156	0.078	0.208	0.071	0.065	0.189	0.116
		(0.150)	(0.174)	(0.113)	(0.195)	(0.212)	(0.143)	(0.117)	(0.131)	(0.087)
	N	444	374	818	246	195	441	690	569	1,259
Winter	Control Mean	2.213	2.611	2.380	2.444	2.315	2.393	2.281	2.531	2.384
		[1.776]	[1.539]	[1.691]	[1.657]	[1.809]	[1.716]	[1.743]	[1.619]	[1.697]
	Treatment Effect	-0.127	0.128	-0.006	0.061	0.598	0.217	-0.067	0.295	0.086
		(0.143)	(0.184)	(0.113)	(0.196)	(0.245)**	(0.151)	(0.115)	(0.145)**	(0.090)
	N	441	340	781	242	183	425	683	523	1,206
Full Year	Control Mean	4.579	5.215	4.846	5.176	5.011	5.112	4.754	5.160	4.921
		[3.345]	[2.841]	[3.157]	[3.172]	[2.955]	[3.084]	[3.303]	[2.869]	[3.137]
	Treatment Effect	-0.047	0.429	0.184	0.122	0.944	0.329	-0.018	0.582	0.235
		(0.259)	(0.334)	(0.205)	(0.358)	(0.403)**	(0.269)	(0.209)	(0.251)**	(0.161)
	N	441	339	780	242	181	423	683	520	1,203

Notes: The “Control Mean” rows list averages and standard deviations of the number of courses with grades over 69 per cent within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of the number of courses with grades over 69 per cent on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. The dependent variable is the number of courses for which the student earned a grade above 69 per cent. Full-year courses are included only in the “Winter” and “Full-Year” analyses and contribute two to the course count. Students are included in the full year regressions only if they have grades in both semesters. Robust standard errors are in parentheses; standard deviations are in square brackets.

** significant at 5%

Table 10. Treatment Effects on Total Grade Percentage Points over 70 Per Cent

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	20.458	21.644	20.984	24.986	23.674	24.469	21.804	22.165	21.959
		[25.247]	[22.503]	[24.057]	[26.341]	[25.595]	[26.004]	[25.634]	[23.317]	[24.652]
	Treatment Effect	-1.529	4.080	1.210	1.233	1.344	0.168	-0.707	2.732	0.782
		(2.130)	(2.359)*	(1.562)	(2.702)	(2.789)	(1.985)	(1.659)	(1.771)	(1.216)
	N	444	374	818	246	195	441	690	569	1,259
Winter	Control Mean	18.389	22.470	20.100	19.993	21.185	20.462	18.860	22.121	20.203
		[22.636]	[23.459]	[23.053]	[22.858]	[26.530]	[24.319]	[22.689]	[24.297]	[23.405]
	Treatment Effect	-2.337	0.722	-0.912	-1.375	5.557	0.641	-2.356	2.556	-0.216
		(1.889)	(2.584)	(1.528)	(2.204)	(2.981)*	(1.846)	(1.419)*	(1.964)	(1.176)
	N	441	340	781	242	183	425	683	523	1,206
Full Year	Control Mean	38.942	43.332	40.783	45.472	45.011	45.293	40.857	43.780	42.057
		[46.153]	[42.065]	[44.500]	[47.410]	[50.360]	[48.468]	[46.571]	[44.360]	[45.672]
	Treatment Effect	-3.990	6.278	0.745	-0.392	7.947	0.935	-3.275	6.353	0.839
		(3.749)	(4.663)	(2.876)	(4.607)	(5.455)	(3.633)	(2.880)	(3.519)*	(2.249)
	N	441	339	780	242	181	423	683	520	1,203

Notes: The “Control Mean” rows list averages and standard deviations of the total grade percentage points above 70 per cent across all classes within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of the total percentage points over 70 per cent on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. The dependent variable is the sum of the grade percentage points above 70 per cent across all classes taken by the student. Full-year courses are included only in the “Winter” and “Full-Year” analyses and their grade points over 70 per cent are weighted doubly in the sum. Students are included in the full-year regressions only if they have grades in both semesters. Robust standard errors are in parentheses; standard deviations are in square brackets.
* significant at 10%

Table 11. Treatment Effects on (Real and Hypothetical) Program Earnings for Those Who Answered the First Earnings Test Question Correctly

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	727.246	747.296	736.424	788.507	758.293	777.037	747.268	750.159	748.524
		[679.279]	[603.175]	[645.010]	[681.799]	[624.213]	[659.219]	[679.878]	[607.739]	[649.086]
	Treatment Effect	-62.137	130.900	35.386	75.654	59.377	53.793	1.599	98.851	45.033
		(56.078)	(67.985)*	(43.854)	(76.164)	(75.066)	(55.209)	(45.934)	(49.628)**	(34.007)
	N	361	319	680	214	174	388	575	493	1,068
Winter	Control Mean	650.584	738.037	688.934	658.626	655.443	657.429	653.185	715.768	679.456
		[625.240]	[624.021]	[625.574]	[606.773]	[679.705]	[633.586]	[618.587]	[639.382]	[627.707]
	Treatment Effect	-65.569	71.864	-1.093	-8.818	173.232	60.636	-45.414	116.863	26.303
		(55.293)	(73.451)	(44.434)	(63.084)	(81.637)**	(51.825)	(41.624)	(54.702)**	(33.977)
	N	359	294	653	211	163	374	570	457	1,027
Full Year	Control Mean	1379.927	1454.206	1412.500	1460.458	1407.692	1440.766	1405.975	1441.781	1420.976
		[1261.596]	[1128.837]	[1204.540]	[1247.043]	[1235.699]	[1240.108]	[1255.930]	[1156.365]	[1214.490]
	Treatment Effect	-126.753	257.374	54.452	50.752	270.677	120.317	-49.281	258.730	85.512
		(104.168)	(134.807)*	(82.925)	(130.878)	(145.168)*	(100.466)	(81.863)	(98.206)***	(63.907)
	N	359	293	652	211	162	373	570	455	1,025

