

Do Men and Women Respond Differently to Competition? Evidence from a Major Education Reform

Louis-Philippe Morin, *University of Ottawa*

This paper provides new evidence of gender differences in response to increased competition, focusing on important life tasks performed in a regular social environment. The analysis takes advantage of a major education reform in Ontario that exogenously increased competition for university grades. Comparing students prereform and postreform using rich administrative data, I find that male average grades and the proportion of male students graduating “on time” increased relative to females. Further, the evidence indicates that these changes were due to increased relative effort rather than self-selection. The findings have implications for the delivery of education and incentive provision more generally.

I. Introduction

Experimental evidence from several recent studies, notably by Gneezy, Niederle, and Rustichini (2003) and Gneezy and Rustichini (2004), indicates that males perform better than females in the face of increased

I would like to thank Michael Baker, Branko Boskovic, Pierre Brochu, Victor Couture, Fernanda Estevan, Nicole Fortin, Caroline Hoxby, Josh Lewis, Rob McMillan, Hessel Oosterbeek, Phil Oreopoulos, Uros Petronijevic, Joseph Price, Richard Romano, Aggey Semenov, Aloysius Siow, and seminar and conference participants at Carleton University, l'Université de Sherbrooke, the University of Toronto, Wilfrid Laurier University, the November 2009 NBER Education Program Meeting, and the 2010 EALE/SOLE Meeting for many helpful comments. I am grateful to George Altmeyer, Khuong Doan, Susan Pfeiffer, Pekka Sinervo, and the Faculty of Arts and Science at the University of Toronto for making this

[*Journal of Labor Economics*, 2015, vol. 33, no. 2]

© 2015 by The University of Chicago. All rights reserved. 0734-306X/2015/3302-0002\$10.00

Submitted August 4, 2012; Accepted November 13, 2013; Electronically published January 15, 2015

competition when carrying out specific laboratory tasks.¹ These findings are highly suggestive. They point to possible underlying differences in the cost of effort by gender, with effort expenditure becoming less costly for men relative to women in more competitive environments. To the extent that these differences carry over to important long-term tasks, they have the potential to explain—at least in part—several striking features of the US labor market, including the disproportionate presence of males in highly competitive jobs and gender inequalities in pay.²

It is natural to view differences in labor market outcomes as the result of long-run exposure to competitive environments in which individuals perform regular work-related tasks and are then evaluated according to a competitive selection process or tournament. Ideally, to learn about possible differential responses by gender to increased competition, one would like to mimic the essential features of such a setting. This would include individuals performing day-to-day tasks over an extended period—tasks that they are highly incentivized to care about—and drawing a contrast between performance in low- versus high-competition environments while controlling for all other relevant factors.

The pioneering experimental studies referred to above are able to vary competition exogenously and cleanly control for potentially confounding factors. Yet the tasks these papers have focused on—solving computer mazes and sprinting over short distances, for instance—are somewhat atypical. To lend credence to the notion that underlying gender differences in response to greater competition do indeed extend beyond the laboratory to regular activities in day-to-day life, this study analyzes the impact of an exogenous change in competition on gender performance differences in a regular social environment—the university classroom—focusing on life tasks that individuals have strong incentives to care about. In this instance, the tasks involve trying to master the material covered in university courses—somewhat analogous to skill accumulation in the workplace—with rewards (in the form of grades) being based on relative performance over the course of the academic year.

project possible. All remaining errors are mine. Disclaimer: the views, opinions, findings, and conclusions expressed in this paper are strictly those of the author. Contact the author at lmorin@uottawa.ca. Information concerning access to the data used in this article is available as supplementary material online.

¹ Gneezy et al. (2003) find that the relative and absolute performance of males increases as competition intensifies: males improve their performance, while females do not.

² A variety of possible explanations for female underrepresentation in competitive jobs appear in the literature, including career-family trade-offs (Bertrand, Goldin, and Katz 2010), discrimination (Wolfers 2006), and tastes for competition (Niederle and Vesterlund 2007), as well as differential performance in competitive environments (Gneezy et al. 2003).

More specifically, the analysis makes use of data from a major education reform in Ontario's secondary school system that shortened high school by 1 year, from 5 years to 4 years. As a consequence of this reform, which abolished the final year of high school (or grade 13), two cohorts graduated from high school in June 2003, creating a so-called double cohort of students and drastically increasing competition for university places that year. Importantly, given that capacity did not increase anything close to proportionately, this in turn increased the quality of students who ended up being admitted to university. For universities that grade on a fixed bell curve, like the one studied here,³ the reform made it harder to receive higher grades. That is, each student was now competing with higher quality (and more homogeneous) students for the same fixed grade. In such circumstances, one would expect student effort to play a larger role in the determination of university grades. The incentives to exert higher effort would increase alike for males and females, but if males are more positively (or less negatively) affected by an increase in competition,⁴ then such cross-gender differences could result in a differential shift in effort, manifesting itself as a relative change in performance when comparing male and female outcomes.

It is worth noting that the word "competition" in the context of classroom performance does not have the same meaning as in the traditional experimental economics literature (e.g., Gneezy et al. 2003). In this literature, researchers introduce competition by moving from a piece-rate payoff scheme, where rewards are independent of the performance of others, to a payoff scheme where the relative performance of other subjects matters (a "tournament" pay scheme). In this paper, student performance depends on the performance of the student's peers, even in the "noncompetitive" (prereform) environment. Therefore, I study the impact of an increase in the level of competition.⁵ One could argue that the change in competition level observed here is closer to what we observe in the workplace.

The first empirical goal of this paper is to see whether a differential performance shift occurred based on the increase in the level of competition in university classrooms due to the Ontario double cohort.⁶ With

³ I present the actual marking guidelines suggested to professors in Sec. IV, and I show evidence of bell curve marking in Secs. IV and VI.

⁴ For instance, men might enjoy being in a competitive environment more than females (or dislike it less).

⁵ In fact, the change in the level of competition observed in this paper makes it also close to some of the experimental work done on the effects of affirmative action on effort (by changing the likelihood to win the tournament for some individuals). See, e.g., Calsamiglia, Franke, and Rey-Biel (2013).

⁶ I focus on differential effort by gender while in university, though the empirics below address the possibility that there may have been effects in high school also.

that goal in mind, I examine the impact of increased competition on the gender performance gap using data from the University of Toronto. In this case, the University of Toronto provides a valuable context to shed new light on the gender effects of increased competition for several reasons. First, with more than 55,000 undergraduate students, it is one of the largest universities in North America and so affords very large samples. Second, I have access to rich administrative data that allow me to observe virtually error-free performance measures for students who attended university in a normal environment (before the double cohort) versus a more competitive environment (entering university in September 2003), along with a set of useful controls. Third, the grading scheme guidelines at the University of Toronto clearly indicate a “bell curve”-like marks distribution throughout this time period, unlike other Ontario universities.

The data make it possible to estimate a set of regressions showing whether the relative university performance of male students compared to females, that is, the university gender grade differential, improved in moving from a normal environment to a more competitive one, controlling for student ability. Using this approach, I find that male students gained about 1 percentage point (or 11% of a standard deviation) over female students during their first year in university in response to the increased competition. While modest in size, the effect persists, remaining around 1 percentage point throughout students’ 4 years of university. The increased competition not only affected student grades but also influenced credit accumulation (especially for “below-average” students) and on-time graduation rates, both suggesting that males became relatively more effective as competition increased.

The second empirical goal of the paper is to shed light on the reasons for this differential change in performance. In the observational setting of the double cohort, the performance change might be attributable to differential changes in effort by gender—the hypothesized channel—but also to differential selection and other changes by gender. Although student effort is not directly observable, my approach makes use of the richness of the administrative data to address the most likely potential self-selection issues and alternative channels. To the extent that selection and other confounding factors do not appear to be significant, this points to differential changes in effort by gender as the primary source of the measured performance differences.

On the selection front, the main concern is that competitive environments attract competitive individuals and repel noncompetitive individuals. Thus one might expect males (and especially competitive males) to differentially select into the double cohort relative to females. Looking at information on all applicants, regardless of whether they actually enrolled

or not, the evidence indicates that the double cohort did not have differential selection effects by gender. Specifically, the proportion of females among applicants and enrolled students did not change significantly, comparing the prereform period with the double cohort; nor did female representation within the ability distribution (by high school average quartile); and the proportion of female “fast-trackers”—students from the grade 13 program who graduated early and applied to university in 2002—did not change significantly either. I also investigate whether the increased competition affected dropout and program selection decisions differently by gender: the evidence does not show any gender differences in these decisions. Overall, self-selection bias in this context does not appear to be a cause for concern.

Ideally, one would randomly assign the competition treatment to some students and the control to others. But in the current observational setting, the students in the “high competition” environment are clearly of higher ability than those attending university in the prereform period, on average. Therefore, one necessary assumption for identifying the impact of competition on the gender performance gap is that observable controls for ability are adequate.

One violation of this assumption serves as a potential threat to identification. Since competition might have been tougher, not only in university classrooms but in high school also, the mapping of student ability into high school grades could have changed in the following confounding way. Suppose high school grades in the double cohort overestimated female student ability (relative to males); then this could result in one overestimating male student performance in university relative to females and falsely attributing this increase in performance to heightened university competition. This scenario would require that females were more stimulated by competition than males in both high school and university, but this differential would need to be smaller in university. This seems unlikely given my findings relating to university performance: (i) the unconditional and conditional university grades both suggest that the relative performance of males increased following the double cohort, and (ii) the female representation does not seem to have changed within the ability distribution of applicants as competition increased.

Nevertheless, I address this possibility using information on out-of-province students and students who attended a university prior to the University of Toronto, on the basis that the mapping of ability into grades is not likely to have changed for these students. Estimating regressions using only these students gives results that are very similar to the ones obtained using Ontario secondary school students. Also, quantile regressions do not suggest that the differential impact of competition on performance by gender is driven by a specific type of student—whether of low or high

ability. The evidence points to differential changes in effort by gender over the entire ability distribution.

The rest of the paper is organized as follows: The next section describes existing research on gender and competition and then explains how the Ontario double cohort can shed new light on this topic. Section III sets out a model examining the potential impact of increased competition in university, analyzing the implications for effort choice. Section IV presents the data and also evidence of increased competition at the University of Toronto. The estimation strategy is described in Section V, and in Section VI the main results are presented. Section VII consists of a series of robustness checks, in particular shedding light on whether the results are driven by self-selection. Section VIII then investigates whether increased competition also affected performance beyond students' first year in university, and Section IX concludes, drawing out implications of the analysis.

II. Background

A. Prior Literature

This subsection discusses the well-known research in experimental economics that first examined how males and females respond to increased competition and also the small number of papers that have looked at this issue in an observational setting.⁷ The experimental economics literature on competition and gender performance differences has been motivated primarily by the findings in Bertrand and Hallock (2001). Those authors noted that women only represent 2.5% of the 1992–97 ExecuComp data set, consisting of the top-five executives in each firm of the S&P 500, S&P Midcap 400, and S&P Smallcap 600.⁸ Gneezy et al. (2003) and Gneezy and Rustichini (2004) proposed and experimentally tested an explanation—aside from discrimination or occupational self-selection—as to why we might observe large gender differences in highly ranked (and highly competitive) labor market positions. Specifically, Gneezy et al. (2003) conducted a laboratory experiment in which participants (university students) had to solve as many computer mazes as possible in a given amount of time. While men and women performed equally in the noncompetitive environment, when placed in a more competitive environment, men significantly improved their performance while women did not. The authors concluded that “women may be less effective than men in competitive environments.” Gneezy and Rustichini (2004) presented evidence that this gender differ-

⁷ For a more complete review of the literature on the topic, see Croson and Gneezy (2009) and Niederle and Vesterlund (2010).

⁸ Wolfers (2006) suggests that the female CEO representation is not improving, as only 2.5% of new CEO appointments between 2000 and 2004 were females.

ence might hold at a young age as well, based on a field experiment in which children had to sprint over a certain distance in different competition settings. There also boys' performance was enhanced by direct competition while girls' performance was not.

My paper investigates whether the findings in these pioneering studies also emerge in long-term tasks involving a regular social environment. There is a clear link between my paper and Gneezy et al. (2003): both studies compare performance across individuals, and in both cases one group receives the (increased) competition treatment while the other does not. The main differences lie in the contrasting nature of the tasks studied (solving computer mazes vs. performing in university courses) and the random assignment, which is only possible in the experimental setting. Note that the nature of the task studied here—long-term and mattering for future life outcomes—is key in gauging the potential of competitive gender differences to explain features of the labor market; Section VII of my paper tackles the selection issue.

Few papers have studied the effects of an exogenous change in competition on performance in observational (nonexperimental) environments. Price (2008) looked at the effect of the introduction of the Graduate Education Initiative (GEI), which increased competition within PhD programs, on the time to candidacy for students attending elite US universities. The GEI sponsored fellowships intended to reward "students who made the quickest progress toward completing their degree" (Price 2008). While males decreased their time to candidacy by 10%, women were not affected by the GEI. My paper analyzes a more common type of activity that a large fraction of the population will have to deal with at some point in their life. In the course of doing so, it takes a close look at the potential self-selection impacts of increased competition.

Two recent studies have looked at potential gender differences in performance on university admission tests. Jurajda and Münich (2011) looked at whether females become relatively less likely to be admitted in a university program as it becomes more competitive (as the admission rate of the program decreases). Based on an entire cohort of Czech secondary school graduates, each applying to about three university programs, they find that female applicants are significantly less likely to be admitted to most selective university programs relative to males with similar ability.

