ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 19-96

Effects of the shift to English-only instruction on college outcomes: Evidence from Central Asia

Aizat Nurshatayeva
University of Pittsburgh

Lindsay C. Page
University of Pittsburgh

English-only college education in non-English speaking countries is a rapidly growing phenomenon that has been dubbed as the most important trend in higher education internationalization. Despite worldwide popularity, there is little empirical evidence about how the transition to English-only instruction affects students' academic outcomes. Using a natural experiment at a selective university in Central Asia and a difference-in-differences strategy, we estimate the causal effect of switching to English-only instruction on students' college outcomes. We find that the introduction of English-only instruction led to a decrease of GPAs and probability of graduation and an increase in the number of failed course credits. Although negative, the effects were short-lived. The difference-in-differences estimates and the examination of potential mechanisms suggest that at least in selective universities in non-English speaking countries, the switch to English-only instruction may affect college outcomes negatively at the time of transition but may not necessarily imply longer-run negative effects.

VERSION: June 2019

Suggested citation: Nurshatayeva, Aizat, and Lindsay C. Page. (2019). Effects of the shift to English-only instruction on college outcomes: Evidence from Central Asia. (EdWorkingPaper: 19-96). Retrieved from Annenberg Institute at Brown University: http://www.edworkingpapers.com/ai19-96

Effects of the shift to English-only instruction on college outcomes: Evidence from Central Asia

Aizat Nurshatayeva
University of Pittsburgh
5904 WWPH, 230 South Bouquet Street, Pittsburgh, PA 15260, USA
Phone: 412-378-0526
Email: ain7@pitt.edu

Lindsay C. Page
University of Pittsburgh
5918 WWPH, 230 South Bouquet Street, Pittsburgh, PA 15260, USA
Phone: 412-648-7166
Email: lpage@pitt.edu

June 2019

Abstract

English-only college education in non-English speaking countries is a rapidly growing phenomenon that has been dubbed as the most important trend in higher education internationalization. Despite worldwide popularity, there is little empirical evidence about how the transition to English-only instruction affects students' academic outcomes. Using a natural experiment at a selective university in Central Asia and a difference-in-differences strategy, we estimate the causal effect of switching to English-only instruction on students' college outcomes. We find that the introduction of English-only instruction led to a decrease of GPAs and probability of graduation and an increase in the number of failed course credits. Although negative, the effects were short-lived. The difference-in-differences estimates and the examination of potential mechanisms suggest that at least in selective universities in non-English speaking countries, the switch to English-only instruction may affect college outcomes negatively at the time of transition but may not necessarily imply longer-run negative effects.

1. Introduction

In recent decades, many countries where English is not the native language have introduced English as a language of instruction in some or all of their universities. Some consider the switch to English as the language of college instruction as the most significant trend in higher education internationalization (Parr, 2014). The geography of countries making this change within higher education is notable and includes almost all European, former Soviet, and Asian countries, among others (Ackerley, Guarda, & Helm, 2017; Doiz, Lasagabaster, & Sierra, 2011; Goodman, 2014; Zhao & Dixon, 2017). The pace with which countries are adopting English as the language of instruction in colleges is equally remarkable; in Europe alone, the number of college programs taught in English grew from 725 in 2001 to 8,029 in 2014 (Wächter & Maiworm, 2014). Although these figures do not refer to exclusively English-only programs (which are the focus of this paper), they do serve as the best available international indicator of the growth in the adoption of English as a language of instruction in tertiary education.

By implementing English-only instruction,[1] colleges aim to attain ambitious goals such as entering and improving their positions in international rankings (Drljača Margić & Vodopija-Krstanović, 2017), enhancing research output by attracting faculty publishing in high-tier journals (Doiz et al., 2013), attracting more international students (Macaro et al., 2018; Wilkinson, 2017), and contributing to economic development by training an internationally competitive labor force (Dearden, 2014; Salmi, 2009). In pursuing these goals, an assumption is that colleges introducing English-only instruction will succeed in doing so.

[1] Researchers and practitioners use many terms to refer to the use of English as the language of instruction to educate students whose native language is other than English. Common terms include "immersion", "content and language integrated learning", and "English-medium instruction." See Macaro et al. (2018) for a comprehensive review of the phenomenon and terms used to describe it. We use the term *English-only instruction* and define it as the teaching of academic disciplines in English in non-English speaking countries.

However, implementation of English as the language of instruction might not be successful because instructors and students may be less than proficient in English (Doiz, Lasagabaster, & Sierra, 2013; Macaro et al., 2018). Studies from a variety of non-English speaking countries consistently have documented the challenges students and instructors experience when instruction is delivered in English (Bolton, Botha, & Bacon-Shone, 2017; Bradford, 2016; Hu & Lei, 2014; Nguyen, Hamid, & Moni, 2016). Yet, we lack reliable evidence about whether these self-reported challenges translate into actual negative effects on academic outcomes (Macaro et al., 2018).

We contribute to filling this gap in the literature on English-only instruction by examining the impact of an institutional switch to English as the language of instruction on student-level academic outcomes. We make use of an arguably exogenous policy shift at a selective university in Central Asia, to which we refer as Anon U., which introduced English-only instruction within one of its schools while making no changes in the language of instruction in its other schools. We employ a difference-in-differences (DID) design and administrative data on cohorts of students from before and after the policy change to estimate the causal effect of the transition to English-only instruction on a range of academic outcomes. In addition, we examine how the effect of switching to English-only instruction changed across cohorts and explore the mechanisms through which the effect of the language reform may have occurred.

To preview our findings, the programmatic switch to English has negative effects on some academic outcomes for the students in the first cohort affected by the policy change. We find no effect of switching to English-only instruction on third-year retention and the probability of dropout. However, in the first post-treatment cohort, the number of failed course credits increased by about 6 (0.47 SD), students' GPAs declined by 0.32 points (effects size of 0.36

standard deviations), and the probability of graduation within five years decreased by about nine percentage points (0.21 SD). In the subsequent post-treatment cohorts, effects of the language shift on GPA and probability of five-year graduation are not statistically significant, and the negative effect on the failed course credits persists in only one subsequent cohort. In other words, the negative effects were driven by a decline in the academic performance of the first cohort of students exposed to English-only instruction, with little evidence of detrimental effects for the cohorts that followed. This rapid fade out of negative academic consequences implies that concerns about potential sustained negative impacts of English-only higher education on students' academic outcomes may not always be warranted. We find no evidence that a particular, dominant mechanism drove the pattern of effects that we observe.

Our study makes several contributions to the existing literature on English-only instruction. First, we contribute to filling the gap in understanding the effect of switching to English-only instruction on students' academic performance by using a rigorous quasi-experimental approach. To our knowledge, our paper is the first study of English-only instruction conducted in the causal inference framework. We estimate the impact of shifting to English-only instruction on a set of measurable and policy-relevant academic outcomes to provide a more complete picture of the language reform impacts that can inform institutional policy and further research. Although we recognize this investigation as likely having limited generalizability, it nevertheless represents a research approach that may be applicable in many other settings where English-only instruction is being newly implemented. Second, we extend the debate about English-only instruction by considering how and why the impact changes across cohorts. The short-lived nature of the negative effects across cohorts we identified and the mechanisms we examined provide the basis for studying the factors that hinder or contribute to the successful

implementation of English-only instruction. Third, we use data that is often hard to obtain. Colleges that have implemented English-only instruction may be reluctant to share their data to avoid publicizing any negative impacts of their reforms. We show in our study that it is possible to gain access to such institution-level administrative data and conduct policy-relevant research while protecting the privacy of the data source.

We structure the remainder of the paper as follows: Section 2 presents the literature review; Section 3 describes the data, background, and analytic approach; Section 4 presents estimation results and discusses the mechanisms driving the observed effects; Section 5 discusses the findings and concludes.

2. Literature review

2.1 Why colleges choose to switch to English-only instruction

Motivation for implementing English-only college-level programs varies by geographical context. For example, most European and certain Asian countries that aim to expand as international education hubs have introduced English-only programs (Airey et al., 2017; Macaro et al., 2018). In addition, in Europe, the focus on ensuring comparability of the standards and quality of higher education stimulated many universities to develop English-taught programs to tap into the funding and growth opportunities of the common European higher education area (Ackerley, Guarda, & Helm, 2017; Wilkinson, 2017).

In other contexts, improving university ranking and research prestige is a more salient rationale (Drljača Margić & Vodopija-Krstanović, 2017). Sometimes, entering international university rankings or improving the standing of national universities in those rankings distills into explicit government policies. For example, Russia's Project 5-100 was begun in 2012 with

the goal of at least five Russian universities entering the top 100 of three leading international university rankings (Moed, Markusova, & Akoev, 2018). Establishing English-taught programs allows colleges to hire faculty with western training who are capable of producing high quality research and thereby is instrumental in improving research output scores, the key metric of the international university rankings.