Notes: The “Control Mean” rows list averages and standard deviations of program earnings within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of program earnings on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, and parental education. For treatment individuals, earnings are equal to actual earnings from the experiment. For control individuals, earnings are equal to what their grades would have earned them if they had been in the treatment group. Full-year courses are included only in the “Winter” and “Full Year” analyses and are weighted doubly in the earnings calculation. Students are included in the full-year regressions only if they have grades in both semesters. Robust standard errors are in parentheses; standard deviations are in square brackets.

* significant at 10%; ** significant at 5%

Table 12. Treatment Effects on Average Grades for Those Who Answered the First Earnings Test Question Correctly

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	69.613	71.710	70.573	70.975	72.744	71.646	70.058	71.979	70.893
		[11.195]	[8.361]	[10.043]	[10.901]	[7.990]	[9.915]	[11.105]	[8.266]	[10.010]
	Treatment Effect	-0.030	0.701	0.416	1.223	-0.535	0.216	0.531	0.276	0.366
		(0.968)	(1.034)	(0.705)	(1.243)	(1.072)	(0.845)	(0.767)	(0.734)	(0.537)
	N	361	319	680	214	174	388	575	493	1,068
Winter	Control Mean	68.363	71.351	69.673	68.860	69.739	69.191	68.524	70.916	69.528
		[11.272]	[9.383]	[10.580]	[11.266]	[10.813]	[11.080]	[11.259]	[9.797]	[10.727]
	Treatment Effect	-1.125	1.191	-0.004	-0.652	2.659	0.636	-0.911	1.675	0.265
		(1.289)	(1.122)	(0.864)	(1.613)	(1.475)*	(1.106)	(0.998)	(0.914)*	(0.686)
	N	359	294	653	211	163	374	570	457	1,027
Full Year	Control Mean	69.059	71.378	70.076	70.093	71.474	70.608	69.393	71.404	70.236
		[10.509]	[8.067]	[9.576]	[10.364]	[8.521]	[9.719]	[10.461]	[8.176]	[9.615]
	Treatment Effect	-0.674	1.505	0.391	0.162	1.144	0.394	-0.287	1.305	0.408
		(1.033)	(0.966)	(0.710)	(1.269)	(1.117)	(0.872)	(0.797)	(0.742)*	(0.553)
	N	359	293	652	211	162	373	570	455	1,025

Notes: The “Control Mean” rows list averages and standard deviations of average grades within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of average grades on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, and parental education. Average grades are on a 100-point scale. Full-year courses are included only in the “Winter” and “Full-Year” analyses and are weighted doubly in the average grade calculation. Students are included in the full-year regressions only if they have grades in both semesters. Robust standard errors are in parentheses; standard deviations are in square brackets.

* significant at 10%

Table 13. Treatment Effects on Number of Courses with Grade over 69 Per Cent for Those Who Answered the First Earnings Test Question Correctly

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	2.598	2.764	2.674	2.746	2.780	2.759	2.646	2.768	2.699
		[1.765]	[1.608]	[1.695]	[1.676]	[1.507]	[1.610]	[1.735]	[1.580]	[1.670]
	Treatment Effect	-0.030	0.314	0.155	0.174	0.241	0.171	0.072	0.266	0.161
		(0.158)	(0.188)*	(0.123)	(0.211)	(0.216)	(0.154)	(0.128)	(0.139)*	(0.094)*
	N	361	319	680	214	174	388	575	493	1,068
Winter	Control Mean	2.409	2.659	2.518	2.481	2.354	2.433	2.432	2.577	2.493
		[1.779]	[1.569]	[1.693]	[1.666]	[1.812]	[1.719]	[1.742]	[1.640]	[1.700]
	Treatment Effect	-0.138	0.300	0.074	0.093	0.609	0.286	-0.050	0.418	0.160
		(0.161)	(0.198)	(0.126)	(0.206)	(0.253)**	(0.161)*	(0.127)	(0.154)***	(0.099)
	N	359	294	653	211	163	374	570	457	1,027
Full Year	Control Mean	5.015	5.360	5.166	5.260	5.103	5.201	5.094	5.291	5.176
		[3.338]	[2.885]	[3.149]	[3.173]	[2.913]	[3.072]	[3.284]	[2.890]	[3.124]
	Treatment Effect	-0.166	0.773	0.292	0.230	0.976	0.492	0.008	0.816	0.371
		(0.282)	(0.359)**	(0.225)	(0.385)	(0.412)**	(0.287)*	(0.230)	(0.266)***	(0.176)**
	N	359	293	652	211	162	373	570	455	1,025

Notes: The “Control Mean” rows list averages and standard deviations of average grades within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of average grades on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, and parental education. The dependent variable is the number of courses for which the student earned a grade above 69 per cent. Full-year courses are included only in the “Winter” and “Full-Year” analyses and contribute two to the course count. Students are included in the full-year regressions only if they have grades in both semesters. Robust standard errors are in parentheses; standard deviations are in square brackets.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 14. Treatment Effects on Total Grade Percentage Points over 70 Per Cent for Those Who Answered the First Earnings Test Question Correctly