In the same vein, Ors, Palomino, and Peyrache (2013) studied gender difference in performance in the highly competitive entry exam to the master of science in management program at one of France's top business schools (HEC Paris). They found that males outperformed females in the entry exam, while admitted female students outperformed their male counterparts during their first year at HEC—a less competitive environment, according to the authors. They also found that female HEC appli-

cants outperformed male applicants in the *baccalauréat* exam, assumed to be a less competitive preapplication exam.⁹ Note that the three types of exam studied by Ors et al. (2013) differed not only in their competition level but also in their format (e.g., oral vs. written exams) and the skills (e.g., general vs. management-specific knowledge) that they tested. As with Price (2008), Ors et al. (2013) does not focus on the potential impact of competition on gender differentials in self-selection.

My paper is also related to Lavy (2013), in that both studies look at a regular competitive environment. Lavy (2013) examines a tournament-like pay scheme rewarding Israeli teachers who performed better than their peers (based on the difference between the predicted and the actual performances of their students). Contrary to the studies cited above, Lavy (2013) does not find evidence of gender differences in performance in this competitive workplace environment.

B. The Ontario Double Cohort

A key challenge in estimating the impact of competition on performance in a natural environment is to find an exogenous source of variation in the competition level faced by individuals. The abolition of grade 13 offers such a source.

As part of a major reform to its secondary school system, the government of Ontario announced the abolition of grade 13 in 1997. Prior to this reform, Ontario students entered college or university after completing grade 13, which contrasted with the other secondary school programs in Canada. Following the reform, students were now expected to complete secondary school in 4 years (after grade 12) instead of 5 years. The first cohort of the new program (henceforth, the G12 program) began secondary school (grade 9) in September 1999. Since students completed the new program faster than its predecessor (the G13 program), the first cohort of the G12 program and the last cohort of the abolished program were expected to graduate and apply to postsecondary institutions in spring 2003, giving rise to the Ontario “double cohort.”

Figure 1 shows the effect of the double cohort on the number of post-secondary institution applicants. There was a significant spike in the number of applicants in 2003, increasing from about 60,000 to more than 100,000 between 2001 and 2003. Since Ontario universities did not expand capacity in proportion, this large increase in the number of applicants made university access more difficult in 2003. We can see the impact of the double cohort on the selectivity of universities by looking at

⁹ Niederle and Vesterlund (2010) suggest using caution when comparing the performance distribution of admitted students to the performance distribution of applicants as the former is a (highly) truncated subsample of the latter.

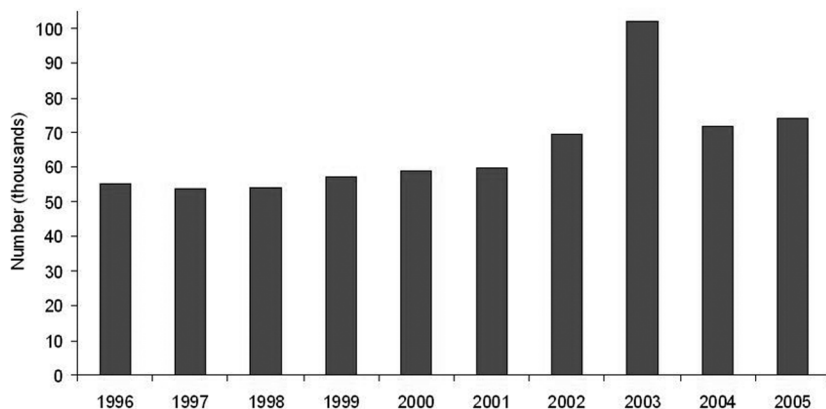


FIG. 1.—Number of Ontario university applicants (in thousands). SOURCE: Ontario Universities' Application Centre (<http://www.ouac.on.ca>).

the application numbers for the University of Toronto. The relevant numbers for 2001 and 2003 are presented in table 1.¹⁰

The increase in the number of applicants between 2001 and 2003 was 61.3%, which is significantly greater than the increase in enrollment (25%). This led the enrollments-to-applicants ratio to drop from 71% to 55%, indicating that university admissions became much more competitive than in previous years.

By increasing competition for university admissions, the double cohort also affected the quality of students enrolled in university in 2003. It is natural to expect students admitted to a university during the double cohort year to be better than students admitted to the same institution a few years before. The level of competition in university should also have been greater since each student would be facing better (and more homogeneous) classmates in terms of high school averages. This is especially true if relative performance is an important component of university grades, as the increased homogeneity in student ability and a fixed “bell curve” can increase the role of effort in determining university grades (as explained in the next section).

It is worth noting that there was a significant increase in the number of applicants in 2002 relative to 2001, looking ahead to some of the selection issues that will be addressed in Section VII. Although this increase is smaller than in 2003, it does suggest that some G13 students escaped the

¹⁰ This is the number of students who put the University of Toronto as their first choice institution when applying through the Ontario Universities' Application Centre (OUAC). The number of students who put the University of Toronto as their first, second, or third choice was around 64,000 for 2003–4 (source: University of Toronto Admissions and Awards, <http://www.adm.utoronto.ca>).

Table 1
University Applicants and Enrollment

	Applicants	Enrollment	Enrollment/ Applicants	Applicant Increase (%)	Enrollment Increase (%)
2001	10,349	7,300	.71		
2003	16,697	9,124	.55	61.3	25.0

SOURCE.—University of Toronto Admissions and Awards.

double cohort by “fast-tracking” secondary school.¹¹ If more females tried to avoid the double cohort by fast-tracking, the populations of male and female students who entered university in 2003 may not be comparable. The evidence presented below indicates that this is not an issue.¹² Also, comparison of the female representation by high school average quartile between 2001 and 2003 indicates that the female representation did not change significantly within the student ability distribution (see table 2 in Sec. IV.A).

III. Increased Competition, Bell Curve Marking, and Effort

In this section, I first examine the impact of increased competition following the double cohort on the link between high school grades and university grades. In particular, I present predictions regarding the expected signs of the parameter estimates from regressing university grades on high school grades if a university grades its students based on a bell curve. This will allow me to verify later on that competition really did increase in university classrooms following the double cohort. I then present a simple model that illustrates a possible channel through which performance gender differences can emerge as competition increases: differential changes in effort incentives.

A. Increased Competition and Grades

This subsection makes explicit the way that increased competition can affect the link between ability (measured by high school grades) and university performance. Suppose that, for every student i enrolled in university in year t , we observe a university grade, U_{it} , and a high school grade, H_{it} . Assume that the university grades on a fixed bell curve such that every year university grades follow a normal distribution with the same mean μ_U and variance σ_U^2 . Further, assume that H_{it} also follows a normal

¹¹ It was possible for G13 students to fast-track secondary school, but only a small number of them would do so. Between 1996 and 2001, the proportion of Ontario university students who had graduated from high school after 4 years fluctuated between 7.7% and 8.5% (King et al. 2002).

¹² I investigate this possibility in Sec. VII.B, but I do not find any significant change in the proportion of female fast-trackers in 2002.

distribution but with mean $\mu_{H,t}$ and variance $\sigma_{H,t}^2$ that are allowed to vary over time.¹³ We can link the two distributions using

$$\mu_U = a_t + b_t \mu_{H,t} \quad (1)$$

and

$$\sigma_U^2 = b_t^2 \sigma_{H,t}^2, \quad (2)$$

where b_t is expected to be positive.

Figure 2 illustrates the link between high school and university grades. The top panel of figure 2 presents the distribution of high school grades of students admitted to university, while the bottom panel presents the distribution of university grades.

Assume that competition for university admission increases between t and t' . As competition increases, the distribution of high school grades admitted to university will shift to the right—the average will increase from $\mu_{H,t}$ to $\mu_{H,t'}$ —and, in turn, the variance, $\sigma_{H,t}^2$, will decrease. These changes in mean and variance are simply due to the limited admission capacity of the university. This phenomenon is illustrated in the top panel of figure 2.

Since the university keeps the same performance distribution from year to year (i.e., μ_U and σ_U^2 remain unchanged), the increased competition will affect the link between high school grades and university grades. In the top panel of figure 2, we can see that $\sigma_{H,t}^2$ decreases as competition increases. Therefore $b_{t'}$ will be larger than b_t (from a less competitive year) to guarantee that equation (2) still holds. Also, since both $\mu_{H,t}$ and b_t increase, a_t will decrease to ensure that equation (1) is satisfied.

The framework presented above makes two predictions regarding the parameter estimates from estimating a regression equation of the following form:

$$U_{i,t} = \alpha + \gamma H_i + \pi DC_t + \rho(DC_t \times H_i) + \omega_{i,t}, \quad (3)$$

where $U_{i,t}$ is student i 's university grade, H_i is student i 's high school grade, and DC_t is a dummy variable equal to one if student i is observed during the more competitive year. Equation (3) can be seen as the base regression equation used in this paper. Here π and ρ will capture changes in the constant and slope parameters in equation (1)—that is, $\pi = a_{t'} - a_t$ and $\rho = b_{t'} - b_t$ from the earlier discussion—following the increase in competition. If the university grades on a bell curve, we should first expect the slope coefficient to be greater ($\rho > 0$) and, second, the intercept to be smaller

¹³ Fig. 4 suggests that a normal distribution is a reasonable approximation to the actual grade distribution.

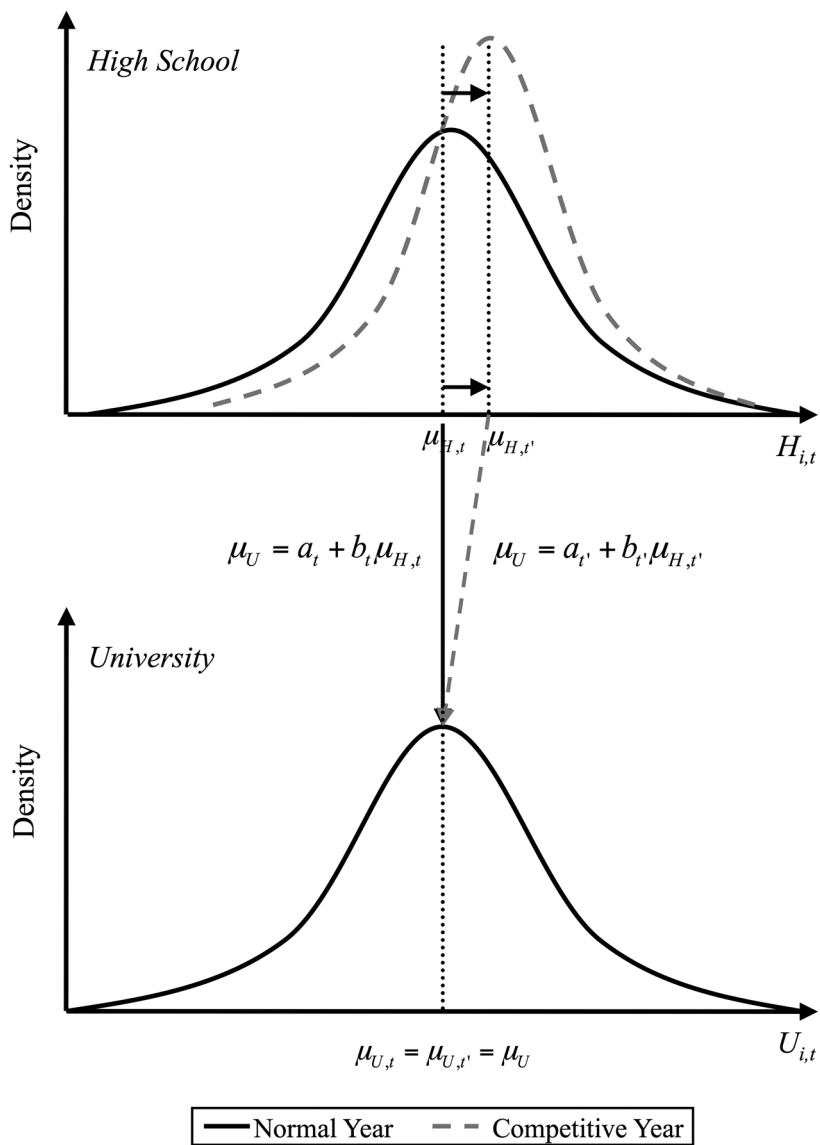


FIG. 2.—Increased competition, bell curve marking scheme, and grades. *Top*, distribution of high school grade of students admitted to university; *bottom*, distribution of university grades.

($\pi < 0$) in more competitive years. Note that these predictions are easily testable with the available data. The regression results in Section VI suggest that $\rho > 0$ and $\pi < 0$.

B. Increased Competition and Effort in University

The previous model made explicit the link between competition, grades, and bell curve marking. I now present a simple model to analyze the potential impact of increased competition and bell curve marking on the effort incentive in university classrooms.

Imagine students with ability a will exert effort e when studying to acquire knowledge, $K(e, a)$. Assume that $K(e, a)$ is continuously differentiable and increasing and concave in e and a . The cost of studying is given by a continuously differentiable function $C_g(e, a, l)$ (in e and a) and increasing and convex in e —these assumptions yield well-behaved marginal cost and marginal benefit of effort schedules. The cost to studying also depends on the student's gender $g \in \{f, m\}$ and on the level of competition l in the classroom. Increases in the level of competition will decrease the marginal cost of studying if students enjoy competition. Suppose that in the absence of competition (at $l = 0$), males and females have the same cost functions:

$$C_f(e, a, 0) = C_m(e, a, 0) = C(e, a, 0).$$

Therefore, females and males are expected to exert the same level of effort and perform equally for a given a —they have the same knowledge production and cost functions at $l = 0$.