In certain instances, the growth of English-only college programs is linked to the goal of establishing world-class universities to foster economic development by training the country's talented youth to be globally competitive (Doiz et al., 2013). There is general consensus among policy-makers that universities aspiring to be world-class benefit from operating in the English language (Dearden, 2014; Salmi, 2009). For example, since the early 1990s, China has established national higher education policies to build world-class universities to drive and support further economic growth (Hu & Lei, 2014).

This push for English-only instruction is built on an implicit assumption that colleges implementing English as the language of instruction are successful in doing so and that students can succeed in English-only programs (Doiz et al., 2013; Drljača Margić & Vodopija-Krstanović, 2017). Indeed, the goals listed above can hardly be achieved if the implementation of English-only instruction fails. Therefore, the academic performance of the students in English-only programs may be considered a prerequisite or proxy for the ambitious goals set by education policy-makers who design and enact English-only programs. In sum, to understand whether English-only instruction is likely to attain the goals listed above, it is informative to investigate how the switch to English-only instruction affects students' academic outcomes.

2.2 Possible effects of English-only instruction on students' academic performance

Numerous studies suggest that English-only instruction may adversely affect students' academic

outcomes. For students whose native language is not English, writing, reading, listening to and speaking about academic material in English present additional challenges to the learning process. For example, Bolton et al. (2017) report that both undergraduate and graduate students in Singapore experience difficulties in academic communication even though English-only instruction has been the language of core higher education for several decades. Such difficulties, particularly among graduate students, may signal that lack of English proficiency is detrimental to understanding complex material.

That lack of English language proficiency prevents students from adequately understanding college-level material has also been documented in other contexts, including South Korea, China, Hong Kong, and Turkey. Across these contexts, students reported having difficulty understand course content taught in English; requesting that instructors use their native language during course instruction (Byun et al, 2011); failing to understand large parts of lecture content, particularly related to abstract disciplinary concepts and technical vocabulary (Hu & Lei, 2014; Evans & Morrison, 2011; Sert, 2008); and struggling to write in an appropriate academic style (Evans & Morrison, 2011). These studies tend to agree that the challenges students experience stem from a lack of sufficient proficiency in English.

In addition to students' language difficulties, instructors' limited English proficiency may prevent them from adequately supporting students to master complex concepts. For example, when required to teach in English, professors tend to reduce the amount of academic content covered in lectures and slow down lecture pace. Across a variety of contexts, including Japan (Bradford, 2016), South Korea (Byun et al., 2011), China (Hu & Lei, 2014), Hong Kong (Evans & Morrison, 2011), Turkey (Sert, 2008), and others (Doiz et al., 2013), professors cover at least 20-30% less material when teaching in English compared to when teaching in their native

language. If faculty are restricted in their capability to deliver the full course content due to the required language of instruction, student learning may be hindered.

Despite these substantial challenges, both students and instructors in the contexts studied report overall positive attitudes towards English-only instruction and report using various coping strategies to address the challenges they experience. Prospects of better academic and labor market opportunities motivate students to work harder and use peer support (Evans & Morrison, 2011) or to review course content both in English and in their native language (Sert, 2008; Hu & Lei, 2014). Better professional opportunities and the desire to help their students motivates instructors to improve their English proficiency and their English-based instruction (Bradford, 2016; Sert, 2008; Zhao, 2017). That is, along with the difficulties, evidence suggests that both students and faculty can be highly motivated to study and teach in English and to utilize coping strategies to help themselves. Such ability to adjust and respond may mitigate potential negative effects of language-related challenges on learning in English-only programs.

Our review highlighted two notable gaps in the literature on English-only instruction. First, although research has documented substantial challenges experienced by both students and faculty, it is unclear whether these challenges translate into negative effects on students' academic outcomes. Second, a related question is to what extent both instructors and students whose native language is not English can adjust to college-level coursework in English. It is possible, for example, that any negative effects of switching the language of instruction to English may dissipate as students and professors adapt to the change (Macaro et al., 2018). These issues deserve careful study. If the documented challenges of English-only instruction convert into persistent deficiencies in content learning and achievement of academic milestones (e.g., degree attainment), then the ambitious goals set by policy-makers implementing English-only

instruction may be hindered and resources wasted. We contribute to filling these gaps in the literature by answering the following research questions:

1) How does the switch to English-only instruction within a postsecondary institution affect student-level academic outcomes?
2) How do the effects of switching to English-only instruction change across student cohorts, and what drives these patterns?

3. Data, background, and research method

3.1 Dataset

We use administrative data from Anon U (pseudonym), a university in Central Asia, for six cohorts of students (N = 2,884) who entered the university between 2007 and 2012. The data come from students' administrative records and include information on their demographic characteristics, high school language of instruction and achievement measures, university language of instruction and achievement measures, and university financial aid status.

3.2 Setting

The data from Anon U is well-suited to inform our research questions for three main reasons. First, Anon U is similar in most respects to a typical university that is likely to implement English-only instruction in a non-English speaking country. Specifically, it is selective, western-oriented, and well resourced. It is a small university, enrolling an annual cohort of approximately 350 students, selected from among the top students in the country. As shown in Table 1, the average scores in the national standardized college admission test taken at the end of high school range from 85 to 95 points (out of 100), placing Anon U students in the top quartile of the national distribution for these tests. Other indicators reflect the relatively

advantaged background of the average Anon U student. For example, about 85% of Anon U students are from urban areas which are better off economically than rural areas, more than half of the students come from Russian language high schools where academic achievement tends to be higher,[2] and about 90% of Anon U students are of the local ethnicity. In short, the student body is fairly homogeneous and continued to be so throughout the period of our examination. Almost all Anon U graduates find employment within three months of graduation in the national labor market or enroll in competitive graduate programs. Furthermore, Anon U was created in the post-USSR era and modeled after western universities. Expatriates from the U.S. and European countries comprise a considerable share of Anon U's administrators and faculty. Anon U's academic process, grading, and instructional practices are structured similarly to U.S. universities. Anon U is well-funded by the government and enjoys financial support from transnational and local businesses.

Second, document analysis shows that the external and internal motivations of Anon U were similar to those discussed above.[3] Externally, the government encouraged the university to pursue entering international rankings and to produce more impact-factor publications (most commonly written in English). In addition, Anon U experienced competition from newer world-class-aspiring universities in the region that offered instruction in English. Internally, the university felt the need to employ more professors capable of producing high quality research and to make graduates more competitive by offering highly demanded training in English.

Third, Anon U's administrative data is of high quality. Modeled after western universities

[2] Given that this Central Asian country has been a part of the USSR where the Russian language was the dominant language in education and other spheres of life, graduates of the Russian language high schools tend to perform better academically.
[3] Specifically, we examined the country's government plans for strategic development, Ministry of Education plans and reports, and Anon U's senior administrators' interviews in mass media. We do not include sources of these documents so as to protect the identity of the university.

and in order to prevent corruption, Anon U has maintained a comprehensive electronic database of student files and transcripts since its founding. In addition, upon our request, Anon U administrators checked the key variables in the electronic database against the university's internal documents.

3.3 Language of instruction shift at Anon U

We make use of a natural experiment that took place in 2010 at Anon U to examine the impact of shifting to English as the language of instruction. Before 2010, Anon U allowed its incoming students to choose their language of instruction, with Russian and the local language as options. In the 2010-2011 academic year, Anon U's School of Computer Science changed its language of instruction policy such that from 2010 onward, all students were able to study in Russian or the local language in the first year only but were required to switch to English starting in the second year. We exploit the fact that only the School of Computer Science switched to English-only instruction in 2010 while all other schools continued to teach in Russian and the local language. Therefore, we use all other schools within the university as the comparison group.

Two features of the language policy shift at the School of Computer Science make identification of causal effects possible. First, the decision was made by the senior management of the university and was imposed on the School of Computer Science. Expecting strong resistance from the faculty and administration of the School of Computer Science, Anon U's senior administration conducted the reform in a very abrupt manner. All possible measures were taken so that faculty and mid-level administrators could not delay implementation of the reform. Within a few months, several non-English speaking faculty members were replaced with lecturers who could teach in English. In short, the policy shift could not have been easily

anticipated by faculty and students, making the change an arguably exogenous shock.