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	23.373	23.545	23.452	25.694	24.012	25.056	24.132	23.667	23.930
		[26.482]	[23.365]	[25.079]	[26.998]	[24.860]	[26.161]	[26.641]	[23.724]	[25.399]
	Treatment Effect	-2.955	4.977	0.992	2.913	1.765	1.835	-0.278	3.613	1.444
		(2.276)	(2.645)*	(1.734)	(2.946)	(2.890)	(2.133)	(1.818)	(1.930)*	(1.335)
	N	361	319	680	214	174	388	575	493	1,068
Winter	Control Mean	20.485	23.607	21.855	20.527	21.000	20.705	20.499	22.904	21.509
		[23.561]	[24.589]	[24.042]	[23.175]	[26.384]	[24.371]	[23.408]	[25.066]	[24.130]
	Treatment Effect	-2.589	2.093	-0.426	-0.905	5.616	1.602	-2.019	3.753	0.517
		(2.130)	(2.884)	(1.725)	(2.317)	(3.022)*	(1.926)	(1.571)	(2.112)*	(1.300)
	N	359	294	653	211	163	374	570	457	1,027
Full Year	Control Mean	43.923	45.911	44.795	46.725	44.872	46.033	44.830	45.634	45.166
		[48.300]	[43.807]	[46.347]	[48.375]	[49.113]	[48.542]	[48.282]	[45.201]	[46.984]
	Treatment Effect	-5.507	9.001	1.264	1.390	8.655	3.555	-2.506	8.858	2.422
		(4.140)	(5.253)*	(3.246)	(4.926)	(5.531)	(3.811)	(3.173)	(3.814)**	(2.481)
	N	359	293	652	211	162	373	570	455	1,025

Notes: The “Control Mean” rows list averages and standard deviations of average grades within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of average grades on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, and parental education. The dependent variable is sum of the grade percentage points above 70 per cent across all classes taken by the student. Full-year courses are included only in the “Winter” and “Full-Year” analyses and their grade points over 70 per cent are weighted doubly in the sum. Students are included in the full-year regressions only if they have grades in both semesters. Robust standard errors are in parentheses; standard deviations are in square brackets.

* significant at 10%; ** significant at 5%

Table 15. Treatment Effects on Credits Earned

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	2.009	2.077	2.039	2.058	2.005	2.037	2.023	2.058	2.038
		[0.567]	[0.523]	[0.549]	[0.507]	[0.569]	[0.532]	[0.550]	[0.536]	[0.544]
	Treatment Effect	-0.030	0.015	-0.008	0.042	0.105	0.061	0.002	0.042	0.017
		(0.058)	(0.055)	(0.040)	(0.059)	(0.078)	(0.049)	(0.042)	(0.044)	(0.031)
	N	449	377	826	246	199	445	695	576	1,271
Winter	Control Mean	2.026	1.835	1.941	1.932	1.753	1.859	1.998	1.813	1.918
		[0.608]	[0.803]	[0.707]	[0.648]	[0.744]	[0.692]	[0.621]	[0.787]	[0.703]
	Treatment Effect	0.005	0.068	0.030	0.152	0.143	0.105	0.061	0.073	0.061
		(0.063)	(0.089)	(0.052)	(0.079)*	(0.108)	(0.063)*	(0.048)	(0.066)	(0.040)
	N	449	377	826	246	199	445	695	576	1,271
Full Year	Control Mean	4.034	3.912	3.980	3.990	3.758	3.896	4.021	3.871	3.956
		[1.076]	[1.111]	[1.093]	[1.006]	[1.098]	[1.048]	[1.055]	[1.108]	[1.080]
	Treatment Effect	-0.025	0.083	0.021	0.195	0.248	0.166	0.063	0.115	0.078
		(0.107)	(0.123)	(0.080)	(0.121)	(0.161)	(0.098)*	(0.079)	(0.095)	(0.061)
	N	449	377	826	246	199	445	695	576	1,271

Notes: The “Control Mean” rows list averages and standard deviations of credits earned within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of credits earned on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. A credit counts as earned if the student received a 50 per cent grade or better in the course. The credits earned variable is constructed from course level data and it includes only courses eligible for program earnings (i.e., pass/fail classes are excluded). Credit for full-year courses is split half and half between fall and winter. If students left the university, they are coded as having earned zero credits. Robust standard errors are in parentheses; standard deviations are in square brackets.

* significant at 10%

Table 16. Treatment Effects on Being in Good Standing or on Academic Probation

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Good Standing	Control Mean	0.771	0.856	0.808	0.815	0.889	0.845	0.784	0.864	0.819
		[0.421]	[0.352]	[0.394]	[0.390]	[0.316]	[0.363]	[0.412]	[0.343]	[0.386]
	Treatment Effect	0.036	-0.019	0.008	0.052	-0.016	0.010	0.043	-0.024	0.009
		(0.041)	(0.042)	(0.029)	(0.047)	(0.049)	(0.033)	(0.030)	(0.031)	(0.022)
	N	449	377	826	246	199	445	695	576	1,271
Probation	Control Mean	0.198	0.007	0.113	0.130	0.010	0.082	0.178	0.008	0.104
		[0.399]	[0.085]	[0.317]	[0.338]	[0.101]	[0.274]	[0.383]	[0.089]	[0.306]
	Treatment Effect	-0.018	0.012	-0.005	0.002	0.011	0.014	-0.010	0.016	0.004
		(0.039)	(0.013)	(0.021)	(0.044)	(0.020)	(0.025)	(0.029)	(0.012)	(0.016)
	N	449	377	826	246	199	445	695	576	1,271

Notes: The “Control Mean” rows list averages and standard deviations of credits earned within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of credits earned on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. The dependent variables are dummies equal to one if the student is in good standing/on probation at the end of the 2008/2009 school year. These designations are based on cumulative GPA cutoffs and require that the student has completed 2.5 credits. Robust standard errors are in parentheses; standard deviations are in square brackets.