Assume that males enjoy an increase in the level of competition more than females, in the sense that the marginal cost of effort decreases more for males as the level of competition increases:¹⁴

$$\frac{\partial C_f^2(e, a, l)}{\partial e \partial l} > \frac{\partial C_m^2(e, a, l)}{\partial e \partial l}.$$

Now, the university grades its students according to the following scheme:

$$G(e, a, l) = \eta_l + \lambda_l K(e, a),$$

where the university adjusts η_l and λ_l to satisfy its bell curve marking scheme. In particular, the university is expected to increase the benefit to effort by increasing λ_l (and to decrease η_l) in a more competitive year.¹⁵

¹⁴ The results are exactly the same if they dislike competition less than females.

¹⁵ In app. B, secs. B.1 and B.2, I present a model that shows the impact of increased competition on the effort incentives set by a university that grades on a bell curve. I show that under mild conditions (depending on the complementarity between effort and ability), the university will increase λ_l as competition for university admission increases.

Student utility is simply defined by the difference between the university grade and the cost of studying. Therefore, the student utility maximization problem is

$$\max_e \eta_l + \lambda_l K(e, a) - C_g(e, a, l).$$

The first-order condition for this problem is

$$\lambda_l \frac{\partial K(e_g^*, a)}{\partial e} - \frac{\partial C_g(e_g^*, a, l)}{\partial e} = 0. \quad (4)$$

The question is whether effort will increase more for males than for females when we move from a noncompetitive environment ($l = 0$) to a competitive environment ($l > 0$). Figure 3 illustrates how female and male levels of effort will change following an increase in the level of competition. The details relating to this comparative statics exercise are in appendix B.

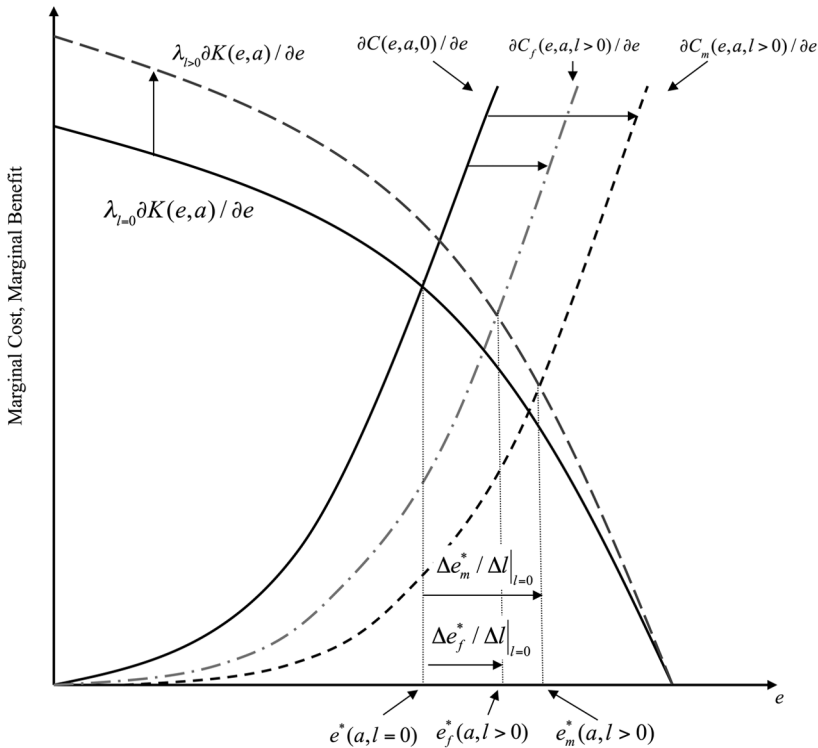


FIG. 3.—Increased competition and effort level

First, recall that in the absence of competition (at $l = 0$), females and males are assumed to have the same knowledge production and cost functions. Therefore, in the absence of competition, females and males will both choose an effort level equal to $e^*(a, l = 0)$. An increase in competition will affect both the marginal benefit (by increasing λ_l) and the marginal cost of exerting effort. The increase in marginal benefit will be the same for females and males, but the change in marginal cost will differ if males enjoy competition more than females. More specifically, males will experience a more pronounced rightward shift of their marginal cost curve than females. Consequently, male effort will shift from $e^*(a, l = 0)$ to $e_m^*(a, l > 0)$, while female effort level will only shift from $e^*(a, l = 0)$ to $e_f^*(a, l > 0)$, as in figure 3. Under these circumstances, an increase in competition is expected to differentially affect the effort level of females and males.

IV. Data

In order to look at the effect of the increased competition in university classrooms on academic performance, I focus on students who enrolled at the University of Toronto—one of the largest universities in North America. I use an administrative data set provided by the University of Toronto Faculty of Arts and Science that is composed of first-year students who started in Arts and Science in September 2001 or September 2003.¹⁶ In 2010, close to 22,000 undergraduate full-time students were attending University of Toronto's Faculty of Arts and Science, making it the largest faculty of the University.

The Faculty of Arts and Science combines three features necessary for analyzing the effect of competition on grades. First, and most important, the institution analyzed must grade its students based (at least implicitly) on a bell curve. Unlike a number of Ontario university faculties, the Faculty of Arts and Science had specific guidelines for the marks distribution, clearly suggestive of a bell curve marking scheme.¹⁷ Second, the faculty has large introductory classes—many of these classes are compulsory (for a given program), and some of them have more than 1,000 students—making it possible to get relatively precise estimates. Third, most study

¹⁶ For related studies using college administrative data, see Sacerdote (2001) using Dartmouth College data; Stinebrickner and Stinebrickner (2006) using Berea College data; Angrist, Lang, and Oreopoulos (2009), Hoffmann and Oreopoulos (2009a, 2009b), and Lindo, Sanders, and Oreopoulos (2010) using data from a large Canadian university.

¹⁷ Until 2009, in the "Academic Handbook: Course Information for Instructors" that was available on the University of Toronto Faculty of Arts and Science website (<http://www.artsci.utoronto.ca>), we could read that, although not required, experience suggests that there will normally be between 5% and 25% of A's, not over 75% of combined A's and B's, and not over 20% of combined E's and F's in large classes. Since 2009, the marks distribution guidelines are slightly broader.

subjects offered by the faculty were not affected by the Ontario secondary school reform.¹⁸ This is not the case with other faculties.

The data come from two sources of information that were linked using students' identification numbers: preuniversity admission information and university academic histories. The university academic history contains (i) the numerical grades for all Arts and Science courses that the student completed, (ii) the list of courses that the student dropped (along with the dates these courses were dropped), and (iii) a dummy variable, "on-time graduation," indicating whether the student had graduated from the university by July 1 of his/her fourth year (e.g., July 1, 2007 for 2003 students).

The main dependent variable consists of individual course grades, on a 100-point scale. I focus on performance in larger classrooms (i.e., with at least 30 students) since bell curve marking in smaller classrooms is less likely and this type of marking scheme is assumed to be a driving factor behind my results. About 90% of first-year grades come from these larger sections. As a robustness check, I present the results from smaller classrooms in Section VI.C. I also look at a student's first-year (credit-weighted) averages (on a 100-point scale), GPA, dropped courses, credit accumulation, and "on-time graduation."¹⁹ Although I eventually look at student performance from the first through fourth university years, I concentrate the analysis on first-year performance. An obvious advantage of doing so is that it mitigates much of the potential impact of course selection, as many first-year courses are, as noted, compulsory given a specific program of study.

A nice feature of the data is that they contain preuniversity admission information for students who applied to the Faculty of Arts and Science, regardless of whether their application was accepted or rejected. Hence, this information will be used not only to control for students' backgrounds but also to look for evidence of increased competition in university ad-

¹⁸ The compression of the Ontario secondary school curriculum affected the delivery of material for some subjects and not others. For example, the mathematics curriculum was affected, while the biology curriculum was not (Morin 2013). See King et al. (2002, 2004, 2005) and Morin (2013) for more details. As explained below, all mathematics courses are excluded from the analysis.

¹⁹ Students earn 0.5 credits for successfully completing a 1-semester course and a 1.0 credit for successfully completing a 2-semester course. As students take 1- and 2-semester courses, the student's first year average is weighted by the credit units. I computed first-year averages for all students for which I observe at least two numerical grades by the end of the first year. I consider that students who report grades under 30% have, in fact, dropped the course. These course outcomes will be used, with the courses that were officially dropped, to see whether females (or males) were more likely to drop courses (or "stop competing") in 2003. As discussed in Sec. VII.A, including these grades would actually make my results slightly stronger.

missions. This information will also be crucial to investigate whether females tried to avoid the increased competition (more than males) by choosing less competitive programs or simply by not applying/enrolling to the university in 2003, which is plausible if one thinks that females are more likely to “shy away from competition” (Niederle and Vesterlund 2007).

For each applicant, I have the following information: a student identification number, the applicant’s high school average, his/her year and month of birth, his/her gender, the name of the school attended by the applicant, the faculty program applied to (commerce, computer science, humanities and social sciences, or life science), and his/her application status (enrolled, accepted, canceled, or refused).²⁰

I restrict the sample to Ontario high school graduates born in 1984 and 1985 for the 2003 student cohort and in 1982 for the 2001 cohort in order to avoid having the results affected by older students.²¹ Out-of-province students and students who already had some university experience before applying to the faculty will be used in Section VII.C for robustness checks.²²

The data contain an indicator of the secondary school curriculum (grade 12 or grade 13) the student graduated from. The G12/G13 indicator is necessary for performing the analysis with or without G12 students. There are pros and cons to including G12 students in the sample. Focusing on G13 students (excluding G12 students from the sample) guarantees that, aside from potential differences in academic ability, students from 2001 and 2003 should be quite similar in terms of academic background (e.g., they come from the same secondary school program) and other dimensions such as maturity—since students have the same age. But, if grade 13 had a significant impact on students’ university preparation and if gender composition differs across G12 and G13, then not including G12 students could overestimate or underestimate the effect of the double cohort on gender differences in performance. In addition, Morin (2013) presents evidence that G12 students who entered university

²⁰ Students interested in a specific study program offered by the University of Toronto Faculty of Arts and Science have first to apply to one of the following general programs: commerce, computer science, humanities and social sciences, or life science. Enrolled status entails application acceptance (from the faculty) and the student choosing to enroll.

²¹ The birthday cutoff date for primary school enrollment is December 31 in Ontario. Hence, students from the first G12 cohort are supposed to be born in 1985, while students from the last cohort of the G13 program should be born in 1984. As a robustness check, I also estimated the regression model including older students. The inclusion of these students does not affect the results.

²² Preadmission information from many out-of-province students is missing, since they do not necessarily apply through the same process as Ontario high school graduates. Admission information for these out-of-province students is kept by the colleges to which they applied to and not by the Faculty of Arts and Science.

in 2003 were better-than-average students. If this is the case, and if the gender composition differs across G12 and G13 students, then excluding G12 students could also result in biased estimation of the competition effect even if we assume no difference in university preparation across curricula. For this reason, the estimations were all done with and without G12 students. Results show that including or excluding G12 students give very similar estimates.

Finally, the secondary school mathematics curriculum was clearly affected by the reform (Morin 2013). Since female and male students might have reacted differently to the change in the mathematics curriculum, all mathematics courses were excluded from the analysis.²³ This should mitigate any potential bias due to differences across G12 and G13 university preparation. In practice, however, including mathematics performance in the university average does not affect my results.

The drop between 2001 and 2003 in the enrollment/application ratio presented in table 1 already suggests a significant impact of the double cohort on university admission. Table 2 presents descriptive statistics used to look at the effect of increased competition on student university performance.

The key point coming out of panel A is that the double cohort significantly increased competition for university admission and the quality (based on students' high school averages) of enrolled students. The average high school grade increased by 2.6 percentage points between 2001 and 2003. This difference is statistically significant and considerable—representing an increase of close to 55% of a standard deviation (relative to 2001).

The increase in student quality between 2001 and 2003 can be illustrated by comparing the distributions of enrolled students' high school averages. The top panel of figure 4 plots estimated densities of high school averages for students enrolled at the Faculty of Arts and Science in 2001 and 2003. Clearly students enrolled in 2003 have higher high school averages than students who enrolled in 2001.

What further suggests that competition increased in university classroom is the use of bell curve grading at the university. The university average was around 70% in 2001 and increased by about 1.4 percentage points (less than 15% of a standard deviation) in 2003. As can be seen from the bottom panel of figure 4, this increase is significantly less (in absolute or relative terms) than the increase in high school averages, suggesting that it became harder to get higher grades. Credit accumulation and “on-time graduation” rates also suggest a modest performance increase between 2001 and 2003. In sum, evidence from figure 4 and tables 1 and 2

²³ There is a large literature on the gender gap in mathematics. Interested readers should consult Ellison and Swanson (2010), Fryer and Levitt (2010), and Niederle and Vesterlund (2010) for recent developments in the topic.

Table 2
Descriptive Statistics

	2001	2003	
	G13	G12	G13
A. Main descriptive statistics:			
High school average	85.4 (4.86)	88.2 (4.41)	87.9 (4.43)
Age	19.2 (.28)	18.2 (.28)	19.2 (.28)
Female (%)	61.1	62.4	61.3
Observed in year 2 (%)	90.0	91.7	92.4
Observed in year 3 (%)	77.3	84.5	82.1
Observed in year 4 (%)	71.9	78.1	75.3
First year university average*	69.7 (9.30)	71.2 (9.94)	70.8 (9.42)
Attempted credits (first year)†	4.01 (.80)	4.10 (.70)	4.13 (.79)
Earned credits after 1 year†	3.84 (.96)	3.92 (.91)	3.94 (.98)
Earned credits after 4 years†	13.1 (4.55)	14.9 (4.77)	14.5 (4.90)
“On-time graduation” (%)	39.9	42.1	38.4
B. Females by high school average quartile (%):			
Bottom quartile	56.0	56.5	55.3
Second quartile	60.4	64.2	60.1
Third quartile	62.3	66.4	64.2
Top quartile	65.8	62.7	65.7
Number of observations	2,463	1,702	1,835

NOTE.—Standard deviations are in parentheses.