Second, course schedules at Anon U are rigidly structured by cohort which enables us to identify the effect of the cohort-based language of instruction change. Specifically, Anon U does not allow students from multiple entry cohorts in a single course. For instance, a computer science course for second-year students would not have any first-, third- or fourth-year students in it. Given this cohort-based curriculum and schedule structure, the 2009 entering cohort experienced no effects of the switch to English and graduated from Anon U taking classes in their originally chosen languages. Such characteristics of Anon U's educational offerings (customary in the country) ensures that the language of instruction switch impacted only the cohorts enrolled in 2010 and onwards.

Several factors motivated the Anon U administration's choice to transition the the School of Computer Science specifically to English-only instruction. First, the undergraduate computer science program at Anon U has traditionally attracted high performing students from the country's secondary schools. Further, coding and professional terminology in computer science is heavily English-dominated. Russian and local language textbooks were often translated editions of English-language originals. They perceived computer science-related professions to be sufficiently grounded in English, making the switch to English appear as a reasonable policy solution to eliminating the dependency on translated instructional materials. Third, the mathematics department at Anon U (separate from the School of Computer Science) had several applied mathematicians with teaching experience at U.S. universities who were capable of teaching math-heavy computer science courses in English. Given these favorable conditions for transitioning to English-only instruction, the immediate negative effects for students that we observe may reasonably be interpreted as a lower-bound for possible effects in other contexts

that may be less well positioned to manage such a change.

3.4 Research method

We estimate the effects of the language shift on academic outcomes using a difference-in-differences (DID) framework. The School of Computer Science is the treatment group, and other schools of Anon U serve as a comparison group. In essence, we compare changes in student outcomes before and after the shift to English instruction to analogous changes in the comparison group to identify the causal effect of the treatment.

The model specification is shown in equation 1.

$$Y_{it} = \alpha + \beta_1 Post_{it} + \beta_2 English_{it} + \beta_3 Post_{it} * English_{it} + \beta_4 \boldsymbol{X}_{it} + \lambda_t + \theta_m + \epsilon_{it} \quad (1)$$

In equation 1, $Post_{it}$ is a dummy variable indicating whether student *i* in cohort *t* is from a post-treatment cohort. $Post_{it}$ is equal to 1 if student *i* entered Anon U in 2010 or later and is equal to 0 if a student entered Anon U before 2010. The pre-treatment cohorts in the data are 2007, 2008, and 2009, whereas the post-treatment cohorts are 2010, 2011, and 2012. $English_{it}$ is a dummy specifying whether a student was enrolled in the school that implemented the shift to English-only instruction (i.e., the School of Computer Science). The coefficient β_3 on $Post_{it} * English_{it}$ is the DID estimator of the effect of the language switch on the outcome variable. $\boldsymbol{X}_{it}$ denotes a vector of student baseline characteristics including gender, ethnicity, home locality (urban or rural), language of instruction in high school, national standardized college admission test score, financial aid status (state grant, Anon U grant, or self-supported), and a dummy for whether a student transferred in from another college. When fitting this model, we additionally include fixed effects for cohort (λ_t) and major (θ_m).

Y_{it} represents each of our outcomes of interest. All cohorts have five years of follow-up for the outcomes. We consider five primary outcomes. The first outcome is retention into the

third year captured by a dummy equal to 1 if a student is engaged in coursework after completing the second year of studies.[4] The second outcome is dropout, equal to 1 if a student drops out of Anon U within five years. The third outcome is the number of failed course credits, a continuous variable. To put this variable in context, one credit at Anon U is similar to one credit unit at four-year U.S. colleges, and most courses at Anon U are three-credit courses. The fourth outcome is GPA at graduation, a continuous variable on a 4-point scale calculated just as in U.S. universities. For those who did not graduate, we use GPA at the end of their studies. The final outcome variable is degree completion, represented by a dummy variable equal to 1 if a student graduates within five years. Arguably, five-year graduation has a limitation of being a potentially censored version of true graduation rates, as some students may take longer than five years to graduate. Nevertheless, we focus on five-year graduation, because we have data as of 2017, so we can track the most recent 2012 cohort in the study only for five years. As a sensitivity check, we additionally considered graduation within six years for the 2007-2011 cohorts and obtained estimates similar to those from using five-year graduation. Therefore, we are confident that five-year graduation captures the graduation outcome quite well in the context of Anon U.

We explore whether the effect of the shift to English-only instruction is stable across cohorts by re-estimating equation 1 using each one of the post-treatment years alone as a post-treatment period. That is, we estimate equation 1 using 2010 as the only post-implementation cohort and then do the same for 2011 and 2012. Comparing cohort-specific estimates of β_3 allows us to examine whether the effect of the language switch is stable across cohorts.

[4] We also examined retention into the second year and retention into the fourth year. Estimates of the policy impact were similar to estimates using retention into the third year. Therefore, we present estimates for retention into the third year only.

3.5 Internal validity of the research design

The key assumption underlying causal inference using a DID analytic strategy is the assumption of parallel trends (Imbens & Rubin, 2015; Murnane & Willett, 2011). This assumption, in essence, states that trends in outcomes in the comparison schools serve as a valid counterfactual for how students would have performed in the School of Computer Science had the shift to English-only instruction not occurred. We take several steps to consider the reasonableness of this assumption in the Anon U context.

First, we plotted each independent variable across cohorts for the treatment and comparison schools (Figures 1-9). The trends by treatment status appear parallel for gender (Figure 1), locality from which students come (Figure 2), ethnicity (Figure 3), standardized college admission test scores (Figure 4), Anon U institutional financial aid (Figure 5), and whether students transferred in from another college (Figure 6). In contrast, we observe somewhat divergent trends in the proportions of students from local language high schools (Figure 7). Figure 7 suggests that the School of Computer Science historically attracted more students from local language high schools except for a dip in 2009 and a jump in 2012. In general, this might imply changes in student quality as graduates of the Russian language schools tend to perform better academically due to the dominant role of the Russian language in formal instruction during the USSR times. In addition, Figure 8 suggests that trends in the proportion of students awarded state grants were not completely identical in the treatment and comparison schools.[5] Similarly, Figure 9 implies that trends in the share of self-supported students (i.e., students who do not receive any financial aid) were not completely parallel. Below, we discuss whether these patterns threaten causal inference.

[5] State grants and Anon U grants are mainly merit-based and are awarded to the students whose high school academic performance and national standardized college admission test scores.

Next, we formally model how each covariate changed across the pre-/post-treatment periods in the treatment and comparison groups. The descriptive statistics shown in columns 1-6 of Table 1 align with the graphical analysis. We further explore the covariate changes in the DID framework. In Table 1, column 7, we show that the share of students from local language high schools increased, the share of students receiving state grants increased, the share of students receiving Anon U grants decreased, and the share of students who transferred in from other colleges decreased in the School of Computer Science relative to the other schools.

We then consider whether these shifts in covariates threaten our ability to draw causal conclusions using the DID analytic strategy outlined above. As we show in section 4.1, impacts on student outcomes of interest are concentrated in the 2010 cohort. Therefore, we check whether the changes observed in some of the covariates took place simultaneously with the treatment impacts on the outcome variables. We re-fit the DID models for each covariate using only 2010 as the post-treatment period. These DID estimates presented in column 8 of Table 1 are not statistically significant and are generally smaller in magnitude than those that pool across all cohorts to experience the policy change. These results illustrate that none of the covariates changed significantly at the time the policy impacts on students were concentrated. In other words, the overall differences in covariates are driven by the cohorts further away from the policy shift. Although we do observe some changes in these baseline measures across cohorts, they are not in obvious alignment with the patterns of impacts of the policy on student outcomes. Furthermore, the shifts in covariates for the later cohorts do not imply these cohorts were comparatively stronger than the 2010 cohort. Specifically, as we show in section 4.2, the post-treatment cohorts were similar to the pre-treatment cohorts in terms of overall academic preparation and English language proficiency. In other words, the post-treatment cohorts were

not necessarily better prepared to manage English-only instruction. Taken together, we judge the parallel trends assumption to be reasonably well met. In addition, we control for the covariates discussed here in all of our preferred model specifications.

4. Results

4.1 The effect of switching to English-only instruction on academic outcomes

In Table 2, we present results from estimating equation 1 to examine impacts of the language change on third-year retention, dropout, course failure, GPA, and degree completion within five years. The coefficients associated with the Post*English interaction term denote the estimate of the language switch effect. For each outcome, we ran four model specifications: the model (1) with no covariates, (2) with student characteristics, (3) with student characteristics and fixed effects for students' majors, and finally, (4) with student characteristics and fixed effects for majors and cohorts. Results are robust to these modeling choices. In Table 2, we present results from the fourth specification (additional results are available upon request).