Table 17. Treatment Effects on Dropping Out of School

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Fall	Control Mean	0.009	0.007	0.008	0.000	0.040	0.016	0.006	0.016	0.010
		[0.092]	[0.085]	[0.089]	[0.000]	[0.198]	[0.127]	[0.078]	[0.125]	[0.101]
	Treatment Effect	0.004	0.006	0.004	0.000	-0.047	-0.021	0.001	-0.014	-0.006
		(0.012)	(0.012)	(0.008)	(.)	(0.023)**	(0.010)**	(0.007)	(0.011)	(0.006)
	N	449	377	826	246	199	445	695	576	1,271
Winter	Control Mean	0.020	0.108	0.059	0.021	0.071	0.041	0.020	0.098	0.054
		[0.140]	[0.311]	[0.236]	[0.142]	[0.258]	[0.198]	[0.141]	[0.298]	[0.226]
	Treatment Effect	-0.009	-0.040	-0.020	-0.025	-0.001	-0.011	-0.016	-0.021	-0.017
		(0.014)	(0.033)	(0.017)	(0.015)*	(0.038)	(0.018)	(0.009)*	(0.023)	(0.012)
	N	449	377	826	246	199	445	695	576	1,271
Full Year	Control Mean	0.020	0.108	0.059	0.021	0.091	0.049	0.020	0.104	0.056
		[0.140]	[0.311]	[0.236]	[0.142]	[0.289]	[0.216]	[0.141]	[0.305]	[0.231]
	Treatment Effect	-0.009	-0.029	-0.015	-0.025	-0.021	-0.022	-0.016	-0.022	-0.018
		(0.014)	(0.034)	(0.017)	(0.015)*	(0.040)	(0.019)	(0.009)*	(0.024)	(0.013)
	N	449	377	826	246	199	445	695	576	1,271

Notes: The “Control Mean” rows list averages and standard deviations of average grades within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of average grades on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. The dependent variable is equal to one if the student dropped out of school during the relevant time period (at least one of two semesters in the year-long regressions) and equal to zero otherwise. We define dropping out as attempting zero credits. Robust standard errors are in parentheses; standard deviations are in square brackets.

* significant at 10%; ** significant at 5%

Table 18. Treatment Effects on Grade Outcomes in Fall 2009

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Program Earnings	Control Mean	705.147	780.735	738.696	650.476	761.220	694.135	689.238	775.841	726.500
		[643.549]	[588.436]	[620.281]	[619.629]	[610.875]	[617.101]	[636.453]	[593.256]	[619.325]
	Treatment Effect	5.055	62.516	33.055	71.151	22.837	34.356	19.145	55.453	31.243
		(57.756)	(68.877)	(44.122)	(74.153)	(77.647)	(53.125)	(45.091)	(51.541)	(33.888)
	N	395	334	729	209	165	374	604	499	1,103
Average Grade	Control Mean	70.529	73.242	71.733	69.991	73.725	71.463	70.372	73.363	71.659
		[10.251]	[9.072]	[9.830]	[12.333]	[8.552]	[11.127]	[10.886]	[8.934]	[10.194]
	Treatment Effect	1.411	0.345	0.824	1.370	-2.152	-0.434	1.333	-0.602	0.291
		(0.908)	(1.165)	(0.732)	(1.495)	(1.453)	(1.062)	(0.802)*	(0.908)	(0.601)
	N	395	334	729	209	165	374	604	499	1,103
GPA	Control Mean	2.582	2.838	2.696	2.571	2.860	2.685	2.579	2.843	2.693
		[0.906]	[0.763]	[0.854]	[0.966]	[0.710]	[0.884]	[0.922]	[0.749]	[0.862]
	Treatment Effect	0.146	0.019	0.080	0.084	-0.144	-0.036	0.118	-0.040	0.033
		(0.078)*	(0.095)	(0.061)	(0.127)	(0.121)	(0.088)	(0.068)*	(0.073)	(0.050)
	N	395	334	729	209	165	374	604	499	1,103

Notes: The “Control Mean” rows list averages and standard deviations of the variables in the left-most column within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of the left-most column variables on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. Earnings are equal to what their grades would have earned them in the treatment group during fall 2009. Full-year courses are excluded, since they do not give grades in the fall. Robust standard errors are in parentheses; standard deviations are in square brackets.