* Excludes grades from smaller classes (i.e., with fewer than 30 students) and mathematics grades.

† Excludes mathematics credits. The number of credits earned is zero for year X ($X = 2, 3, \text{ or } 4$) if the student is not observed in year X .

all suggest that competition in classrooms increased significantly as a consequence of the double cohort.

As expected, G12 students are on average exactly 1 year younger than G13 students (see panel A of table 2). The student population is composed of a majority of females, comprising more than 60% of the population. Of note, the proportion of female students did not change significantly between 2001 and 2003. This is important evidence, especially if one is concerned that there could be significant gender differences in the taste for competition.

Panel B of table 2 presents the proportions of female students by high school average quartile to further investigate whether there were significant changes in the female representation. Females tend to be slightly less present in the bottom quartile. There was, however, no drastic change in the female representation by quartile between 2001 and 2003. This is especially true if we compare the 2001 and 2003 G13 students. This is im-

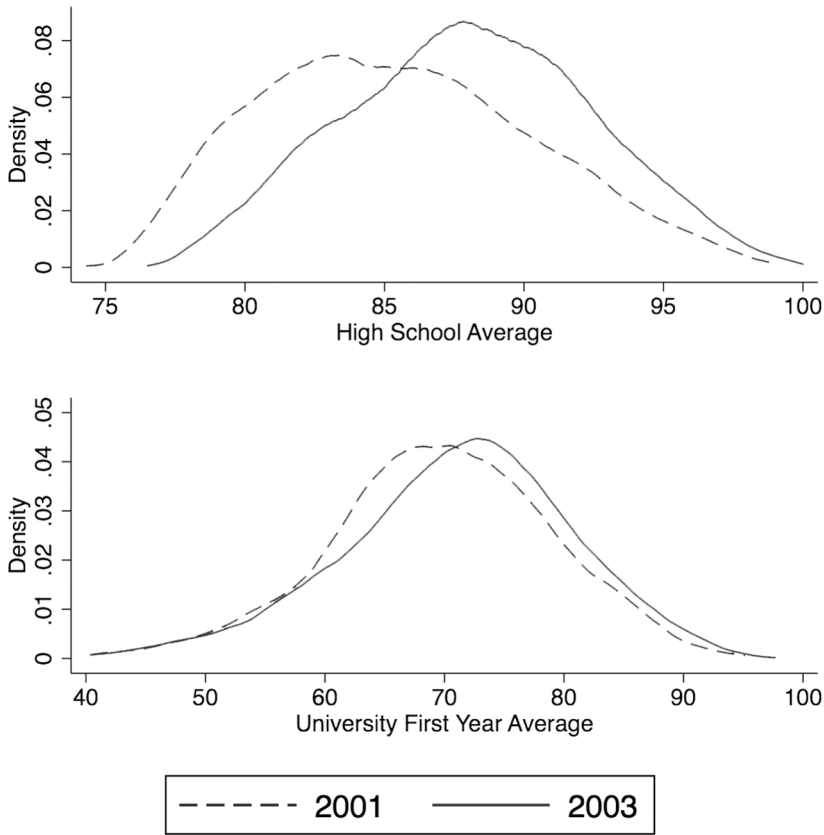


FIG. 4.—High school average and university first-year average distributions of enrolled students in 2001 and 2003.

portant evidence, as it suggests that the female representation did not change significantly within the “ability” distribution—something that could signal some significant selection issues.

V. Estimation

The main estimation strategy used to gauge the effect of increased competition on gender performance differences is captured by the following extension of equation (3):²⁴

$$U_{i,t} = \alpha + \gamma H_i + \delta \text{Male}_i + \pi \text{DC}_t + \rho(\text{DC}_t \times H_i) + \beta(\text{DC}_t \times \text{Male}_i) + X_i \Gamma + \omega_{i,t}. \quad (5)$$

²⁴ Quantile regressions are also estimated in Sec. VI.B to investigate the possibility of heterogeneous effects of competition on the university performance distribution.

In this equation $U_{i,t}$ is a university performance measure (the student's university average, the number of university credits earned, or on-time graduation). The main outcome of interest is a student's performance in a large first-year university course $U_{i,t,c}$ (where c stands for a specific course). For this specific outcome, equation (5) is augmented with course fixed effects, ψ_c . One of the main benefits of looking at first-year performance is that during this year the choice of courses is not as important as for later years, which can mitigate potential course selection issues. Furthermore, first-year courses usually have very high enrollments, as mentioned above.

The variable H_i is a measure of student ability (high school average), while Male_i and DC_i are a male dummy variable and a double cohort year dummy, respectively. Finally, X_i is a vector of other personal characteristics. The coefficient of interest is β , which represents the difference, across genders, in the effect of the double cohort. The coefficients π and ρ will capture common effects (to males and females) of the increased competition. Notice that the estimates of π and ρ are expected to be negative and positive, respectively (as in eq. [3]). The coefficient δ allows one to test whether (*ceteris paribus*) males perform better than females at the university. Vector X_i will consist of controls like program fixed effects and age, which will be added to the equation in some specifications. Controlling for age could be useful when including G12 students in the analysis.

VI. Results

A. Competition and First-Year Performance

Figure 5 shows the evolution of the university's first-year average distributions between 2001 and 2003 for males and females separately, already pointing to gender differences in response to the increased competition following the double cohort. First, the top panel of figure 5 shows that male and female pre-double cohort (2001) performance distributions are similar. A Kolmogorov-Smirnov test for equality of distributions suggests that the two distributions are identical. Things are different in 2003. We see a clear shift to the right for both males and females (in the bottom panel), suggesting that the unconditional student performance increased in 2003. More important, the distributional shift for males is more pronounced—the male and female performance distributions are now statistically different. Interestingly, when comparing the female and male 2003 performance distributions, the male distribution looks, more or less, like a translation (to the right) of the female distribution. Although figure 5 only plots “unconditional” performance distributions, the regression results presented in table 3 and the quantile regression results presented in Section VI.B will support the idea that the difference in the shifts in performance distributions captures the effect of the increased competition (as opposed to changes in student characteristics).

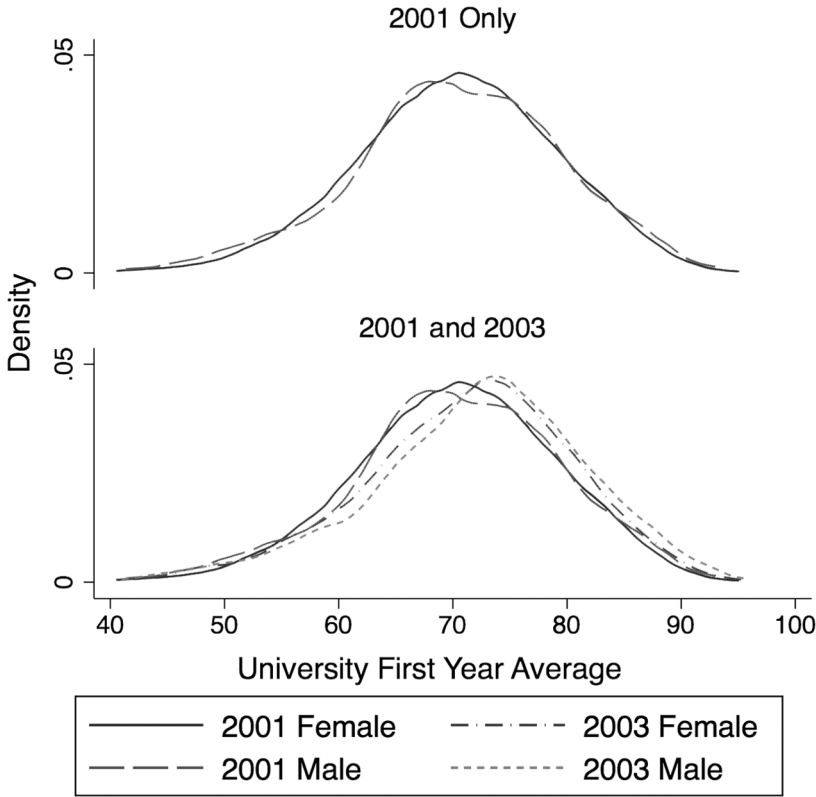


FIG. 5.—University average distributions

Table 3 presents results from estimating the course-level version of equation (5) with a course grade (on a 100-point scale) as the dependent variable for different sets of controls and for different subsamples. The results in table 3 will shed light on (i) whether it became harder to get high grades as a consequence of the double cohort (i.e., whether the university graded on a bell curve) and (ii) whether (on average) males' performance significantly improved relative to females.

Columns 1–3 of table 3 only include students who graduated from the G13 program, while columns 4–6 also include students who graduated from the G12 program. The effect of the increased competition on the university grading policy slope and intercept coefficients (π and ρ in eq. [5]) are captured by "Double cohort" and "High school average \times Double cohort," respectively. "Male \times Double cohort" gives an estimate of the difference across genders in the effect of the increased competition (β).

Results in column 1 of table 3 do not include any personal characteristics aside from students' high school average and gender. The estimates

Table 3
Impact of the Double Cohort on the Gender Performance Gap (Individual Courses)

	G13 Students Only			G12 and G13 Students		
	(1)	(2)	(3)	(4)	(5)	(6)
High school average	1.074*** (.041)	1.176*** (.050)	1.177*** (.050)	1.074*** (.041)	1.152*** (.049)	1.153*** (.049)
Double cohort	-11.86** (4.893)	-15.51*** (4.439)	-15.38*** (4.441)	-18.05*** (4.570)	-20.92*** (4.283)	-20.94*** (4.285)
Male	.908*** (.284)	.843*** (.235)	.862*** (.235)	.908*** (.284)	.782*** (.251)	.790*** (.252)
High school average × Double cohort	.117** (.054)	.152*** (.050)	.150*** (.050)	.187*** (.051)	.215*** (.048)	.214*** (.048)
Male × Double cohort	1.061*** (.362)	1.019*** (.348)	1.009*** (.345)	1.042*** (.341)	1.014*** (.342)	1.015*** (.341)
Age			-.778*** (.232)		-.329** (.151)	
Course fixed effects	No	Yes	Yes	No	Yes	Yes
Number of observations	18,152	18,152	18,152	25,501	25,501	25,501
Students	4,298	4,298	4,298	6,000	6,000	6,000
R ²	.225	.245	.245	.238	.249	.249

NOTE.—The dependent variable in these sets of regressions is a first-year university course grade. Clustered standard errors (at the course level) are in parentheses.

** Significant at the 5% level.

*** Significant at the 1% level.

for the changes in the university grading policy slope and intercept coefficients have the expected signs ($\rho > 0$ and $\pi < 0$), supporting the idea of increased competition in classrooms in 2003. These results suggest that university grades did not fully adjust for the increased quality of students in their classrooms. This is to be expected if we think that many universities recommend implicitly or explicitly using a bell curve marking scheme. Not surprisingly, the effect of the double cohort on university grades is statistically significant. An “average” 2003 female student—with an 88% high school average—had a 1.56 percentage point (about 0.17 SD) disadvantage when compared to a similar student who entered the university in 2001.²⁵

The estimated change in the gender performance gap due to the increased competition (1.06 percentage points) is statistically significant at the 1% level, but modest (0.114 SD), suggesting that males coped better with the increased competition than females. These results are in line with the findings of Gneezy et al. (2003) and Gneezy and Rustichini (2004). Note that, contrary to Gneezy et al. (2003) and Gneezy and Rustichini (2004), I can only measure a relative (as opposed to absolute) change in performance due to the increased competition since the university seems to be grading on a bell curve (given the estimated coefficients for ρ and π).

The effect of the double cohort varies significantly across the student population. For example, imagine two female students, one with a high school average one standard deviation (4.4 percentage points) below and the other with a one standard deviation above the 2003 mean high school average (88%). The student below the mean would have had a 2.05 percentage point disadvantage from being part of the double cohort. For the student above the mean, the disadvantage would be significantly smaller (1.02 percentage points).

Controlling for course fixed effects and age (cols. 2 and 3 of table 3) does not alter the findings. The coefficients estimates on “High school average” and “Male \times Double cohort” are very stable across specifications. Finally, the age effect is statistically significant but modest and negative: when comparing the youngest and oldest students coming out of the G13 program, we expect the youngest students to have a 0.78 percentage point advantage over the oldest.

Columns 4–6 of table 3 replicate the estimates in columns 1–3, using the complete sample of Ontario students, including both G13 and G12 students. The inclusion of G12 students does not affect the estimated impact

²⁵ $88 \times 0.117 - 11.86 \approx -1.56$. The null hypothesis $H_0: \pi + 88\rho = 0$ is rejected at a 1% confidence level. In order to convert the effect in terms of standard deviations, I use the 2001 university average standard deviation (9.30) as found in table 2: $-1.56/9.3 = -0.168$.

of increased competition on the gender difference in performance; it remains around 1 percentage point.

One might think that the effect is modest enough not to have any impact on students' grade point averages (GPAs). Using student GPA (on a 4-point scale) or a credit-weighted average (on a 100-point scale) as the dependent variable gives results very similar to the ones presented in table 3. The estimates for "Male \times Double cohort" when using students' credit-weighted averages fluctuate between 1.03 and 1.13 points and between 0.08 and 0.10 GPA points when using students' GPAs (or between 0.10 and 0.11 SDs). These results are presented in tables A1 and A2 in appendix A, respectively. In the end, all three measures of university performance (i.e., course-level performance, GPA, credit-weighted average) suggest that male students coped better with the increased competition than females.