As shown in Table 2, the switch to English-only instruction had a negative impact on three academic outcomes. First, estimates suggest that the switch to English-only instruction had no effect on third-year retention and probability of student dropout. Second, the course failure rate increased by about three course credits (0.22 standard deviations). A typical course at Anon U is worth three credits, therefore, computer science students failed one more course on average after English-only instruction was implemented. Third, GPA fell by 0.13 points (0.15 standard deviations). We note that when students at Anon U retake a course they failed, their new grade overrides the one previously earned. This implies that the estimated decline in GPA is likely

smaller than the actual decline before adjusting for retaken courses.[6] Finally, the five-year graduation rate dropped by about 7 percentage points (0.17 standard deviations). This impact is considerable, given the high academic capability of Anon U students. As a point of comparison, the university's typical five-year graduation rate is around 80%.

To test the robustness of our estimates to the choice of comparison, we reran the models using each of the non-treatment schools as a comparison group. We present results in Table 3. The coefficients change slightly depending on which non-treatment school serves as the comparison. The estimates are larger in magnitude when comparison schools are those where average academic performance has historically been higher (e.g., the School of Chemistry and the School of Business). In contrast, estimates are somewhat smaller when the comparison school is limited to the School of Engineering, in which students have a relatively higher number of failed courses and lower probability of five-year graduation. Substantively, however, estimates remain similar to those using all other schools together as the comparison group.

Next, we explored whether the effect of the shift to English-only instruction is stable across cohorts. First, we graphed raw outcomes across cohorts. Figures 12-14 suggest the fading out of the negative effect of the policy shift on course failure rate, GPAs, and five-year graduation, with a "bump" for the 2010 cohort followed by a recovery for the 2011 and 2012 cohorts. The overall pattern in the outcome variables suggests a disruptive effect of the shift to English-only instruction on the first treated cohort. In contrast, the subsequent cohorts in the School of Computer Science appear to have been better able to adjust.

Second, we re-estimated impacts for each of the post-treatment cohorts separately. We present results in Table 4. Consistent with the graphical depiction, the policy shift had the most

[6] In the data available to us, we cannot demarcate patterns of course retaking and, therefore, are unable to adjust for course retakes in our estimates.

consistently negative effects for the first cohort of students to experience English-only instruction. The 0.12 point drop in GPA estimated using the full sample (column 1 of Table 4) was driven by a 0.32 point GPA decline for the 2010 cohort, which was the first to experience the switch to English-only instruction (column 2 of Table 4). Similarly, the seven percentage point decrease in the probability of five-year graduation estimated using the full sample was concentrated in the 2010 cohort, which experienced a nine percentage-point decline. The increase of 3.03 course credits failed estimated for the pooled sample is driven mostly by the increase of failed course credits by 5.7 in the 2010 cohort. Unlike the other outcome variables, however, we do observe a negative effect on course failure for the 2012 cohort (column 4 of Table 4). Nevertheless, the negative effect of 3.34 more course credits failed of the 2012 cohort is somewhat smaller in magnitude than for the 2010 cohort.

These results for failed course credits suggest that later cohorts may have still struggled with learning in English. However, we view this more persistent negative effect in conjunction with the fading out of negative effects for the GPA and five-year graduation outcomes and the absence of any impact on dropout and third-year retention. The academic guidelines at Anon U (and other universities in this country) indicate that when a failed course is re-taken, the "Fail" grade is replaced by the new grade. Thus, if a failed course is re-taken with a better grade, the "Fail" grade does not affect GPAs and probability of five-year graduation. In sum, although students in the 2011 and 2012 cohorts may have still struggled with course instruction in English (leading to more course failures), by retaking courses failed, they were able to recover academically and still graduate within five years. Given that a typical course is worth three credits, the typical Anon U student would have needed to retake about one course because of course failure.

4.2 Why was the negative effect of the language shift limited primarily to the first affected cohort?

Here, we consider potential reasons that the negative effects of the shift to English-only instruction on GPAs, failed course credits, and probability of five-year graduation were experienced primarily by the first affected cohort.

Did the School of Computer Science start to enroll students who are academically stronger in later cohorts?

One possibility is that academic ability was higher in the cohorts following the first treated cohort because Anon U's School of Computer Science started to attract students who were more capable academically. To assess this hypothesis, we consider two measures of the academic ability of Anon U's incoming students. The first is the national standardized college admission exam scores. As shown in Table 1, these scores fell by about 5 points in the School of Computer Science after it switched to English-only instruction. Exam scores fell similarly, by about 6 points, in the comparison schools. The DID estimates to examine differential changes in exam scores using the full sample (Table 1, column 7) and using only 2010 as the post-treatment period (Table 1, column 8) were both statistically insignificant suggesting that the treatment school did not experience changes in academic ability of its incoming students relative to the comparison schools. Figure 4 shows further that tests scores remained relatively stable through 2012.

The second measure of academic ability is grades in the first semester at Anon U. There are two advantages to using first-semester grades. First, the courses in the first semester were taught in either Russian or the local language corresponding to the language in which students

studied in high school. Therefore, we can observe student performance free from the effect of English as the language of instruction. Second, several of these first-semester courses at Anon U are not chosen by students but are mandated in the curriculum. Incoming freshmen are automatically enrolled in a sequence of courses including math, physics, history, and languages. The content of these courses is fairly standard and has not changed over the period under study. Given such general content and mandated enrollment, grades in these courses serve as an institution-level standardized measure of students' ability to do college-level coursework.

We plot averages of grades in the first semester courses for each cohort in the School of Computer Science in Figure 15. Grades in these first semester courses remained quite stable across cohorts suggesting that there were no substantial improvements in student body. If anything, the 2011 incoming cohort struggled somewhat more in math compared to other cohorts. As Table 5 shows, there were no significant differences in students' performance in these first semester courses before and after the English-only policy. The DID estimates using the pooled sample (column 7 of Table 5) show that the treatment school grades in first semester math, physics, and languages did not change while grades in history decreased relative to the comparison schools. We observe a similar pattern of first semester grades when we restrict the post-treatment period to 2010 only (column 8 of Table 5). Taken together, both the descriptive trends and the DID estimates of national standardized college admission test scores and first-semester course grades indicate that there was no consistent improvement in the academic preparation or first semester performance of students in the School of Computer Science. In sum, it is unlikely that the later cohorts were better able to handle the English-only instruction due to comparatively stronger academic ability.

Did later cohorts of students in the School of Computer Science have stronger English skills?

Even though the later cohorts were not academically stronger overall, it is possible that they had stronger English skills, specifically. Anon U tests all incoming students' English proficiency to place them into a relevant English language course at the elementary / pre-intermediate, intermediate, or upper-intermediate / advanced levels. We were able to obtain English language course placement for the 2010, 2011, and 2012 cohorts, but not for early cohorts. A limitation is that placement test scores for the 2010 cohort were incomplete, with 22% of observations missing. To handle this, we imputed values of the missing scores.[7]

As shown in Table 6 (Panel A), there were no drastic improvements in students' English proficiency across cohorts in the School of Computer Science. Although the share of upper-intermediate and advanced students increased from 8% in 2010 to 16% in 2012, the share of elementary and pre-intermediate level students remained relatively stable at about 46% in 2010 and 2011 and reached 68% in 2012. In addition, the proportion of intermediate level students declined every year from 43% in 2010 to 16% in 2012. Thus, it is unlikely that improved English proficiency explains the general recovery in the 2011 and 2012 cohorts.

Further, English language proficiency trends in the comparison schools were similar to those of the treatment school (Table 6, Panel B). Overall, the proportion of top English proficiency level students did not change dramatically while the share of lower English proficiency level students increased in the treatment school and remained stable in the

[7] We impute English language placement test scores because they are important for examining whether Anon U started to recruit students with better English language skills. We used multiple imputations with chained equations based on the following variables: gender, ethnicity, urbanicity, language of instruction in high school, national standardized college admission test score, GPA at graduation, number of fails by graduation, five-year graduation status, grade in first semester English language course, cohort, major, and financial aid status. We estimated the proportions of students by English proficiency level using 20 multiply imputed samples.

comparison schools. In sum, these patterns provide no affirmative support for higher levels of English proficiency in the later cohorts driving our main results.

Did the School of Computer Science change its criteria for course passage and graduation?

The negative effect of switching to English-only instruction may have been temporary because the School of Computer Science changed its criteria for course passage and graduation after observing negative effects for the first cohort of students to experience English-only instruction. We rule out this mechanism based on our understanding of the Anon U context. Although the USSR-era total state control over the education system has been relaxed, state standards and regulations still govern all important aspects of college instruction in the context that we study. More than 50 regulatory statutes direct the day-to-day academic process. The skills and themes to be mastered in every course are prescribed by the state curriculum and faculty members have limited space to substantially alter the content covered in their courses. State standards regulate student admissions, student assessment, passage from year to year based on academic performance, and graduation requirements. Every five years the Ministry of Education audits all universities for compliance with the state standards. The Ministry of Education auditing teams examine the university administrative records (including course syllabi and department meeting minutes) for the preceding 5 years, interview faculty and students, and conduct teacher evaluations by attending several classes selected at random. Noncompliance with state regulations uncovered during these audits leads to the recall of the university's license to operate, forcing either the university or the non-compliant school within it to close. The Ministry of Education has closed several universities and, in some cases, rescinded licenses partially so that colleges could not offer certain majors that did not comply with the regulations.