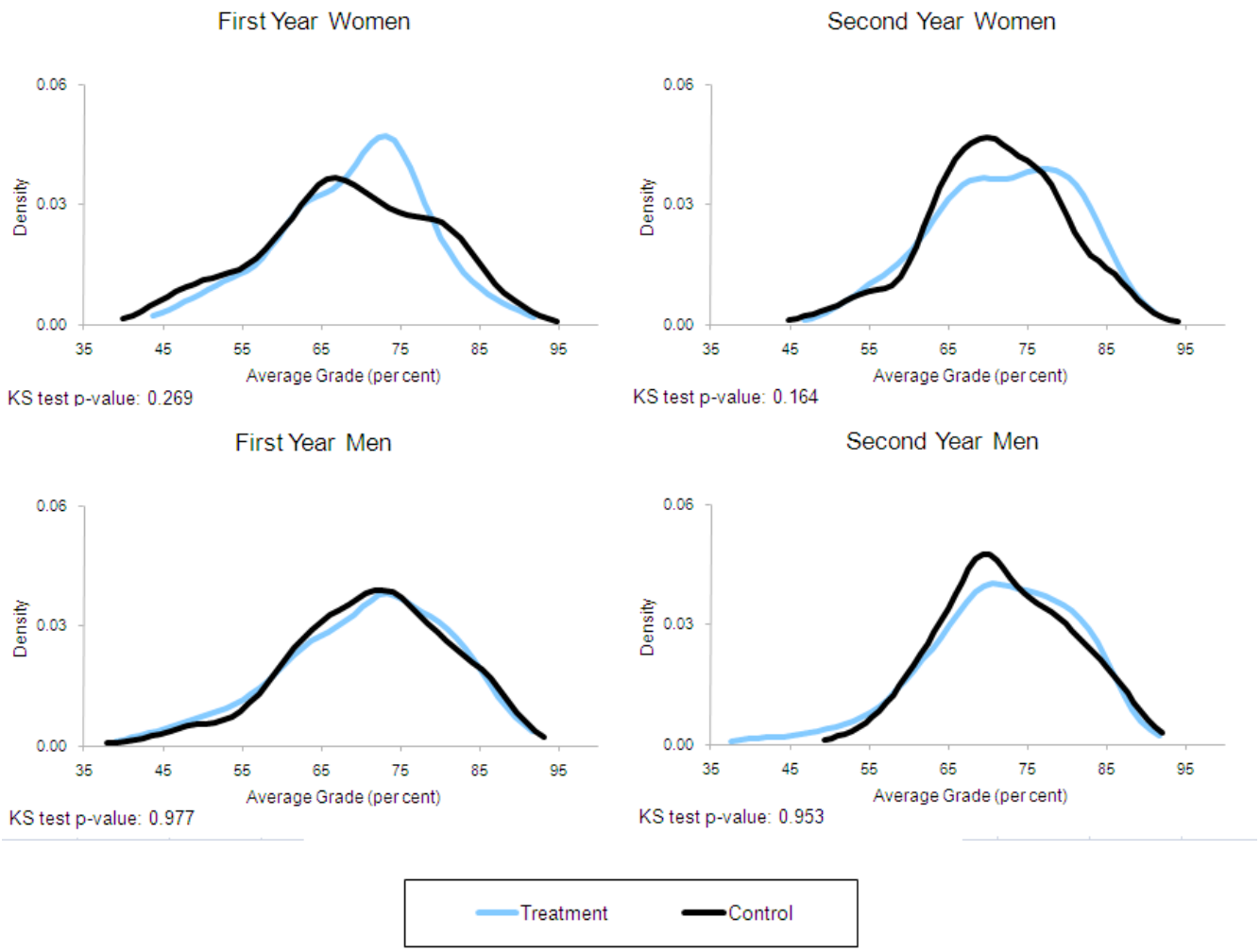
* significant at 10%

Table 19. Treatment Effects on Credit Outcomes in Fall 2009

		Women			Men			All First Years	All Second Years	Whole Sample
		First Year	Second Year	All	First Year	Second Year	All			
Credits Attempted	Control Mean	1.908	1.912	1.910	1.740	1.674	1.713	1.859	1.849	1.854
		[0.779]	[0.760]	[0.770]	[0.768]	[0.884]	[0.816]	[0.779]	[0.800]	[0.788]
	Treatment Effect	-0.112	0.071	-0.012	0.043	0.167	0.084	-0.050	0.100	0.024
		(0.083)	(0.089)	(0.061)	(0.111)	(0.129)	(0.083)	(0.066)	(0.073)	(0.049)
	N	449	373	822	246	199	445	695	572	1,267
Credits Earned	Control Mean	1.723	1.690	1.708	1.545	1.510	1.531	1.671	1.642	1.658
		[0.831]	[0.835]	[0.832]	[0.839]	[0.906]	[0.865]	[0.837]	[0.857]	[0.845]
	Treatment Effect	-0.042	0.032	0.008	0.125	0.161	0.121	0.029	0.075	0.053
		(0.085)	(0.100)	(0.065)	(0.115)	(0.135)	(0.086)	(0.069)	(0.079)	(0.052)
	N	449	377	826	246	199	445	695	576	1,271
Dropped Out	Control Mean	0.109	0.094	0.102	0.123	0.162	0.139	0.113	0.112	0.113
		[0.312]	[0.292]	[0.303]	[0.330]	[0.370]	[0.346]	[0.317]	[0.315]	[0.316]
	Treatment Effect	0.018	-0.003	0.004	0.035	-0.021	0.015	0.027	-0.006	0.009
		(0.036)	(0.033)	(0.024)	(0.048)	(0.054)	(0.035)	(0.029)	(0.028)	(0.020)
	N	449	377	826	246	199	445	695	576	1,271

Notes: The “Control Mean” rows list averages and standard deviations of the variables in the left-most column within the relevant gender-year subgroup of the control group. The “Treatment Effect” rows report coefficients from regressions of the left-most column variables on a variable equal to one if the student was in the treatment group and zero otherwise, with sampling strata controls (each interaction of gender, year in school and high school grade quartile within the gender/year cell) and controls for high school grade average, whether students' first language is English or not, parental education, and whether they answered a question correctly on how earnings would be calculated. Full-year courses contribute half of their total credit to the credits attempted dependent variable. “Dropped out” is defined as attempting zero credits in fall 2009. Robust standard errors are in parentheses; standard deviations are in square brackets.