B. Heterogeneity

The least-squares regressions capture the effect of increased competition on the average female-male performance difference. Although informative, they may mask important heterogeneity in the impact of competition. The least-squares estimates could, for example, be driven by a specific group of students, such as male students in the upper tail of the performance distribution, who react more to competition than the rest of the student population. In this case, the increased competition would affect the shape of the performance distribution. In order to investigate this possibility, I estimated equation (5) using a quantile regression methodology proposed by Firpo, Fortin, and Lemieux (2009) and where the dependent variable is students' first-year (credit-weighted) averages (the averages plotted in fig. 5). This methodology allows one to estimate the impact of "Male \times Double cohort" on the quantiles of the unconditional university performance distribution, shedding light on a possible increase in performance dispersion due to the increased competition.

Figure 6 plots the quantile regression "Male \times DC" estimates for the 5th to the 95th quantiles and its 90% confidence interval band.²⁶ The effect of competition on the gender performance gap seems relatively stable over the performance distribution, as the least-squares point estimate is covered by the entire quantile regression band. What is clear from the quantile regression results is that the estimated effects of the increased competition on the gender performance gap found above are not overwhelmingly driven by effects at specific points (say top or bottom) of the university performance distribution. They further suggest that the difference in the shift in performance distributions found in figure 5 is a reasonable illustrat-

²⁶ The 90% confidence band is based on bootstrapped standard errors. The standard errors of each quantile are obtained from 500 replications.

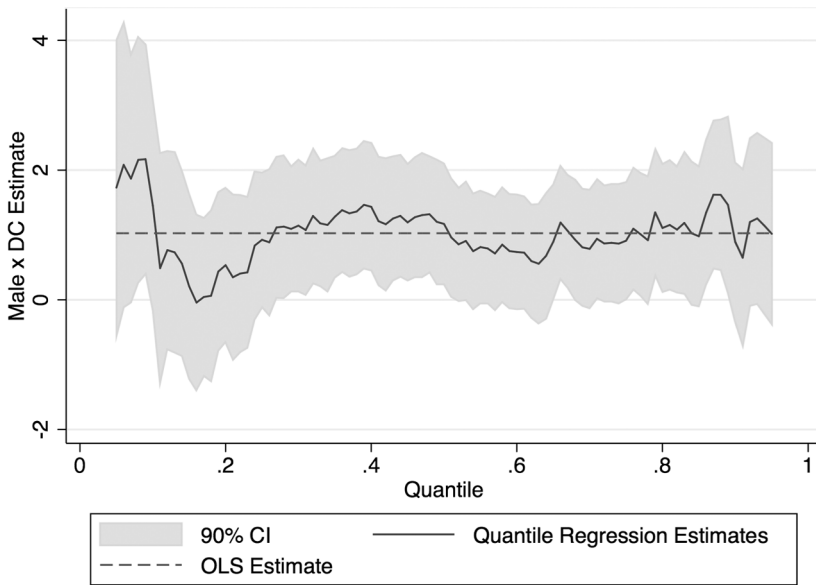


FIG. 6.—Quantile regression estimates

tion of the impact of the increased competition on the gender performance gap.²⁷

C. Smaller Sections

Evidence from table 3 suggests that instructors graded on a bell curve in large classes—recall that the sign of coefficient estimates for “Double cohort” and “High school average \times Double cohort” are in line with bell curve marking. I argue earlier that this bell curve marking combined with the increased quality of students gave rise to an increase in the level of classroom competition. However, we should not expect the same kind of effects in smaller classes, as the suggested grading scheme (see footnote 17) was intended for larger classes (e.g., those with more than 20–30 students). Therefore, the increase in competition for grades should be lower in smaller classes, and we should expect the effect of the double cohort on the gender performance gap to be minimal.

In order to investigate whether there was evidence of bell curve marking in smaller classrooms, and also to see whether I find any evidence of a double-cohort effect on the gender performance gap, I reestimated

²⁷ An alternative, albeit more restrictive, strategy to investigate the possibility of heterogeneous effects is to add interaction terms (i.e., “Male \times High school average \times Double cohort,” with and without “Male \times High school”) to eq. (5). The parameter estimates of these interaction terms are always small and statistically insignificant at conventional confidence levels.

equation (5) but used only grades from sections with fewer than 30 students.²⁸ Appendix table A3 does not suggest that instructors mark on a bell curve in these smaller classes: the coefficient estimates for “Double cohort” and “High school average \times Double cohort” are small in magnitude (compared to the ones found in table 3), and they have the opposite signs that we would expect if instructors were to grade on a bell curve. Table A3 suggests not only that instructors do not grade on a bell curve but also that the increased student quality did not affect the gender performance gap in these smaller classes: the coefficient estimates for “Male \times Double cohort” are all small, fluctuating between -0.166 and 0.271 , and they are statistically insignificant. The combination of these two findings support the idea that bell curve marking could be a crucial driving factor behind the change in the gender performance gap observed following the double cohort—when there is an increase in student quality.

VII. Selection Issues

The results presented so far suggest that male students coped better with the increased competition than females. These results do not take into account the possibility that students are more or less free to drop out of courses/programs and to enroll and to apply to the university, or that competition might have increased not only in universities but in secondary schools as well. This section investigates self-selection problems and looks at the potential consequences of increased competition at the secondary school level on the estimates presented so far. I start by considering the possibility that males dropped out of courses disproportionately more than females as a consequence of the double cohort, which would (probably) lead me to overestimate the gender difference in performance in 2003. Next, I use the available admissions data to investigate any gender differences in the enrollment or application decisions—the admissions data allow me to observe any student who applied to the Faculty of Arts and Science (whether they were accepted or not). I also look at pre-university competition and its potential impact on the link between students’ ability and their high school averages (the ability measure used in this paper).

A. Dropouts

If there are important gender differences when it comes to dropping a course, then the estimates presented above may overestimate or underestimate the full impact of the increased competition on university performance.

²⁸ The class size cutoff was put at 30 students per section, instead of 20, to have a decent number of observations when running the regressions while having a class size small enough not to expect bell curve marking: 11.1% of first-year grades come from sections with fewer than 30 students, while only 4.7% come from sections with fewer than 20 students.

If, for example, the 2003 male students who were more adversely affected by the increased competition dropped out of courses in a disproportionate way relative to their female counterparts, then the findings presented in Section VI.A could overestimate the gender performance differential—the results could be due to selection. In order to investigate this issue, I looked at individual courses, and I constructed a dummy variable equal to one if a course was dropped, and zero otherwise.²⁹ I then estimated a linear probability model, using the same six specifications used to estimate the effect of competition on grades to test whether female and male drop-out decisions were affected differently by the increased competition.

Table 4 shows that males did not drop out more than females when facing increased competition, supporting the idea that the gender differences found above are not due to omitting students who failed to complete courses. The double cohort did increase the percentage of courses being dropped, all else equal: the estimated increase in dropout rates for an “average” student is between 1.2 and 2.4 percentage points.³⁰ This is significant since dropped courses represent about 8% of observed course outcomes in the sample. Overall, results from table 4 do not suggest that the results presented in table 3 are due to gender differences in the dropping out decision process.

B. Participation

The experimental economics literature suggests not only that males perform better when faced with more intense competition than females but also that females might shy away from it.³¹ If a disproportionate number of females avoided the double cohort (e.g., by delaying their university ap-

²⁹ Note that courses for which the final grade is below 30% are also considered as dropped. Adding these grades to the sample analyzed in table 3 would make the estimates for “Male \times Double cohort” slightly larger—the estimates being around 1.2–1.3 points.

³⁰ Although the parameter estimates for “Double cohort” and “High school average \times Double cohort” are not individually statistically significant when looking at G12 and G13 students, the null hypothesis $H_0: \pi + 88\rho = 0$ is rejected at a 10% confidence level under every specification.

³¹ Experimental studies such as Niederle and Vesterlund (2007) and Gupta, Poulsen, and Villevall (2013) find that males are more inclined to participate in competitive activities. Flory, Leibbrandt, and List (2010) suggest that this gender difference also holds in a natural environment. Dohmen and Falk (2011) suggest that risk attitudes play a role in explaining this finding, while Gneezy, Leonard, and List (2009) and Booth and Nolen (2012) suggest that the role of nurture is important in this context. Interestingly, Gneezy et al. (2009) and Andersen et al. (2013) find that females’ and males’ tastes for competition differ when we are looking at a patriarchal or a matrilineal society. Males are more competitive than females in a patriarchal society, while it is not the case in a matrilineal society. See Croson and Gneezy (2009) and Niederle and Vesterlund (2011) for more complete (and more general) discussions of gender differences in preferences.

Table 4
Impact of the Double Cohort on Dropped Courses

	G13 Students Only			G12 and G13 Students		
	(1)	(2)	(3)	(4)	(5)	(6)
High school average	-.007*** (.001)	-.007*** (.001)	-.007*** (.001)	-.007*** (.001)	-.007*** (.001)	-.007*** (.001)
Double cohort	.214** (.100)	.222** (.102)	.218** (.101)	.086 (.079)	.090 (.077)	.089 (.077)
Male	.007 (.008)	1.26e-4 (.006)	-4.05e-4 (.006)	.007 (.008)	4.37e-4 (.006)	7.08e-05 (.006)
High school average × Double cohort	-.002** (.001)	-.002** (.001)	-.002** (.001)	-.001 (.001)	-.001 (.001)	.001 (.001)
Male × Double cohort	-.001 (.008)	.002 (.008)	.002 (.008)	-.006 (.008)	-.003 (.008)	-.004 (.008)
Age			.025*** (.006)			.017*** (.004)
Course fixed effects	No	Yes	Yes	No	Yes	Yes
Number of observations	19,793	19,793	19,793	27,640	27,640	27,640
R ²	.021	.019	.020	.019	.016	.017

NOTE.—The table reports the estimates of regressions of a dummy dependent variable equal to one if the course was dropped, zero otherwise, using a linear probability model. Clustered standard errors (at the course level) are in parentheses.

** Significant at the 5% level.

*** Significant at the 1% level.

plication), and if these females would have been more adversely affected by the increased competition, then the estimates presented in table 3 might underestimate the impact of competition on the performance gap. In order to investigate this possibility, table 5 presents numbers on the proportion of female students by status (applied and enrolled) and by year. Both the proportions of female applicants (panel A) and the proportions of female among enrolled students (panel B) suggest that females did not try to avoid the double cohort more than males. Whether I look at the female proportions among all students or by quartiles (based on high school average), I do not see any economically or statistically significant differences in the proportions of female applicants or enrolled students. These results are not sensitive to the inclusion or exclusion of G12 students. The cost of delaying university enrollment by a year might be large compared to the cost of entering a more competitive environment, which could explain why we do not see changes in the female proportions of applicants and enrolled students between 2001 and 2003.

Figure 1 suggests that 2002 was not a typical year in terms of university applications, which is why I use 2001 (not 2002) as a control year in my regression analysis. Nevertheless, the 2002 data can be used to further investigate whether relatively more females tried to avoid the double cohort. This is not the case: the female proportion among fast-trackers de-

Table 5
Female Participation by Student Status and Year

Student Status	Number of Students		Female Proportion		Difference in Proportions (% ₂₀₀₃ - % ₂₀₀₁)
	2001	2003	2001	2003	
A. Applied:					
Bottom quartile	2,176	4,072	.571	.557	-.014
Second quartile	2,226	4,267	.576	.572	-.004
Third quartile	2,174	4,082	.591	.598	.007
Top quartile	2,169	4,230	.607	.611	.004
Total	8,745	16,651	.586	.585	-.001
B. Enrolled:					
Bottom quartile	661	929	.560	.550	-.011
Second quartile	627	898	.596	.615	.018
Third quartile	660	978	.629	.650	.022
Top quartile	660	901	.658	.644	-.014
Total	2,608	3,706	.611	.615	.004

NOTE.—The table displays the number of grade 13 and grade 12 (combined) students by status (i.e., applied or enrolled) and by year. The quartiles are based on the distributions of student high school averages. The last column presents the difference in female student proportion and results from testing the null hypothesis of equal proportions across years against the alternative hypothesis that they are different. None of these differences are statistically significant at conventional confidence levels.

creased by (a non-statistically significant) 1.9 percentage points between 2001 and 2002, and it stayed virtually unchanged for the general student population. Another way female students could have tried to avoid the increased competition is by enrolling in programs that were “less competitive.” Looking at the proportion of female students per program across time does not suggest that this is cause for concern. The only program for which the proportion of female students changed significantly between 2001 and 2003 is computer science. Unfortunately, a couple of specializations that were associated with computer science in 2001 were associated with life science in 2003, making it impossible to know whether the change in the female proportion was due to the double cohort or the change in program. In any case, this change should not affect my results. This program is the smallest in terms of enrollment—470 students out of the 6,000 observed in this paper are enrolled in computer science—and it has the smallest proportion of female students (around 23%).

C. Preuniversity Competition

Results presented so far assume that a student’s high school grade is a good indicator of academic ability. Since students knew (since 1997) that university admissions would be more competitive in 2003 than in previous years, it is possible that they reacted by studying more in high school. The competition level would therefore not only increase at the university but also prior to attendance at the university. In this case, the link between

academic ability and high school grades might have changed between 2001 and 2003, representing a potential source of bias for the estimator. Two scenarios must be considered.

First, assume that competition was intense at the high school level, which stimulated males more than females, as suggested by the experimental economics literature, resulting in males outperforming females with similar academic ability. Consequently, 2003 male academic ability could be overestimated by the male high school average, which would then translate into underestimating the impact of increased competition on the gender university-performance gap. This situation should not be a major concern, since I already find a positive effect, and it would only make it larger.