The most recent Ministry of Education audits of Anon U took place in 2010 and 2015 and concluded that Anon U was fully compliant with the state standards. Given the high stakes of noncompliance with the state standards and given the results of the most recent audits of Anon U, it is unlikely that Anon U would risk its license by allowing the School of Computer Science to change its course passage and graduation requirements.

Did a change in instructor characteristics drive the temporary dip in academic outcomes?

The introduction of English-only instruction required hiring faculty who could teach in English. Thus, it is possible that the dip in academic outcomes in the first treatment cohort is a function of changes in instructor characteristics. To explore this potential mechanism, we examined how the proportion of new instructors hired at Anon U changed between 2007 and 2012. As Figure 16 shows, the proportion of new professors has historically been quite high – between 30 and 40% in both the treatment and comparison schools. Importantly, there are no significant fluctuations in academic outcomes in the pre-treatment period where we observe substantial variation in the proportion of new instructors. This suggests that student performance is not linked to whether instructors are newly hired at Anon U.

To further explore this mechanism, we examined the relationship between instructors' years of experience at Anon U and students' grades.[8] Table 7 presents estimates from regressing students' grade in each course on the years of experience at Anon U of the instructor who taught every given course. These results suggest that instructors' years of experience at Anon U are not correlated with students' performance.

[8] We were not able to obtain data on the average overall years of experience or other instructor characteristics.

We also examined the average years of experience at Anon U among instructors who taught courses that were the most challenging for the students. Table 8 shows the courses in which the majority of "Fail" grades are concentrated. As the rightmost column of Table 8 indicates, these "top three" most challenging courses were taught by experienced instructors, with years of experience at Anon U ranging between 2 to 6 years, on average. For the 2010 cohort which had the biggest decline in academic performance, none of the most difficult courses were taught by the newly hired instructors. In contrast, students struggled the most in the courses delivered by professors who previously taught at Anon U and who themselves had to adapt to teaching their courses in English starting from 2010. Collectively, we conclude that changes in instructor characteristics were not the mechanism through which the decline and subsequent rebound of academic performance occurred. Rather, students performed similarly when taught by more experienced instructors and as when taught by more recently hired instructors and both students and faculty needed to adapt to the requirement of English-only instruction.

Which English-only courses were the most challenging for students?

Another important insight from Table 8 is that the lists of the most difficult courses for each cohort are represented predominantly by the mathematics courses and elective courses taught by departments outside of the School of Computer Science. Table 8 shows that students found mathematics courses most difficult both before and after the switch to English-only instruction in 2010. The "Calculus 2" and "Probability theory and statistics" courses were challenging for students to master prior to as well as after the language reform had been implemented. In the 2011 and 2012 cohorts, where we observe recovery in academic outcomes according to the DID estimates, students were able to avoid failing "Calculus 2". However, the

courses in “Probability theory and statistics” and “Differential equations” were still challenging for some of the post-treatment cohorts. Overall, the courses that posed the biggest challenges both before and after the language reform were mathematics courses and elective courses such as “Principles of economics” and “English for professional purposes” taught by departments outside of the School of Computer Science.

It is sensible that the computer science courses did not make it to the top of the lists of the most challenging courses in Table 8. Computer coding is done in English even when lectures and tutorials are in the native language, so students did not experience drastic shocks in studying their core computer science courses due to overall familiarity with computer technologies and already possessing introductory-level coding skills. In contrast, studying advanced mathematics in English requires abstract thinking using the English language and involves the use of terms that should be learned in English. Elective courses outside of the School of Computer Science also require intensive learning of English terms and expose students to discourses, types of assignments and academic activities (e.g., essays, reading literature not related to computer science directly, etc.) that are not as common in the core computer science courses. At the same time, both mathematics and non-core electives taken outside of one’s home school were challenging for students prior to the language reform suggesting that the switch to English-only instruction made the traditionally difficult courses even more difficult for the students.

5. Discussion

We estimated the effects of switching to English-only instruction on college outcomes using a DID framework and administrative data from a Central Asian university. Effects are generally in the hypothesized direction. Students failed more courses and earned lower GPAs, on

average, when English-only instruction was introduced. In addition, although students did not drop out, more students did not graduate within five years, suggesting that students took longer to graduate. Our results suggest that the challenges of English-only instruction discussed in the literature (Bolton et al., 2017; Doiz et al., 2013) may indeed translate into lower academic achievement but may not necessarily prevent students from continuing their studies. Despite having the top students of the country and giving them one year within Anon U to prepare for the language transition, the negative effects of the switch to English-only instruction on academic performance are considerable in magnitude.

Nevertheless, the negative effects of the language shift fade out rapidly across cohorts. Our estimates suggest that there was a dip in academic performance in the year of transition from which the system subsequently recovered. Specifically, the switch to English-only instruction posed challenges for the first cohort of students who experienced the language of instruction transition but did not present sustained challenges to the cohorts that followed. This finding, consistent with Macaro et al. (2018), suggests that at least in contexts similar to Anon U, it is possible that students and faculty can adapt quickly to such a change in the language of instruction. The improvement in academic performance in the cohorts following the first treated cohort might indicate adaptation of students and teachers to English-only instruction.

To shed more light on the nature of the adaptation of subsequent cohorts to English-only instruction, we examined several possible mechanisms for the pattern of effects we observe. Our analyses show that the decline and subsequent improvement in academic outcomes was not driven by the School of Computer Science beginning to enroll students who were more capable academically. Neither did the English proficiency of incoming students improve nor course passage and graduation requirements change. Finally, the data suggest that changes in instructor

characteristics did not drive the academic outcomes of students. Our interpretation of the lack of a single mechanism driving the impacts we observe is that, more generally, both students and faculty likely adapted to the change.

The initial dip in academic performance followed by a recovery across cohorts at Anon U links our study to educational research that has studied the introduction of new educational policies or assessments in other contexts. For example, Linn (2000) examined such a phenomenon focusing on several waves of tests introduced in the U.S. from the 1950s to the 1990s. Linn (2000) highlights that whenever a new version of a math or reading test was introduced, either at the federal or state level, students in the first cohort to take the new test would perform less well than subsequent cohorts taking the same test. Linn (2000) refers to this as the "saw-tooth pattern" because the trends in average test scores across cohorts look like teeth of a saw, going down in the first new-test-taker cohort and up in the subsequent cohorts. Our findings showing an initial dip in academic performance followed by a recovery across cohorts at Anon U fit this saw-tooth pattern discussed in the education assessment implementation literature (Koretz et al., 1991; Linn, 2000).

Researchers who have detected such saw-tooth patterns repeatedly have warned that using such trends in test scores should not be interpreted as indication of improving academic performance in the subjects that these test evaluated (Linn, 2000). Instead, the pattern suggests that students simply get used to the new test format and teachers focus on content necessary to perform well on the test (Koretz et al., 1991). Our findings and the lack of a dominant measurable mechanism to explain them suggest that an adaptation of a similar sort likely took place at Anon U. Arguably the cohorts following the first treated cohort at Anon U had more resources to navigate the challenges of English-only instruction. For example, these subsequent

cohorts had an opportunity to learn about the pitfalls of studying in English from the first cohort and therefore could deploy more proactive coping strategies. In line with prior literature, we argue that our findings should not be interpreted as evidence of better learning of the cohorts following the first English-only cohort at Anon U but instead as evidence of adaptation to learning in English.

In addition to contributing to the conversation about distinguishing between academic performance and learning, our study generated insights relevant to institutions contemplating shifts to English-only instruction and its potential negative effects on students' academic performance. Specifically, the analysis of courses in which students failed most frequently revealed that students performed the worst in two types of courses. The first type includes courses that were challenging for students prior to the English-only reform, such as advanced mathematics courses. The second type comprises mostly elective courses requiring the need to master new terms, discourses, and types of assignments to which students were not previously exposed. In contrast, computer science core courses were not among those which students found most challenging when delivered in English. These findings suggest that while implementing English-only instruction, colleges may be able to mitigate the negative effects of the reform on students' academic outcomes by providing more support to instructors and/or offering academic support services to students in courses requiring acquisition of new terminology and in courses exposing students to the academic work less common in the core courses.