Figure 2. Kernel Densities of Fall 2008 Average Grade (per cent) for Treatment and Control

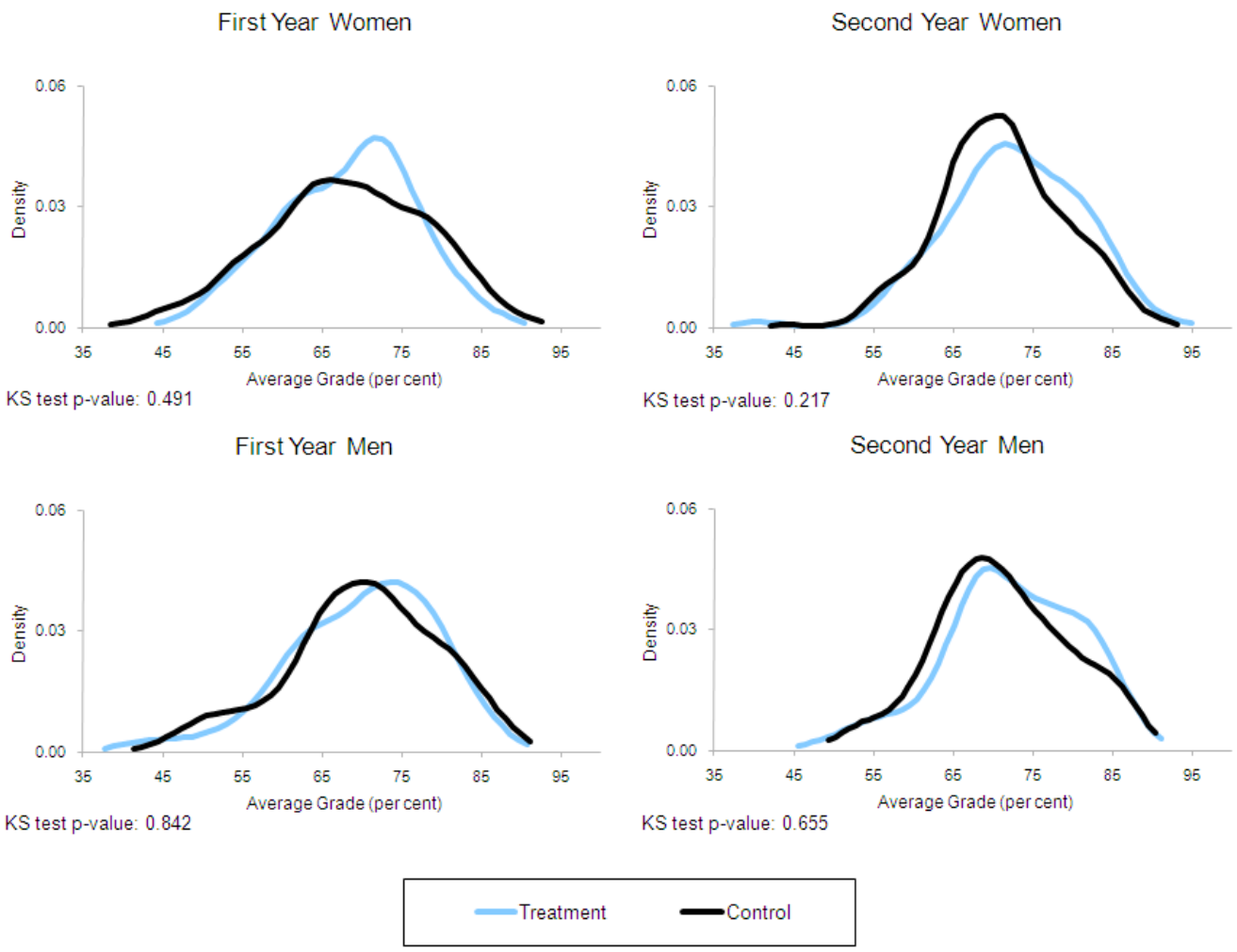


Note: The figures plot the smoothed kernel densities of average grades in fall 2008. Average grades are on a 100-point scale.

Figure 3. Kernel Densities of Winter 2009 Average Grade (per cent) for Treatment and Control

Note: The figures plot the smoothed kernel densities of average grades in winter 2009. Average grades are on a 100-point scale.

Figure 4. Kernel Densities of Full-Year Average Grade (per cent) for Treatment and Control