The second scenario would be that females were actually more stimulated by the increased high school competition than males. This seems unlikely given that table 3 suggests that males increased their university performance relative to females in the more competitive year. Nevertheless, it is still possible in principle that females were more stimulated by competition both in high school and at the university. For that to be true, and to find a positive coefficient to “Male \times Double cohort”—as I do—females’ high school performance would have to have improved significantly more than their university performance. The 2003 female high school average would have to severely overestimate ability. In this case, the 2003 high school average could overestimate females’ ability, which would result in overestimating the impact of increased competition on the gender university-performance gap.

I investigate this possibility by reestimating the regressions presented in table 3 but using only students whose ability measurement is not likely to have been affected by the double cohort (i.e., out-of-province students and students that attended a university prior to the University of Toronto). The results are very similar to the ones using Ontario secondary school students. In particular, the coefficient estimates for “Male \times Double cohort,” although imprecise, fluctuate between 0.70 (when not controlling for course fixed effects or age) and 3.2 (when controlling for course fixed effects, but not for age). This finding supports the idea that the estimates presented in table 3 are not driven by Ontario female student academic ability being significantly overestimated in 2003.³²

³² As an additional robustness check, I reestimated the regressions presented in table 3 on a pooled sample of Ontario and out-of-province students and students who attended a university prior to the University of Toronto. Using a full set of dummy variables and interaction terms (i.e., using a triple-difference estimation approach), I test whether the academic ability of Ontario female students is significantly overestimated in 2003 (relative to Ontario male students). If this is the case, then the coefficient estimate of an interaction term between an Ontario sec-

VIII. After the First Year

Results from table 3 suggest that the increased competition following the Ontario double cohort had an impact on the gender gap in first-year university performance. I now show that the double cohort actually affected the performance gap during most of these students' undergraduate years.

Table 6 presents estimates of the competition effect on the gender performance gap for students' first (year 1) to fourth year (year 4) at the university. Estimates next to "Year 1" are taken from table 3. The estimation strategy is exactly the same as the one used above except that the dependent variables are students' upper-year course grades. Aside from a small drop for students' third year, the effect of increased competition is surprisingly stable across years. The estimated coefficients using students' fourth-year grades vary between 0.75 and 1.08 when controlling for course fixed effects, being close to the ones obtained using students' first-year grades. Despite being of modest size initially, the competition effect could accumulate over time, affecting the gender gap in attrition rates, in credit accumulation, and, ultimately, in on-time graduation rates. I now turn to these.

A. Attrition

From year to year, the sample size of observed students decreases, as can be seen in table 6. Some students will change their program, change their university, or simply quit. If males and females differ in their decision to stay in school (or in the faculty), the estimates found in table 6 could misrepresent the evolution of the competition effect. In particular, if some females were forced to quit school due to bad performances in their first year, then the results from table 6 might be underestimating the effect of competition on the gender performance gap. Table 7 investigates this possibility. Using an estimation strategy similar to the one above, I regress (using a linear probability model) a dummy variable equal to one if the student dropped out of the sample after 1, 2, or 3 years, respectively. Eighteen estimates of the "Male \times Double cohort" parameter are presented in table 7. Most of the estimates are small and all are statistically insignificant. Some estimates (e.g., based on G13 students observed after 2 years), although not statistically significant, are not small when compared to the attrition rate. Note that these (larger but imprecise) estimates are all negative, suggesting, if anything, that males might have dropped out of the sample

ondary school dummy (say "ON") and "Male \times Double cohort" would be significantly different from zero. The coefficient estimate of this interaction term, although imprecise, is very close to zero in all specifications. This further suggests that the academic ability of Ontario female students is not significantly overestimated in 2003.

Table 6
Evolution of the Impact of the Double Cohort on the Gender Performance Gap

	G13 Students Only			G12 and G13 Students		
	(1)	(2)	(3)	(4)	(5)	(6)
(a) Year 1	1.061*** (.362)	1.019*** (.348)	1.009*** (.345)	1.042*** (.341)	1.014*** (.342)	1.015*** (.341)
Number of students	4,298	4,298	4,298	6,000	6,000	6,000
(b) Year 2	1.036*** (.340)	1.035*** (.313)	1.003*** (.310)	.950*** (.297)	.974*** (.267)	.985*** (.265)
Number of students	3,908	3,908	3,908	5,466	5,466	5,466
(c) Year 3	.801*** (.309)	.821*** (.311)	.809*** (.311)	.918*** (.287)	.884*** (.285)	.893*** (.286)
Number of students	3,409	3,409	3,409	4,847	4,847	4,847
(d) Year 4	1.774*** (.303)	1.055*** (.285)	1.077*** (.284)	1.260*** (.268)	.753*** (.250)	.785*** (.249)
Number of students	3,146	3,146	3,146	4,474	4,474	4,474
Course fixed effects	No	Yes	Yes	No	Yes	Yes
Age control	No	No	Yes	No	No	Yes

NOTE.—The table reports 24 coefficient estimates of “Male \times Double cohort” for university years 1–4. The dependent variable in these sets of regressions is a university grade average. Each estimate comes from a separate regression. The estimates next to Year 1 are from table 3. Clustered standard errors (at the course level) are in parentheses.
*** Significant at the 1% level.

Table 7
Impact of the Double Cohort on Attrition Rates

Dependent Variable: Dropped from Sample	G13 Students Only			G12 and G13 Students		
	(1)	(2)	(3)	(4)	(5)	(6)
(a) After 1 year	-.015 (.018)	-.014 (.018)	-.014 (.018)	-.007 (.016)	-.006 (.016)	-.006 (.016)
Attrition rate	.091	.091	.091	.089	.089	.089
(b) After 2 years	-.021 (.025)	-.024 (.025)	-.023 (.025)	-.007 (.021)	-.007 (.021)	-.007 (.021)
Attrition rate	.207	.207	.207	.192	.192	.192
(c) After 3 Years	-.003 (.028)	-2.89e-4 (.028)	2.83e-4 (.028)	.004 (.024)	.009 (.024)	.009 (.024)
Attrition rate	.268	.268	.268	.254	.254	.254
Number of observations	4,298	4,298	4,298	6,000	6,000	6,000
Program fixed effects	No	Yes	Yes	No	Yes	Yes
Age control	No	No	Yes	No	No	Yes

NOTE.—The table presents 18 coefficient estimates for “Male \times Double cohort” capturing the effect of competition on the gender gap in the probability to drop out of the sample (used in table A1). These estimates are obtained using a set of dummy dependent variables “Dropped from sample after X years” equal to one, for a specific student, if I do not observe any grades for this student after X years (in university). These dummy variables are regressed on the same regressors and under the same six specifications used in table A1. The estimation was done using a linear probability model. Robust standard errors are in parentheses. None of the parameter estimates are statistically significant at conventional confidence levels.

less than females. Overall, these results suggest that dropouts are not driving the results in table 6.

B. Credit Accumulation

While not strong enough to affect important decisions like dropping out of university, the competition effect might have been strong enough to slow down students and affect credit accumulation and therefore the probability that they graduate on time. In order to better measure the magnitude of the long-lasting effect of competition on the gender performance gap, I now look at student credit accumulation. Table 8 presents coefficient estimates for the effect of competition on the gender differential in credit accumulation. The number of credits accumulated after X years includes the number of credits earned during year X but also the number of credits earned over the previous years. Therefore, the estimated effects presented in table 8 should be seen as cumulative effects.

The results in the top panel of table 8 suggest that if I look at all students, the effect of the increased competition on the gender performance gap was not strong enough to significantly affect the number of earned credits. The estimates for “Male \times Double cohort” seem to increase in magnitude as we look at the effect after 3 and 4 years, though the standard errors also become larger. The results are different if I concentrate on students with a high school average of 88% and below—students who

Table 8
Impact of the Double Cohort on Credit Accumulation

Dependent Variable: Accumulated Credits	G13 Students Only			G12 and G13 Students		
	(1)	(2)	(3)	(4)	(5)	(6)
A. All Students:						
(a) After 1 year	.048 (.062)	.001 (.054)	3.47e-4 (.054)	.048 (.052)	-.003 (.045)	-.003 (.045)
(b) After 2 years	.273** (.131)	.174 (.124)	.171 (.124)	.175 (.111)	.069 (.105)	.069 (.105)
(c) After 3 years	.412** (.207)	.304 (.202)	.298 (.202)	.266 (.173)	.139 (.167)	.139 (.167)
(d) After 4 years	.419 (.296)	.265 (.290)	.256 (.290)	.284 (.249)	.102 (.242)	.101 (.242)
Number of observations	4,298	4,298	4,298	6,000	6,000	6,000
B. Below-average students:						
(a) After 1 year	.063 (.086)	.046 (.082)	.046 (.082)	.015 (.073)	-.009 (.069)	-.009 (.069)
(b) After 2 years	.485*** (.179)	.444** (.173)	.444** (.173)	.288* (.152)	.230 (.146)	.232 (.146)
(c) After 3 Years	.885*** (.276)	.841*** (.272)	.841*** (.272)	.571** (.232)	.499*** (.226)	.506** (.226)
(d) After 4 Years	1.020*** (.373)	.959*** (.367)	.959*** (.367)	.660** (.312)	.562* (.306)	.574* (.306)
Number of observations	2,687	2,687	2,687	3,506	3,506	3,506
Program fixed effects	No	Yes	Yes	No	Yes	Yes
Age control	No	No	Yes	No	No	Yes

NOTE.—The table presents 48 coefficient estimates for “Male \times Double cohort” capturing the effect of competition on the gender gap in the number of credits accumulated after X years (in university). The number of credits accumulated after X years is regressed on the same regressors and under the same six specifications used in table A1. The number of credits earned is zero for year X ($X = 2, 3$, or 4) if the student is not observed in year X . Robust standard errors are in parentheses.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

are more likely to fail or drop a course. The bottom panel of table 8 clearly suggests that the double cohort affected the gender differential in credit accumulation among below-average students. The estimates for “Male \times Double cohort” increase steadily as students progress toward graduation. By the end of their fourth year, male students will have accumulated between 0.6 and 1.0 credits (more than a “standard” 1-semester course) than similar female classmates as a result of the increased competition.

C. Graduation

One final aspect of the academic performance that I consider is on-time graduation. As the effect of increased competition seems to be long-lasting, it could affect the probability that a student graduates on time (i.e., not later than during the summer following his/her fourth year). In table 9, I present results from regressing an “On-time-graduation” dummy variable on

Table 9
Impact of the Double Cohort on On-Time Graduation Rates

	G13 Students Only			G12 and G13 Students		
	(1)	(2)	(3)	(4)	(5)	(6)
High school average	.023*** (.002)	.026*** (.002)	.026*** (.002)	.023*** (.002)	.025*** (.002)	.025*** (.002)
Double cohort	.506* (.265)	-.357 (.267)	-.351 (.267)	-.563** (.224)	-.353 (.225)	-.355 (.225)
Male	-.087*** (.019)	-.066*** (.020)	-.065*** (.020)	-.087*** (.019)	-.067*** (.020)	-.066*** (.020)
High school average × Double cohort	.005 (.003)	.003 (.003)	.003 (.003)	.005** (.003)	.003 (.003)	.003 (.003)
Male × Double cohort	.081*** (.030)	.070** (.030)	.070** (.030)	.074*** (.025)	.064** (.025)	.064** (.025)
Computer science		-.238*** (.030)	-.241*** (.030)		-.275*** (.027)	-.277*** (.026)
Humanities		-.088*** (.024)	-.089*** (.024)		-.129*** (.021)	-.128*** (.021)
Life science		-.147*** (.026)	-.148*** (.026)		-.164*** (.022)	-.165*** (.022)
Age			-.034 (.026)			-.031** (.013)
Constant	-1.542*** (.166)	-1.677*** (.174)	-1.035** (.527)	-1.542*** (.166)	-1.602*** (.172)	-1.009*** (.300)
Number of observations	4,298	4,298	4,298	6,000	6,000	6,000
R ²	.063	.077	.078	.065	.081	.082

NOTE.—The table reports the estimates of regressions of a dummy dependent variable equal to one if the student graduated by July 1 of his/her fourth university year, and zero otherwise. The estimation was done using a linear probability model. Robust standard errors are in parentheses.

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

similar control variables used in the previous regressions.³³ The “On-time-graduation” variable is equal to one if the student has graduated from university by July 1 of his/her fourth year. Unfortunately, the data do not allow me to observe 2003 students after their fourth year. Estimates of the double cohort effect on the gender difference in the probability of graduating on time are statistically significant and large. They range from 6.4 to 8.1 percentage points. To put the size of the estimates in perspective, about 40% of the students observed in table 3 graduated on time. Prior to the double cohort, male students were less likely to graduate on time than females, the difference being around 7–9 percentage points. This difference completely vanished for double cohort students.³⁴

The estimates presented here must be interpreted with caution. As the abolition of grade 13 affected university admission standards in 2003, so it could affect the admission standards for graduate school in 2007. Hence, while part of the effect estimated here could be due to performance, it is also possible that these estimates are capturing the effect of strategic behavior. Some students may have been tempted to delay graduate school applications by a year. As mentioned earlier, females may not “embrace” competition as much as males. If graduate school-bound female students delayed (more than males) their graduation to avoid the double cohort, part of the competition effect presented above would be due to this behavior. Note that this behavior is not likely to explain all of the effect, since only a fraction of undergraduate students will apply to graduate school. In order to have an idea of how much of the estimates found above could be due to strategic behavior, I estimated similar regressions as those presented in table 9 excluding students with a high school average above 88% to concentrate on students who are less likely to apply to graduate school. The “Male \times Double cohort” coefficient estimates are actually larger for “lower-ability” students, as is the case for earned credits, suggesting that the results presented in table 9 are not due to such strategic behavior.³⁵ Interestingly, changes in unconditional on-time graduation proportions between 2001 and 2003 for females and males show that males’ on-time graduation rate increased significantly (from 33.6% to 38.1%) while females’ decreased slightly (43.8%–41.4%). While comparison of changes in unconditional on-time graduation rates

³³ The results presented in table 9 are from linear probability models. The results from estimating Probit models are very similar.