The main policy implications of our study are three-fold. First, our findings suggest that selective universities (e.g., at least those similar to Anon U) aiming to switch to English-only instruction may be successful in doing so in the long run. It is possible to implement English-only instruction relatively successfully even in contexts like Central Asia, located far away from

the centers of the English-speaking world. To the extent that successful implementation of English-only instruction is a pre-requisite for better rankings, attracting more international students, or more research output, our findings imply that colleges indeed may consider introducing English-only programs as a realistic scenario. Overall, that losses in student outcomes of the first cohort dissipate in subsequent cohorts may justify the significant financial and training investment required for implementing English-only instruction. Second, colleges contemplating or already implementing a transition to English-only instruction should consider their capacity to establish English-only programs. We recommend viewing our findings as lower-bound estimates of the possible negative effects. In contexts where the student population is not as strong academically as at Anon U and in programs and disciplines less grounded in English than computer science, the negative effects on the initial cohort may be larger, and the fade-out across cohorts might occur more slowly. Third, colleges intending to become English-only should be prepared to proactively mitigate potential declines in academic performance and pay close attention to the early cohorts exposed to the language transition. The negative effects of the language shift may be alleviated by providing more support to the instructors teaching English-only courses that require mastery of extensive new terminology and in courses expecting students to engage in academic activities less common in their core courses. Offering academic support services to students in such courses may be another strategy for reducing the challenges of the transition.

Of course, there are limits to the generalizability of this study arising from looking at a single discipline at a single institution. Indeed, the impact of the transition to English-only instruction is likely subject to several factors that vary across contexts. However, the nature of the phenomenon under study is inherently single-institution-confined (Macaro, 2018). To the

best of our knowledge, there are no large-scale initiatives to switch to English-only instruction nationwide. Therefore, because English-only instruction is typically implemented at the institution level in non-English speaking countries, future research is also likely to remain restricted to single institutions. To address this limitation, future work on this topic could involve more case studies from various geographic and institutional contexts. We would particularly recommend that future studies of the effects of English-only instruction on students' outcomes utilize rigorous methodologies, where possible, to contribute to building a robust evidence base regarding the causal effects of implementing English-only instruction. We hope that this study may serve as a model for such future efforts.

References

Ackerley, K., Guarda, M., & Helm, F. (2017). *Sharing perspectives on English-medium instruction*. Bern, Switzerland: Peter Lang.

Airey, J., Lauridsen, K. M., Räsänen, A., Salö, L., & Schwach, V. (2017). The expansion of English-medium instruction in the Nordic countries: Can top-down university language policies encourage bottom-up disciplinary literacy goals? *Higher Education, 73*(4), 561-576.

Bolton, K., Botha, W., & Bacon-Shone, J. (2017). English-medium instruction in Singapore higher education: Policy, realities and challenges. *Journal of Multilingual and Multicultural Development*, 1-18.

Bradford, A. (2016). Toward a typology of implementation challenges facing English-medium instruction in higher education: Evidence from Japan. *Journal of Studies in International Education, 20*(4), 339-356.

Byun, K., Chu, H., Kim, M., Park, I., Kim, S., & Jung, J. (2011). English-medium teaching in Korean higher education: Policy debates and reality. *Higher Education, 62*(4), 431-449.

Crawford, J. (1992). *Hold your tongue: Bilingualism and the politics of English only*. Boston, MA: Addison Wesley Publishing Company.

Dearden, J. (2014). *English as a medium of instruction–a growing global phenomenon*. London, UK: British Council.

Doiz, A., Lasagabaster, D., & Sierra, J. (2011). Internationalisation, multilingualism and English-medium instruction. *World Englishes, 30*(3), 345-359.

Doiz, A., Lasagabaster, D., & Sierra, J. M. (2013). *English-medium instruction at universities: Global challenges*. Bristol, UK: Multilingual matters.

Drljača Margić, B., & Vodopija-Krstanović, I. (2017). *Uncovering English-medium instruction: Glocal issues in higher education*. Oxford, UK: Peter Lang.

Evans, S., & Morrison, B. (2011). Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. *English for Specific Purposes, 30*(3), 198-208.

Goodman, B. A. (2014). Implementing English as a medium of instruction in a Ukrainian University: Challenges, adjustments, and opportunities. *International Journal of Pedagogies and learning*, *9*(2), 130-141.

Hellekjær, G. O. (2010). Lecture comprehension in English-medium higher education. *Hermes–Journal of Language and Communication Studies, 45*, 11-34.

Hu, G., & Lei, J. (2014). English-medium instruction in Chinese higher education: A case study. *Higher Education, 67*(5), 551-567.

Imbens, G., & Rubin, D. B. (2015). Causal inference for statistics, social, and behavioral sciences. New York, NY: Cambridge University Press.

Koretz, D. M., Linn, R. L., Dunbar, S. B., & Shepard, L. A. (1991). *The effects of high-stakes testing on achievement: Preliminary findings about generalization across tests.* Paper presented at the American Educational Research Association and the National Council on Measurement in Education, Chicago, IL.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4-16.

Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching*, *51*(1), 36-76.

Moed, H. F., Markusova, V., & Akoev, M. (2018). Trends in Russian research output indexed in Scopus and Web of Science. *Scientometrics, 116*(2), 1153-1180.

Murnane, R. J., & Willett, J. B. (2011). *Methods matter: Improving causal inference in*

educational and social science research. New York, NY: Oxford University Press.

Nguyen, H. T., Hamid, M. O., & Moni, K. (2016). English-medium instruction and self-governance in higher education: The journey of a Vietnamese university through the institutional autonomy regime. *Higher Education, 72*(5), 669-683.

Parr, C. (2014). English language use 'most significant internationalisation trend for HE'. *Times Higher Education,* Retrieved from: https://www.timeshighereducation.com/news/english-language-use-most-significant-internationalisation-trend-for-he/2013009.article

Salmi, J. (2009). *The challenge of establishing world-class universities*. Washington, D.C.: The World Bank Publications.

Sert, N. (2008). The language of instruction dilemma in the Turkish context. *System, 36*(2), 156-171.

Tan, P. K. W., & Rubdy, R. (2008). *Language as commodity: Global structures, local marketplaces*. New York, NY: Continuum.

Tollefson, J. W., & Tsui, A. B. M. (2004). *Medium of instruction policies: Which agenda? Whose agenda?* Mahwah, NJ: Lawrence Erlbaum Associates.

Wächter, B., & Maiworm, F. (2014). *English-taught programmes in European higher education: The state of play in 2014*. Bonn, Germany: Lemmens Medien GmbH.

Wilkinson, R. (2017). Trends and issues in English-medium instruction in Europe. In K. Ackerley, M. Guarda, & F. Helm (Eds.), *Sharing perspectives on English-medium instruction*. Bern, Switzerland: Peter Lang.

World Bank. (2016). *Polarization and populism: Europe and Central Asia economic update*. Washington, DC: World Bank.

Zhao, J., & Dixon, L. Q. (2017). *English-medium instruction in Chinese universities:*

Perspectives, discourse and evaluation. New York, NY: Routledge.