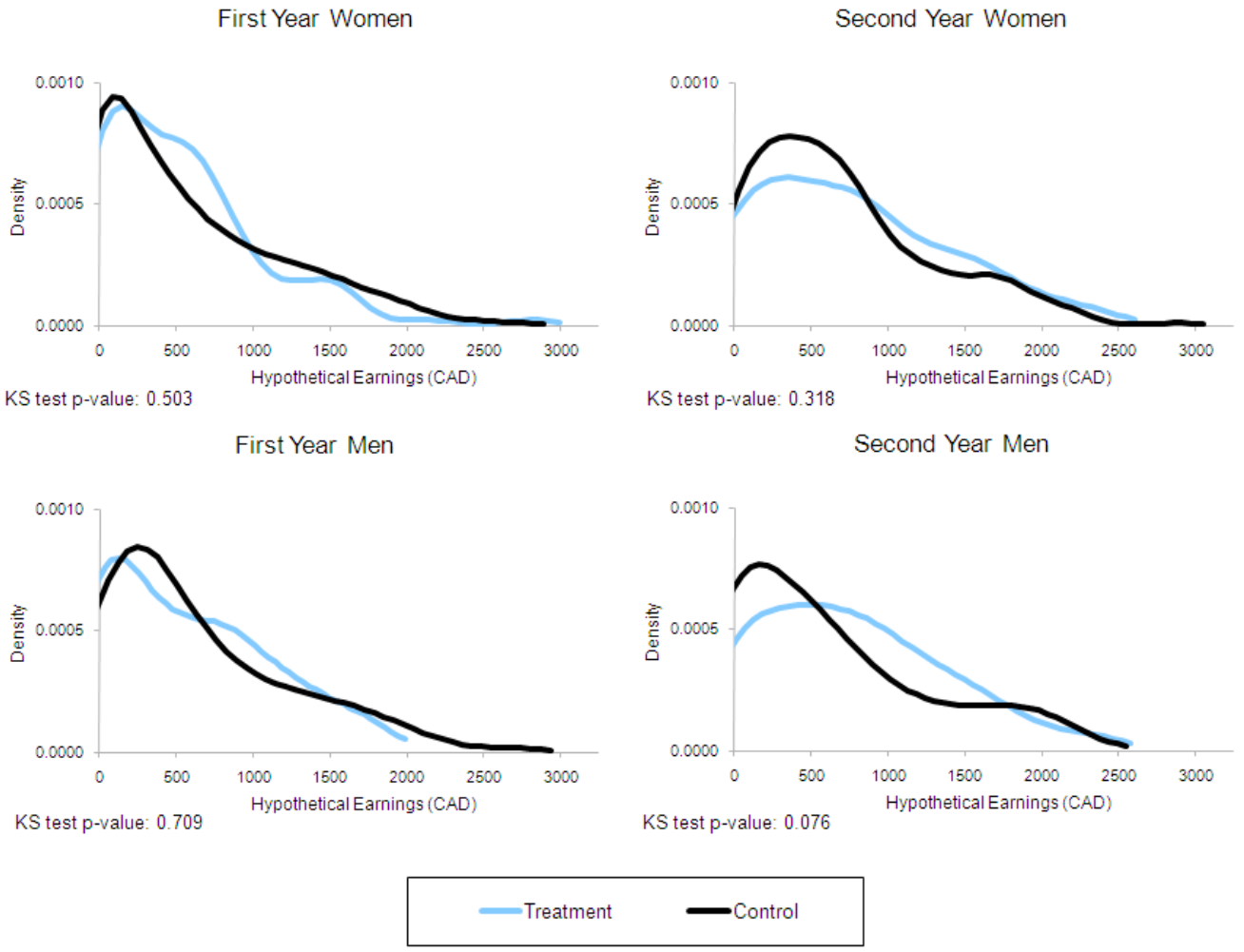


Note: The figures plot the smoothed kernel densities of average grades for the full year from fall 2008 through winter 2009. Average grades are on a 100-point scale. Full-year courses are weighted doubly in the average grade calculation. Students are included in the full-year graphs if they have grades in both semesters.

Figure 5. Kernel Densities of Fall 2008 Hypothetical Earnings (CAD) for Treatment and Control

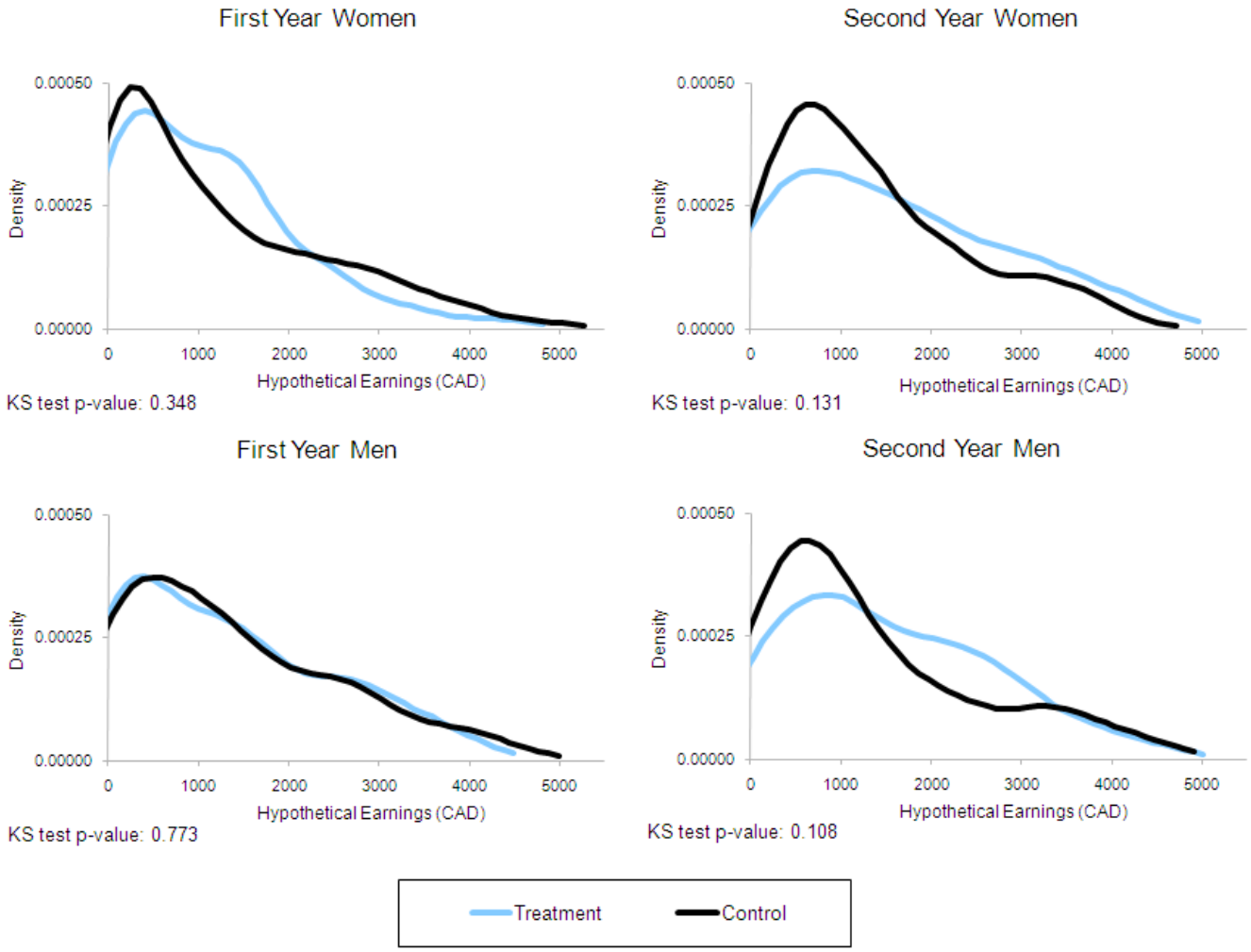
Note: The figures plot the smoothed kernel densities of OK program earnings for fall 2008. For treatment individuals, earnings are equal to actual earnings from the experiment. For control individuals, earnings are equal to what their grades would have earned them if they had been in the treatment group.