³⁴ The null hypothesis $H_0: \delta + \beta = 0$ (i.e., that the sum of the “Male” and “Male \times Double cohort” coefficients is equal to zero) is not rejected at conventional confidence levels in any of the six regressions. Note that the effect of the double cohort is not statistically significant based on the “Double cohort” and “High school average \times Double cohort” coefficient estimates. Once I drop the “High school average \times Double cohort” variable from the regression, the “Double cohort” coefficient estimate changes to around -0.1 and becomes statistically significant at 1%. The “Male \times Double cohort” coefficient estimate remains unchanged at around 0.07 . These results are not presented here but are available upon request.

³⁵ Detailed results for higher- and lower-ability students are available upon request.

suggest that males' performance increased while females' did not change, comparison of changes in conditional graduation rates (results from table 9) suggest that the females' graduation rate was significantly (negatively) affected by the increased competition while the males' graduation rate was not. Overall, results from table 9 suggest that there is a clear difference in the reaction to increased competition by females and males and that results from table 6 are not due to male students taking more time to graduate.

IX. Conclusion

This paper has provided new empirical evidence of gender performance differences in response to increased competition, focusing on long-term tasks carried out in a regular social environment: the university classroom. The exogenous increase in competition generated by the abolition of grade 13 allowed me to assess whether the findings from the experimental economics literature carry over to tasks that a large fraction of individuals of a given age perform over an extended period and have strong incentives to care about. As such, this helps to link the experimental evidence to competitive gender differences in important life outcomes, including those in the labor market.

My results indicate that gender differences in performance under increased competition do carry over to an important nonexperimental setting: performance while at the university. The estimated effects are modest in size but are precisely estimated, and they appear to persist: the increased competition had an economically significant positive impact on males' on-time graduation rates, especially for below-average students. The evidence also supports the view that the results are due to differential changes in effort by gender, rather than self-selection, as indicated by findings regarding the decisions to apply to university, enroll if accepted, and drop out of courses/programs. (The modest nature of the differential effects on performance by gender helps explain why the self-selection effects are muted.)

While the results point to differential effort changes by gender, the nature of the task studied in this paper also allows for competition to affect performance through a peer channel—an important feature of workplace environments.³⁶ The double cohort, combined with the limited capacity of universities to expand rapidly, increased the quality of admitted students, and so of peers. Hence, the estimates presented here are likely to capture the combined (net) effect of two potentially significant mechanisms by which competition could explain stylized facts about the labor market. If female students benefit more from higher-quality peers, as suggested by recent evidence in Stinebrickner and Stinebrickner (2006), then my results are more

³⁶ See Epple and Romano (2011) and Sacerdote (2011) for recent literature reviews on peer effects in postsecondary education.

surprising, as they would represent conservative estimates of the impact of increased competition on the male-female performance differential. Still finding a significant increased-competition impact suggests that the “pure” competition effect due to changes in effort could be more important than the peer effect in explaining gender differences in performance.

The results from this study have several implications. First, they indicate that male and female performance responds differently to an increase in competition for regular everyday tasks that are long-lasting and for which the outcomes matter to participants. This underlines the potential for competition to explain some of the gender inequality found in the labor market. The results also have implications for the delivery of education (e.g., single-sex vs. coed schooling) and for student achievement assessment. In terms of the former, my results suggest that the optimal level of competition associated with classroom activities for student learning (or effort) may be different for females and males and that one should take this finding into account when weighing the pros and cons of single-sex education. One advantage of single-sex classrooms (or schools) would be to allow teachers to adjust the competition level according to the gender of the student body.

In terms of the latter, the results suggest that performance assessment measures (e.g., bell curve marking) found in many universities may not be gender-neutral: a given performance measure can favor females (or males), depending on the level of competition inside the classroom. Such a finding naturally raises the question of whether we should expect similar results for high school students who will not go to college—for potential high school dropouts in particular. If so, the performance assessment measures could be made more “competitive” by grouping males (who represent the majority of high school dropouts) together and using tournament-based evaluations or within-class ability streaming, without negatively affecting female students.

For future work, it would be interesting to investigate the effects of competition in university on gender differences in performance once in the labor market—beyond the scope of the data I analyze in this study. Doing so would speak even more directly to the relative importance of gender competition versus discrimination and occupational self-selection in explaining important gender differences in the labor market, referred to in the Introduction. It would also be interesting to explore the strength of gender differences in tasks where absolute, rather than relative, performance can be measured. This would help clarify whether, on the one hand, the cost of effort rose less for males relative to females, or on the other hand, it fell by more in more competitive environments. One could also examine possible gender effects for tasks where the role of peers is more rather than less important, building on the analysis in Stinebrickner and Stinebrickner (2006).

Appendix A
Alternative Performance Measures

Table A1
Impact of the Double Cohort on the Gender Performance Gap (First-Year Average)

	G13 Students Only			G12 and G13 Students		
	(1)	(2)	(3)	(4)	(5)	(6)
High school average	1.112*** (.030)	1.165*** (.031)	1.166*** (.031)	1.112*** (.030)	1.147*** (.031)	1.148*** (.031)
Double cohort	-11.08*** (4.446)	-12.25*** (4.483)	-12.11*** (4.485)	-17.44*** (3.709)	-18.34*** (3.745)	-18.36*** (3.744)
Male	.703** (.317)	.927*** (.324)	.948*** (.324)	.703** (.317)	.851*** (.322)	.859*** (.322)
High school average \times Double cohort	.104** (.051)	.115** (.051)	.113** (.051)	.176*** (.042)	.185*** (.043)	.184*** (.043)
Male \times Double cohort	1.131** (.489)	1.058** (.490)	1.045** (.489)	1.070** (.419)	1.027** (.419)	1.026** (.419)
Computer science		-.970* (.568)	-1.023* (.570)		-.617 (.503)	-.630 (.503)
Humanities		.709* (.369)	.701* (.369)		.493 (.313)	.503 (.313)
Life science		-.928** (.410)	-.948** (.410)		-.599* (.341)	-.606* (.341)
Age			-.721* (.426)			-.255 (.214)
Constant	-25.62*** (2.601)	-30.18*** (2.692)	-16.42* (8.558)	-25.62*** (2.601)	-28.66*** (2.666)	-23.83*** (4.838)
Number of observations	4,298	4,298	4,298	6,000	6,000	6,000
R ²	.334	.340	.340	.346	.348	.348

NOTE.—The dependent variable in these regressions is the student's credit-weighted first-year university average (excluding mathematics credits). Robust standard errors are in parentheses.
 * Significant at the 10% level.
 ** Significant at the 5% level.
 *** Significant at the 1% level.

Table A2
Impact of the Double Cohort on the Gender Performance Gap (First-Year GPA)

	G13 Students Only			G12 and G13 Students		
	(1)	(2)	(3)	(4)	(5)	(6)
High school average	.100*** (.003)	.104*** (.003)	.104*** (.003)	.100*** (.003)	.103*** (.003)	.103*** (.003)
Double cohort	-.933*** (.390)	-1.052*** (.393)	-1.040*** (.393)	-1.369*** (.323)	-1.456*** (.325)	-1.457*** (.325)
Male	.060** (.029)	.086*** (.030)	.087*** (.030)	.060*** (.029)	.080*** (.029)	.080*** (.029)
High school average × Double cohort	.009** (.004)	.010** (.004)	.010** (.004)	.014*** (.004)	.014*** (.004)	.014*** (.004)
Male × Double cohort	.101** (.045)	.093** (.045)	.092** (.045)	.091** (.038)	.084** (.038)	.084** (.038)
Computer science		-.105** (.051)	.110** (.052)		-.083* (.045)	-.083* (.045)
Humanities		.083** (.034)	.082** (.034)		.060** (.029)	.061** (.029)
Life science		-.072* (.037)	-.073** (.037)		-.045 (.031)	-.046 (.031)
Age			-.062 (.039)		-.014 (.019)	-.014 (.019)
Constant	-6.006*** (.228)	-6.432*** (.237)	-5.253*** (.777)	-6.006*** (.228)	-6.296*** (.234)	-6.034*** (.434)
Number of observations	4,298	4,298	4,298	6,000	6,000	6,000
R ²	.323	.330	.331	.334	.337	.337

NOTE.—The dependent variable in these sets of regressions is the student's first-year university grade point average (excluding mathematics credits). Robust standard errors are in parentheses.
 * Significant at the 10% level.
 ** Significant at the 5% level.
 *** Significant at the 1% level.

Table A3
Double Cohort and Gender Performance Gap in Smaller Sections

	G13 Students Only			G12 and G13 Students		
	(1)	(2)	(3)	(4)	(5)	(6)
High school average	.804*** (.046)	.855*** (.056)	.855*** (.056)	.804*** (.046)	.857*** (.056)	.857*** (.056)
Double cohort	3.662 (5.025)	8.902* (4.537)	8.755* (4.516)	2.860 (4.902)	7.837** (3.708)	7.922** (3.721)
Male	1.354** (.558)	.558 (.448)	.547 (.449)	1.354** (.558)	.600 (.452)	.592 (.454)
High school average × Double cohort	−.046 (.055)	−.113** (.051)	−.111** (.051)	−.039 (.057)	−.100** (.043)	−.099** (.044)
Male × Double cohort	−.048 (.620)	.261 (.509)	.271 (.510)	−.166 (.519)	.167 (.385)	.168 (.385)
Age			.643** (.316)			.454 (.297)
Course fixed effects	No	Yes	Yes	No	Yes	Yes
Number of observations	2,343	2,343	2,343	3,180	3,180	3,180
Number of students	1,964	1,964	1,964	2,691	2,691	2,691
R ²	.171	.188	.188	.165	.180	.181

NOTE.—The dependent variable in these sets of regressions is a first-year university course grade. A course section is labeled as “small” if it contains fewer than 30 students. Clustered standard errors (at the course level) are in parentheses.

- * Significant at the 10% level.
- ** Significant at the 5% level.
- *** Significant at the 1% level.

Appendix B
Competition and Effort

Here I develop some details regarding the model presented in Section III.B. Using equation (4), we can see that as the level of competition increases, effort will increase or decrease depending on whether students enjoy or dislike competition and on how the university changes the link between knowledge and grades:

$$\frac{de_g^*}{dl} = -\frac{\frac{d\lambda_l}{dl}\frac{\partial K(e_g^*, a)}{\partial e}}{\lambda_l\frac{\partial^2 K(e_g^*, a)}{\partial e^2}} - \frac{\frac{\partial^2 C_g(e_g^*, a, l)}{\partial e\partial l}}{\frac{\partial^2 C_g(e_g^*, a, l)}{\partial e^2}}.$$

(B1)

The gender difference in the change in effort following an increase in the level of competition is given by:

$$\Delta \frac{de^*}{dl} \Big|_{l=0} = \frac{\frac{\partial^2 C_m(e_m^*, a, 0)}{\partial e \partial l} - \frac{d\lambda_l}{dl} \frac{\partial K(e_m^*, a)}{\partial e} \Big|_{l=0}}{\lambda_0 \frac{\partial^2 K(e_m^*, a)}{\partial e^2} \Big|_{l=0} - \frac{\partial^2 C_m(e_m^*, a, 0)}{\partial e^2}} - \frac{\frac{\partial^2 C_f(e_f^*, a, 0)}{\partial e \partial l} - \frac{d\lambda_l}{dl} \frac{\partial K(e_f^*, a)}{\partial e} \Big|_{l=0}}{\lambda_0 \frac{\partial^2 K(e_f^*, a)}{\partial e^2} \Big|_{l=0} - \frac{\partial^2 C_f(e_f^*, a, 0)}{\partial e^2}}. \quad (B2)$$

In the absence of competition (at $l = 0$), females and males are assumed to have the same knowledge production and cost functions, and therefore exert the same level of effort, for a given level of ability (a). Hence, we can simplify equation (B2):

$$\Delta \frac{de^*}{dl} \Big|_{l=0} = - \frac{\frac{\partial^2 C_f(e^*, a, 0)}{\partial e \partial l} - \frac{\partial^2 C_m(e^*, a, 0)}{\partial e \partial l}}{\lambda_0 \frac{\partial^2 K(e^*, a)}{\partial e^2} \Big|_{l=0} - \frac{\partial^2 C(e^*, a, 0)}{\partial e^2} \Big|_{l=0}}. \quad (B3)$$

The concavity of the knowledge production function and the convexity of the cost function guarantee that the denominator of equation (B3) is strictly negative. Therefore, the sign of equation (B3) will be determined by whether males enjoy competition more than females. If males enjoy competition more (i.e., if $\partial C_f^2(e, a, l)/\partial e \partial l > \partial C_m^2(e, a, l)/\partial e \partial l$), we would expect to see their effort level and academic performance increase relative to females as competition increases.

1. Competition and Effort Incentives (λ_i)

In this section, I illustrate the potential impact of increased competition and bell curve marking on effort incentives. I abstract from potential gender differences in response to competition to concentrate on the effect of competition on the effort incentives set by a university that grades on a bell curve.

Imagine a student with ability a will exert effort e to acquire some knowledge, $K(e, a)$. Assume that $K(e, a)$ is continuously differentiable, increasing and concave in a and e , and that the marginal benefit of effort increases with ability (i.e., $\partial^2 K(e, a)/\partial e \partial a > 0$). The cost of studying is given by a continuously differentiable function $C(e, a)$ (in e and a) and increasing and convex in e . Further, assume that the marginal cost of effort is decreasing with ability (i.e., $\partial^2 C(e, a)/\partial e \partial a < 0$).