Tables and figures

Table 1. *Descriptive statistics (2007-2012 cohorts)*

	Treatment group: Mean (SD)			Comparison group: Mean (SD)			DiD	
	Before	After	Diff.	Before	After	Diff.	Pooled	2010
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Third year retention	0.80 (0.02)	0.91 (0.02)	0.10*** (0.03)	0.85 (0.01)	0.93 (0.01)	0.08*** (0.01)	0.02 (0.03)	-0.02 (0.04)
Dropout	0.24 (0.02)	0.21 (0.02)	-0.03 (0.03)	0.23 (0.01)	0.16 (0.01)	-0.07*** (0.02)	0.04 (0.03)	0.11* (0.05)
Failed course credits	5.13 (9.71)	10.09 (12.72)	4.96*** (0.85)	8.13 (13.78)	9.09 (13.19)	0.96 (0.58)	4.00*** (1.12)	7.31*** (1.54)
GPA	3.02 (0.81)	2.79 (0.76)	-0.23*** (0.06)	2.89 (0.93)	2.87 (0.81)	-0.02 (0.04)	-0.22** (0.07)	-0.49*** (0.11)
Graduated in 5 years	0.76 (0.43)	0.73 (0.45)	-0.03 (0.03)	0.77 (0.42)	0.80 (0.40)	0.03 (0.02)	-0.07~ (0.04)	-0.12* (0.05)
Female	0.38 (0.48)	0.39 (0.48)	0.01 (0.04)	0.47 (0.50)	0.51 (0.50)	0.04 (0.02)	-0.02 (0.04)	-0.05 (0.06)
Rural	0.14 (0.35)	0.19 (0.40)	0.05 (0.03)	0.13 (0.34)	0.13 (0.34)	0.002 (0.01)	0.05 (0.03)	0.03 (0.04)
Local ethnicity	0.89 (0.31)	0.92 (0.27)	0.03 (0.02)	0.91 (0.29)	0.92 (0.27)	0.01 (0.01)	0.01 (0.02)	0.04 (0.03)
Local language high school	0.46 (0.50)	0.58 (0.49)	0.11* (0.04)	0.40 (0.49)	0.35 (0.48)	-0.05* (0.02)	0.16*** (0.04)	0.04 (0.06)
College admission test score	94.90 (13.67)	89.72 (10.02)	-5.18*** (0.89)	92.06 (15.59)	85.98 (12.54)	-6.09*** (0.61)	0.91 (1.18)	-0.93 (1.71)
State grant recipient	0.60 (0.49)	0.62 (0.49)	0.02 (0.04)	0.32 (0.47)	0.27 (0.45)	-0.05* (0.02)	0.07~ (0.04)	-0.06 (0.06)
Anon U grant recipient	0.10 (0.30)	0.06 (0.24)	-0.03 (0.02)	0.02 (0.16)	0.04 (0.19)	0.02 (0.01)	-0.05* (0.02)	-0.01 (0.02)
Self-supported student	0.30 (0.46)	0.31 (0.47)	0.01 (0.03)	0.66 (0.48)	0.69 (0.46)	0.03 (0.02)	-0.02 (0.04)	0.07 (0.06)
Transferred in	0 (0)	0.003 (0.5)	0.003 (0.003)	0.01 (0.09)	0.05 (0.21)	0.04*** (0.01)	-0.03** (0.01)	-0.02 (0.01)
Observations	352	359		1,177	997		2,884	2,009

~p<0.10, *p<0.05, **p<0.01, ***p<0.001

Note. Column 7 presents the difference-in-differences estimates on each of the baseline characteristics using the full sample. Column 8 presents difference-in-differences estimates using only 2010 as the post-treatment period. The difference-in-differences estimates are based on the equation 1 specification but do not include other covariates.

Table 2. *Impacts of the switch to English-only instruction on retention, probability of dropping out, number of failed course credits, GPA, and probability of five-year graduation at Anon U (2007-2012 cohorts)*

	(1)	(2)	(3)	(4)	(5)
	Third year retention	Probability of dropping out	Number of failed course credits	GPA	Probability of five-year graduation
Post * English	0.01	0.04	3.03**	-0.13*	-0.07~
	(0.03)	(0.03)	(1.05)	(0.07)	(0.04)
Post	0.18***	-0.25***	-2.46**	0.46***	0.14***
	(0.02)	(0.03)	(0.92)	(0.06)	(0.03)
English	0.03	-0.10	2.82	-0.22	0.10
	(0.07)	(0.08)	(2.57)	(0.16)	(0.09)
Observations	2,884	2,884	2,884	2,884	2,884
R-squared	0.06	0.08	0.20	0.25	0.09
Student characteristics	Yes	Yes	Yes	Yes	Yes

~p<0.10, *p<0.05, **p<0.01, ***p<0.001

Note. The parameter estimates for Post * English show the impact of the switch to English-only instruction on the outcomes. Student characteristics include student gender, locality students come from, ethnicity, language of instruction at high school, national standardized college admission test score, financial aid status at entry to Anon U, indicator for whether a student took a year off during studies at Anon U, indicator for whether a student transferred in to Anon U from another institution, and number of years at Anon U.

Table 3. *Sensitivity of results to choice of comparison schools (2007-2012 cohorts)*

Outcome	All other schools	School of Engineering	School of Chemistry	School of Business
Third year retention	0.01 (0.03)	0.04 (0.03)	-0.06 (0.04)	-0.00 (0.03)
Probability of dropping out	0.04 (0.03)	0.01 (0.02)	0.04 (0.08)	0.05 (0.03)
Number of failed course credits	3.03** (1.05)	1.49~ (0.86)	3.30* (1.93)	3.64** (1.19)
GPA	-0.13* (0.07)	-0.11~ (0.06)	-0.11 (0.10)	-0.13~ (0.07)
Probability of five-year graduation	-0.07~ (0.04)	-0.02 (0.02)	-0.17*** (0.04)	-0.08* (0.03)
Observations	2,884	1,499	900	1,907

~p<0.10, *p<0.05, **p<0.01, ***p<0.001

Note. All regressions include student characteristics and fixed effects for college major and cohort. Student characteristics include student gender, locality students come from, ethnicity, language of instruction at high school, national standardized college admission test score, financial aid status at entry to Anon U, indicator for whether a student took a year off during studies at Anon U, indicator for whether a student transferred in to Anon U from another institution, and number of years at Anon U.

Table 4. *Fading out of the treatment effect across the post-treatment period*

	Post-treatment period used in estimating the models			
	All post-treatment years	First post-treatment year only	Second post-treatment year only	Third post-treatment year only
Outcome	2010-2012	2010	2011	2012
Third year retention	0.01 (0.03)	-0.01 (0.04)	0.05 (0.04)	-0.06 (0.05)
Probability of dropping out	0.04 (0.03)	0.08 (0.05)	0.01 (0.05)	0.07 (0.05)
Number of failed course credits	3.03** (1.05)	5.66*** (1.33)	-0.20 (1.10)	3.34* (1.49)
GPA	-0.13* (0.07)	-0.32*** (0.06)	0.02 (0.06)	-0.10 (0.08)
Probability of five-year graduation	-0.07~ (0.04)	-0.09* (0.03)	-0.05 (0.03)	-0.10 (0.06)
Observations	2,884	2,008	2,001	1,931

~p<0.10, *p<0.05, **p<0.01, ***p<0.001

Note. All regressions include student characteristics and fixed effects for college major and cohort. Student characteristics include student gender, locality students come from, ethnicity, language of instruction at high school, financial aid status at entry to Anon U, indicator for whether a student took a year off during studies at Anon U, indicator for whether a student transferred in to Anon U from another institution, and number of years at Anon U.

Table 5. First semester grades before and after the switch to English-only instruction in the School of Computer Science

	Treatment group: Mean (SD)			Comparison group: Mean (SD)			DiD	
	Before	After	diff.	Before	After	diff.	Pooled	2010
Course:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Math	72.14	67.35	-4.79**	58.69	55.40	-3.29**	-1.50	4.17
	(1.02)	(1.12)	(1.51)	(0.69)	(0.78)	(1.04)	(2.07)	(2.89)
Physics	73.23	72.10	-1.13	68.93	65.82	-3.10*	1.98	6.95*
	(0.96)	(0.95)	(1.36)	(0.91)	(0.92)	(1.30)	(1.95)	(2.77)
Language	82.34	81.62	-0.72	79.57	80.17	0.59	-1.31	-2.67
	(0.81)	(0.82)	(1.15)	(0.56)	(0.50)	(0.75)	(1.54)	(2.21)
History	80.49	80.61	0.12	73.10	79.64	6.54***	-6.43**	-5.79~
	(1.07)	(0.88)	(1.39)	(0.79)	(0.55)	(1.01)	(1.94)	(2.98)
Observations	352	359		1,177	997		2,884	2,009

~p<0.10, *p<0.05, **p<0.01, ***p<0.001

Note. All difference-in-differences regressions include student characteristics and fixed effects for college major and cohort. Student characteristics include student gender, locality students come from, ethnicity, language of instruction at high school, financial aid status at entry to Anon U, indicator for whether a student took a year off during studies at Anon U, indicator for whether a student transferred in to Anon U from another institution, and number of years at Anon U.

Table 6. Proportion of students by English proficiency at entry by cohort

	Cohort		
English proficiency at entry:	2010	2011	2012
	A. Treatment school		
Elementary / pre-intermediate	46%	46%	68%
Intermediate	45%	43%	16%
Upper-intermediate / advanced	8%	11%	16%
Observations	120	125	118
	B. Comparison schools		
Elementary / pre-intermediate	62%	64%	57%
Intermediate	30%	27%	20%
Upper-intermediate / advanced	8%	9%	23%
Observations	371	387	268

Table 7. The relationship between instructors' years of experience at Anon U and grades of students in the School of Computer Science

	Average grades
Years of experience at Anon U	0.04
	(0.18)
Constant	52.32***
	(8.41)
Observations	12,426
R-squared	0.08

˜p<0.10, *p<0.05, **p<0.01, ***p<0.001

Note. All regressions include student characteristics and fixed effects for college major and cohort. Student characteristics include student gender, locality students come from, ethnicity, language of instruction at high school, financial aid status at entry to Anon U, indicator for whether a student took a year off during studies at Anon U, indicator for whether a student transferred in to Anon U from another institution, and number of years at Anon U. OLS standard errors are clustered at major-cohort level. Unit of analysis is a grade received by a student in the courses taken during the studies at Anon U. The scale for course grades in 0-100.