Figure 6. Kernel Densities of Winter 2009 Hypothetical Earnings (CAD) for Treatment and Control



Note: The figures plot the smoothed kernel densities of OK program earnings for winter 2009. For treatment individuals, earnings are equal to actual earnings from the experiment. For control individuals, earnings are equal to what their grades would have earned them if they had been in the treatment group.

Figure 7. Kernel Densities of Full-Year Hypothetical Earnings (CAD) for Treatment and Control



Note: The figures plot the smoothed kernel densities of OK program earnings for the full year from fall 2008 through spring 2009. For treatment individuals, earnings are equal to actual earnings from the experiment. For control individuals, earnings are equal to what their grades would have earned them if they had been in the treatment group. Full-year courses are weighted doubly in the earnings calculation. Students are included in the full year graphs if they have grades in both semesters.

References

Anglin, P.M., & Meng, R. (2000). Evidence on grades and grade inflation at Ontario's universities. *Canadian Public Policy/Analyse de Politiques, 26,* 361-368.

Angrist, J., Lang, D., & Oreopoulos, P. (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics, 1,* 136-163.

Angrist, J.D., & Lavy, V. (2002). The effect of high school matriculation awards: Evidence from randomized trials. NBER Working Paper No. 9389.

Ashenfelter, O., & Plant, M.W. (1990). Nonparametric estimates of the labor-supply effects of negative income tax programs. *Journal of Labor Economics, 8*(1, Part 2: Essays in Honor of Albert Rees): S396-S415.

Ashworth, K., Hardman, J., Liu, W-C, Maguire, S., Middleton, S., Dearden, L., Emmerson, C., & Frayne, C. (2001, May). Education maintenance allowance: The first year – A quantitative evaluation. U.K. Department for Education and Evaluation Research, Brief 257.

Association of Canadian Community Colleges and Human Resources and Social Development Canada (2007). *Pan-Canadian study of first year college students: Report 1, Student characteristics and the college experience.* Quebec: Human Resources and Social Development Canada.

Babcock,P.S., & Marks, M. (2010). *The falling time cost of college: Evidence from half a century of time use data.* NBER Working Paper No. 15954.

Bettinger, E.P. (2008). *Paying to learn: The effect of financial incentives on elementary school test scores.* Program on Education Policy and Governance Working Paper 08-03.

Cornwell, C., Mustard, D.B., & Sridhar, D.J. (2006). The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE Program. *Journal of Labor Economics, 24,* 761-786.

DesJardins, S., McCall, B.P. (2008). *The impact of the Gates Millennium Scholars Program on the retention, college finance- and work-related choices, and future educational aspirations of low-income minority students.* Unpublished manuscript.

Dynarski, S. (2002). *The consequences of merit aid.* NBER Working Paper No. 9400.

Dynarski, S. (2005). *Building the Stock of College-Educated Labor.* NBER Working Paper No. 11604.

Fryer, R.G., Jr. (2010). *Financial incentives and student achievement: Evidence from randomized trials.* NBER Working Paper No. 15898.

Garibaldi, P., Giavazzi, F., Ichino, A., & Rettore, E. (2007). *College cost and time to complete a degree: Evidence from tuition discontinuities*. NBER Working Paper No. 12863.

Grayson, J.P., & Grayson, K. (2003). *Research on retention and attrition*. Report to the Canada

Millennium Scholarship Foundation.

Kremer, M., Miguel, E., & Thornton, R. (2009). Incentives to learn. *The Review of Economics and Statistics, 91,* 437-456.

MacDonald, H., Bernstein, L., & Price, C. (2009). *Foundations for success: Short-term impacts report.* Report to the Canada Millennium Scholarship Foundation.

Rae, Bob. (2005, February). *Ontario: A Leader in Learning. Report and Recommendations.* (http://www.edu.gov.on.ca/eng/document/reports/postsec.pdf).

Richburg-Hayes, B., Brock, T., LeBlanc, A., Paxson, C., Rouse, C.E., & Barrow, L. (2009, January). *Rewarding persistence: Effects of a performance-based scholarship for low-income parents.* MDRC Report.

Seligman, J., Milford, R., O'Looney, J., & Ledbetter, J. (2004). *HOPE Scholarship Joint Study Commission report*. Athens, GA: Carl Vinson Institute of Government.