Students expect the university to mark on a bell curve. More specifically, they expect that, every year, the average grade in university will be

$$\mu = \eta_t + \lambda_t E(K(e, a)|t), \quad (\text{B4})$$

where $E(K(e, a)|t)$ is the average knowledge of students enrolled in year t . The variance of university grades is also expected to remain constant across time:

$$\sigma^2 = \lambda_t^2 \text{Var}(K(e, a)|t). \quad (\text{B5})$$

Students take the mean and the variance of knowledge ($E(K(e, a)|t)$ and $\text{Var}(K(e, a)|t)$) as given (i.e., they do not think they can individually affect the average performance or its variance). We can rewrite equation (B5) to see what will affect λ_t :

$$\lambda_t^2 = \frac{\sigma^2}{\text{Var}(K(e, a)|t)}.$$

Therefore, λ_t is expected to increase as $\text{Var}(K(e, a)|t)$ decreases. In order to see what happens to $\text{Var}(K(e, a)|t)$ (and λ_t) as competition increases, we must go back to the student utility maximization problem.

Student utility is simply defined by the difference between his/her university grade and the effort he/she exerted. So his/her utility maximization problem is

$$\max_e \eta_t + \lambda_t K(e, a) - C(e, a).$$

The first-order condition for this problem is

$$\lambda_t \frac{\partial K(e^*, a)}{\partial e} - \frac{\partial C(e^*, a)}{\partial e} = 0. \quad (\text{B6})$$

We can see that student effort will increase as λ_t increases. Importantly, we can use equation (B6) to show that

$$\frac{de^*}{da} = \frac{\lambda_t \frac{\partial^2 K(e^*, a)}{\partial e \partial a} - \frac{\partial^2 C(e^*, a)}{\partial e \partial a}}{\frac{\partial^2 C(e^*, a)}{\partial e^2} - \lambda_t \frac{\partial^2 K(e^*, a)}{\partial e^2}} > 0. \quad (\text{B7})$$

That is, higher-ability students will exert more effort. Since de^*/da is strictly positive and $C(e, a)$ and $K(e, a)$ are continuously differentiable, there is a one-to-one correspondence between a and e . We can therefore define

$$H(a) \equiv K(e^*(a), a).$$

Now we can look at what happens when competition for university admissions increases from one year to the other. It is easy to imagine that

if high school grades are determined in a similar fashion to university grades, the ability threshold that determines whether a student is accepted at the university will go up during the more competitive year. Therefore, as competition for admission increases, the variance of university-student ability should decrease while their average ability will increase as universities become more selective—the ability distribution is being truncated from below.

We can see what will happen to the variance of $H(a)$ if the variance of a decreases, using a Taylor approximation:

$$\text{Var}[H(a)|t] \approx [H'(E(a)|t)]^2 + \text{Var}(a|t). \quad (\text{B8})$$

Under mild conditions $H(a)$ is concave.³⁷ If $H(a)$ is concave, then as competition increases, $H'(E(a)|t)$ will decrease and so will $\text{Var}[H(a)|t]$. This means that the variance of knowledge $\text{Var}(K(e, a)|t)$ will be smaller and λ_i larger (due to eq. [B5]) in a competitive year, and consequently the optimal level of effort for a given a , $e^*(a)$ will be higher in a competitive year (due to eq. [B6]). This discussion indicates that although effort is not observable here, one could imagine that the increased benefit to effort is one important channel through which competition affects students' performance.

2. Concavity of $H(a)$

We know that $H(a)$ is increasing in a :

$$\frac{\partial H(a)}{\partial a} = \frac{\partial K(e(a), a)}{\partial a} + \frac{\partial K(e(a), a)}{\partial e(a)} \frac{de(a)}{da} > 0.$$

The second derivative of $H(a)$ is given by

$$\begin{aligned} \frac{\partial^2 H(a)}{\partial a^2} &= \frac{\partial^2 K(e(a), a)}{\partial a^2} + 2 \frac{\partial^2 K(e(a), a)}{\partial a \partial e(a)} \frac{de(a)}{da} \\ &\quad + \frac{\partial^2 K(e(a), a)}{\partial e(a)^2} \left(\frac{de(a)}{da} \right)^2 + \frac{\partial K(e(a), a)}{\partial e(a)} \frac{d^2 e(a)}{da^2}. \end{aligned} \quad (\text{B9})$$

The first and third terms of equation (B9) are negative, while the second term is positive. In order for the sum of these three terms to be negative, the “complementarity” between ability and effort must not be too large in

³⁷ Loosely speaking, $H(a)$ will be concave if the concavity of the knowledge production function in effort and ability dominates the complementarity of these two inputs. The exact conditions for the concavity of $H(a)$ are presented in app. B.2.

the production of knowledge. That is, the concavity of the knowledge production function must dominate the complementarity effect. A similar condition is sufficient for $d^2e(a)/da^2$ to be negative. To simplify the exposition, I rewrite the first-order condition as

$$\frac{\partial R(e^*(a), a)}{\partial e} = \lambda_t \frac{\partial K(e^*, a)}{\partial e} - \frac{\partial C(e^*, a)}{\partial e}. \quad (\text{B10})$$

Now, from equation (B7), we have

$$\frac{de^*}{da} = - \frac{\frac{\partial^2 R(e(a), a)}{\partial e \partial a}}{\frac{\partial^2 R(e(a), a)}{\partial e^2}}. \quad (\text{B11})$$

So,

$$\begin{aligned} \frac{d^2e^*}{da^2} &= - \frac{\left(\frac{\partial^2 R(e(a), a)}{\partial e^2} \right) \left(\frac{\partial^3 R(e(a), a)}{\partial e \partial a^2} + \frac{\partial^3 R(e(a), a)}{\partial e^2 \partial a} \frac{de^*}{da} \right) + \left(\frac{\partial^2 R(e(a), a)}{\partial e \partial a} \right) \left(\frac{\partial^3 R(e(a), a)}{\partial e^2 \partial a} + \frac{\partial^3 R(e(a), a)}{\partial e^3} \frac{de^*}{da} \right)}{\left(\frac{\partial^2 R(e(a), a)}{\partial e^2} \right)^2} \\ &\equiv \frac{A}{B} \end{aligned}$$

Since B is always positive, the sign of d^2e^*/da^2 will be determined by the sign of A. Using equation (B11), we can rewrite A as

$$\begin{aligned} A &= - \frac{\partial^2 R(e(a), a)}{\partial e^2} \frac{\partial^3 R(e(a), a)}{\partial e \partial a^2} + 2 \frac{\partial^3 R(e(a), a)}{\partial e^2 \partial a} \frac{\partial^2 R(e(a), a)}{\partial e \partial a} \\ &\quad - \left(\frac{\partial^2 R(e(a), a)}{\partial e \partial a} \right)^2 \left(\frac{\frac{\partial^3 R(e(a), a)}{\partial e^3}}{\frac{\partial^2 R(e(a), a)}{\partial e^2}} \right). \end{aligned}$$

In order to guarantee A to be negative, the complementarity between a and e in creating knowledge must decrease with effort and ability, and the third derivative of R with respect to effort must be negative. That is, a sufficient (but not necessary) condition is:

$$\begin{aligned} \frac{\partial^3 R(e(a), a)}{\partial e^3} &< 0, \\ \frac{\partial^3 R(e(a), a)}{\partial e \partial a^2} &< 0, \\ \frac{\partial^3 R(e(a), a)}{\partial e^2 \partial a} &< 0. \end{aligned}$$

References

- Andersen, Steffen, Seda Ertac, Uri Gneezy, John A. List, and Sandra Maximiano. 2013. Gender, competitiveness and socialization at a young age: Evidence from a matrilineal and a patriarchal society. *Review of Economics and Statistics* 95, no. 4:1438–43.
- Angrist, Joshua D., Daniel Lang, and Philip Oreopoulos. 2009. Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics* 1, no. 1:136–63.
- Bertrand, Marianne, Claudia Goldin, and Lawrence F. Katz. 2010. Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American Economic Journal: Applied Economics* 2, no. 3:228–55.
- Bertrand, Marianne, and Kevin F. Hallock. 2001. The gender gap in top corporate jobs. *Industrial and Labor Relations Review* 55, no. 1:3–21.
- Booth, Alison, and Patrick Nolen. 2012. Choosing to compete: How different are girls and boys? *Journal of Economic Behavior and Organization* 81, no. 2:542–55.
- Calsamiglia, Caterina, Jörg Franke, and Pedro Rey-Biel. 2013. The incentive effects of affirmative action in a real-effort tournament. *Journal of Public Economics* 98:15–31.
- Croson, Rachel, and Uri Gneezy. 2009. Gender differences in preferences. *Journal of Economic Literature* 47, no. 2:448–74.
- Dohmen, Thomas, and Armin Falk. 2011. Performance pay and multi-dimensional sorting: Productivity, preferences, and gender. *American Economic Review* 101, no. 2:556–90.
- Ellison, Glenn, and Ashley Swanson. 2010. The gender gap in secondary school mathematics at high achievement levels: Evidence from the American Mathematics Competitions. *Journal of Economic Perspectives* 24, no. 2:109–28.
- Epple, Dennis, and Richard E. Romano. 2011. Peer effects in education: A survey of the theory and evidence. In *Handbook of social economics*, vol. 1B, ed. Jess Benhabib, Matthew O. Jackson, and Alberto Bisin. Amsterdam: North-Holland.
- Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. Unconditional quantile regressions. *Econometrica* 77, no. 3:953–73.
- Flory, Jeffrey A., Andreas Leibbrandt, and John A. List. 2010. Do competitive work places deter female workers? A large-scale natural field experiment on gender differences in job-entry decisions. NBER Working Paper no. 16546, National Bureau of Economic Research, Cambridge, MA.
- Fryer, Roland G., and Steven D. Levitt. 2010. An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics* 2, no. 2:210–40.

- Gneezy, Uri, Kenneth L. Leonard, and John A. List. 2009. Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica* 77, no. 5:1637–64.
- Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. Performance in competitive environments: Gender differences. *Quarterly Journal of Economics* 118, no. 3:1049–74.
- Gneezy, Uri, and Aldo Rustichini. 2004. Gender and competition at a young age. *American Economic Review* 94, no. 2:377–81.
- Gupta, Nabanita Datta, Anders Poulsen, and Marie Claire Villeval. 2013. Gender matching and competitiveness: Experimental evidence. *Economic Inquiry* 51, no. 1:816–35.
- Hoffmann, Florian, and Philip Oreopoulos. 2009a. A professor like me: The influence of instructor gender on college achievement. *Journal of Human Resources* 44, no. 2:479–94.
- . 2009b. Professor qualities and student achievement. *Review of Economics and Statistics* 91, no. 1: 83–92.
- Jurajda, Stěpán, and Daniel Münich. 2011. Gender gap in performance under competitive pressure: Admissions to Czech universities. *American Economic Review* 101, no. 3:514–18.
- King, Alan J. C., Wendy Warren, Will Boyce, Peter Chin, Matthew King, and Jean-Claude Boyer. 2002. Double cohort study: Phase 2 report. Technical report, Social Program Evaluation Group, Queen's University.
- King, Alan J. C., Wendy Warren, Will Boyce, Peter Chin, Matthew King, Jean-Claude Boyer, and Barry O'Connor. 2004. Double cohort study: Phase 3 report. Technical report, Social Program Evaluation Group, Queen's University.
- . 2005. Double cohort study: Phase 4 report. Technical report, Social Program Evaluation Group, Queen's University.
- Lavy, Victor. 2013. Gender differences in market competitiveness in a real workplace: Evidence from performance-based pay tournaments among teachers. *Economic Journal* 123, no. 569:540–73.
- Lindo, Jason M., Nicholas J. Sanders, and Philip Oreopoulos. 2010. Ability, gender, and performance standards: Evidence from academic probation. *American Economic Journal: Applied Economics* 2, no. 2:95–117.
- Morin, Louis-Philippe. 2013. Estimating the benefit of high school for college-bound students: Evidence of subject-specific human capital accumulation. *Canadian Journal of Economics* 46, no. 2:441–68.
- Niederle, Muriel, and Lise Vesterlund. 2007. Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics* 122, no. 3:1067–1101.
- . 2010. Explaining the gender gap in math test scores: The role of competition. *Journal of Economic Perspectives* 24, no. 2:129–44.

- . 2011. Gender and competition. *Annual Review of Economics* 3:601–30.
- Ors, Evren, Frédéric Palomino, and Eloïc Peyrache. 2013. Performance gender-gap: Does competition matter? *Journal of Labor Economics* 31, no. 3:443–99.
- Price, Joseph. 2008. Gender differences in the response to competition. *Industrial and Labor Relations Review* 61, no 3:320–33.
- Sacerdote, Bruce. 2001. Peer effects with random assignment: Results for Dartmouth roommates. *Quarterly Journal of Economics* 116, no. 2:681–704.
- . 2011. Peer effects in education: How might they work, how big are they, and how much do we know thus far? In *Handbook of the economics of education*, vol. 3, ed. Erik A. Hanushek, Stephen Machin, and Ludger Woessmann. Amsterdam: North-Holland.
- Stinebrickner, Ralph, and Todd R. Stinebrickner. 2006. What can be learned about peer effects using college roommates? Evidence from new survey data and students from disadvantaged backgrounds. *Journal of Public Economics* 90, no. 8–9:1435–54.
- Wolfers, Justin. 2006. Diagnosing discrimination: Stock returns and CEO gender. *Journal of the European Economic Association* 4, no. 2–3:531–41.

Copyright of Journal of Labor Economics is the property of University of Chicago Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.