Table 8. Instructors' years of experience at Anon U in courses with largest proportions of "Fail" grades

Cohort	Course (language of instruction)	% of "Fail" grades among students who took the course	Instructors' average years at Anon U
2009	Calculus 2 (Russian)	17%	2
	Probability theory and statistics (Russian)	15%	2
	English for Professional Purposes (English)	9%	4
2010	Calculus 2 (English)	24%	2
	Probability theory and statistics (English)	16%	3.5
	Basics of circuit theory (English)	10%	2
2011	Algorithms and data structures (English)	21%	2.5
	Databases (English)	20%	4
	Probability theory and statistics (English)	12%	3.5
2012	Differential equations (English)	17%	5
	Principles of economics (English)	15%	6
	Algorithms and data structures (English)	13%	3

Figure 1. *Testing parallel trends assumption: proportion of female students in the treatment and comparison groups by cohort.*

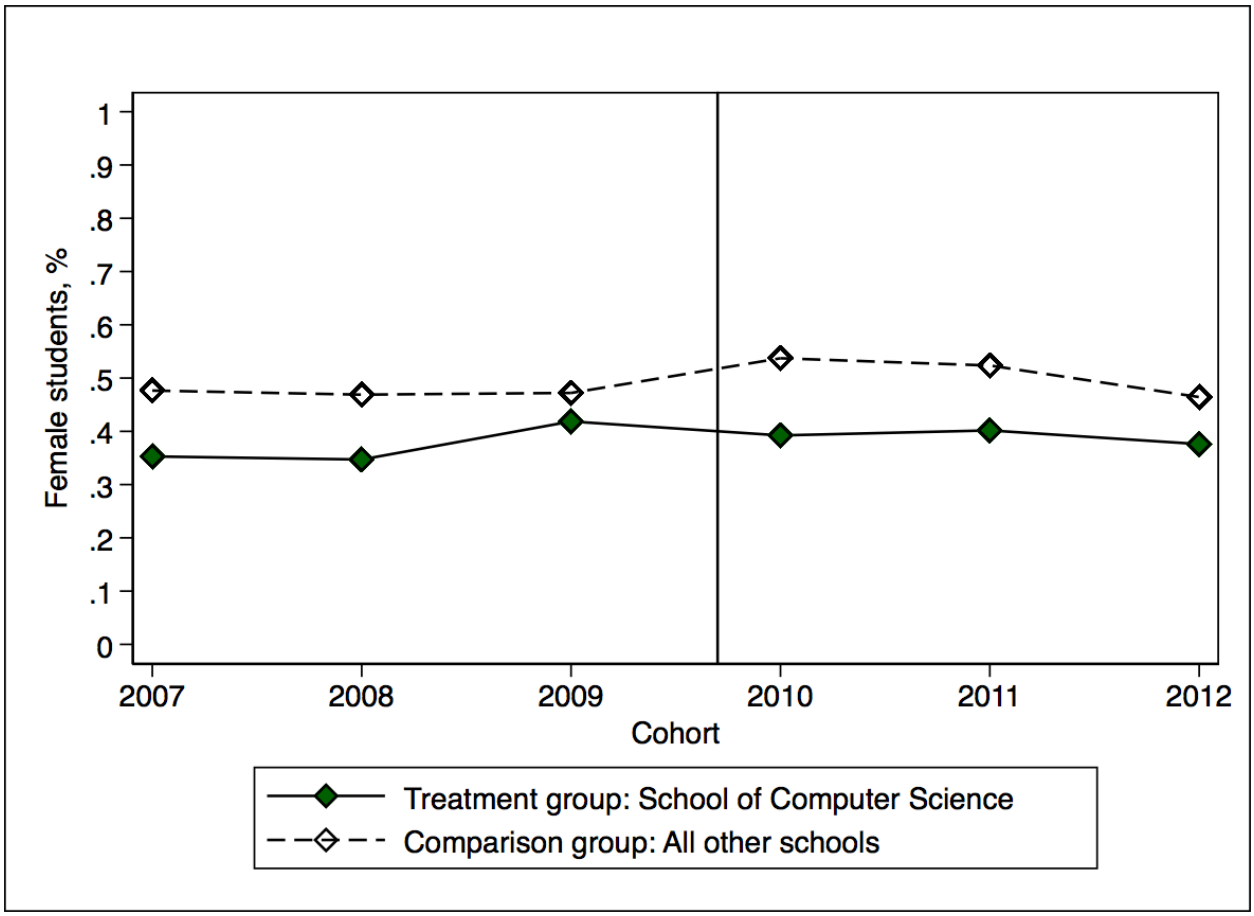

Figure 2. *Testing parallel trends assumption: proportion of rural students in the treatment and comparison groups by cohort.*

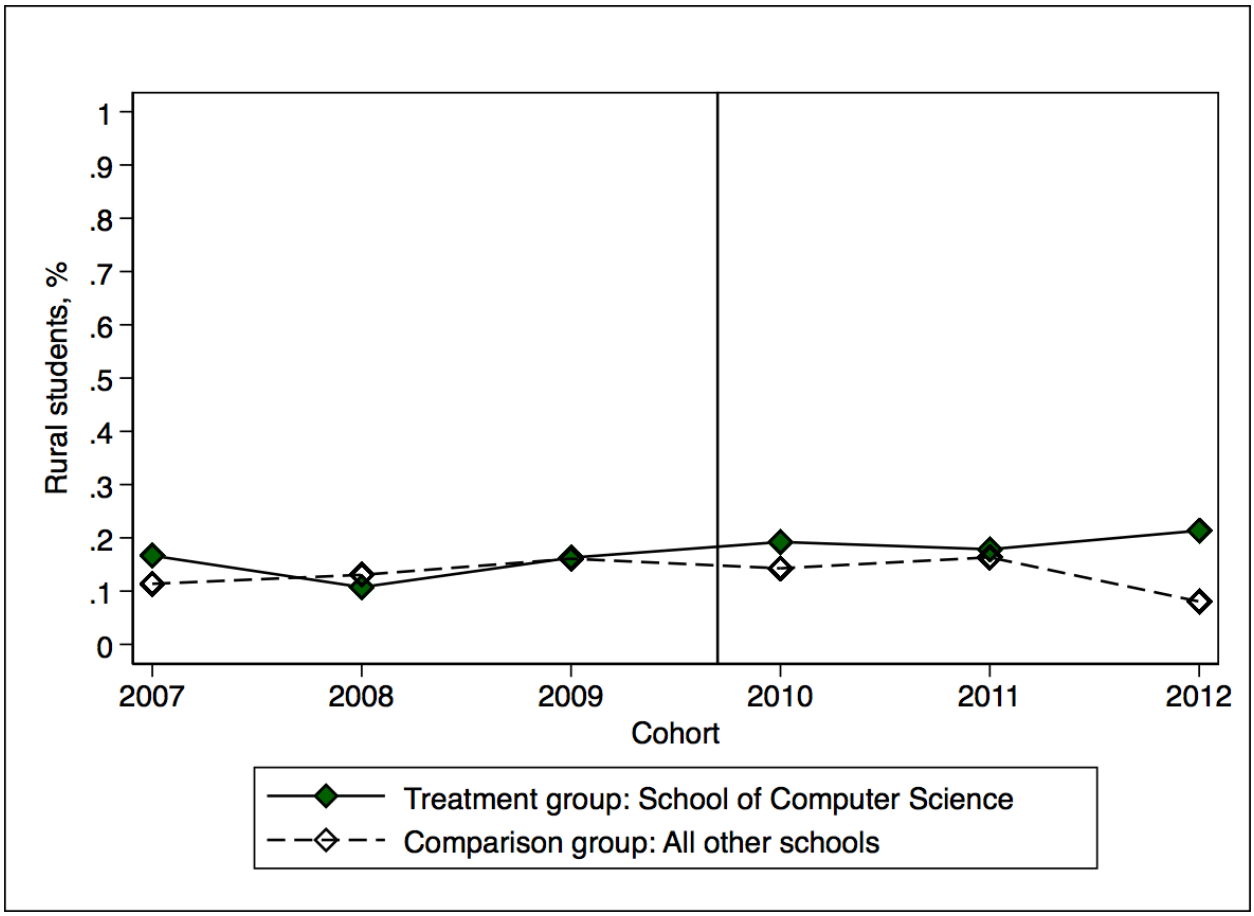

Figure 3. *Testing parallel trends assumption: proportion of local ethnicity students in the treatment and comparison groups by cohort.*

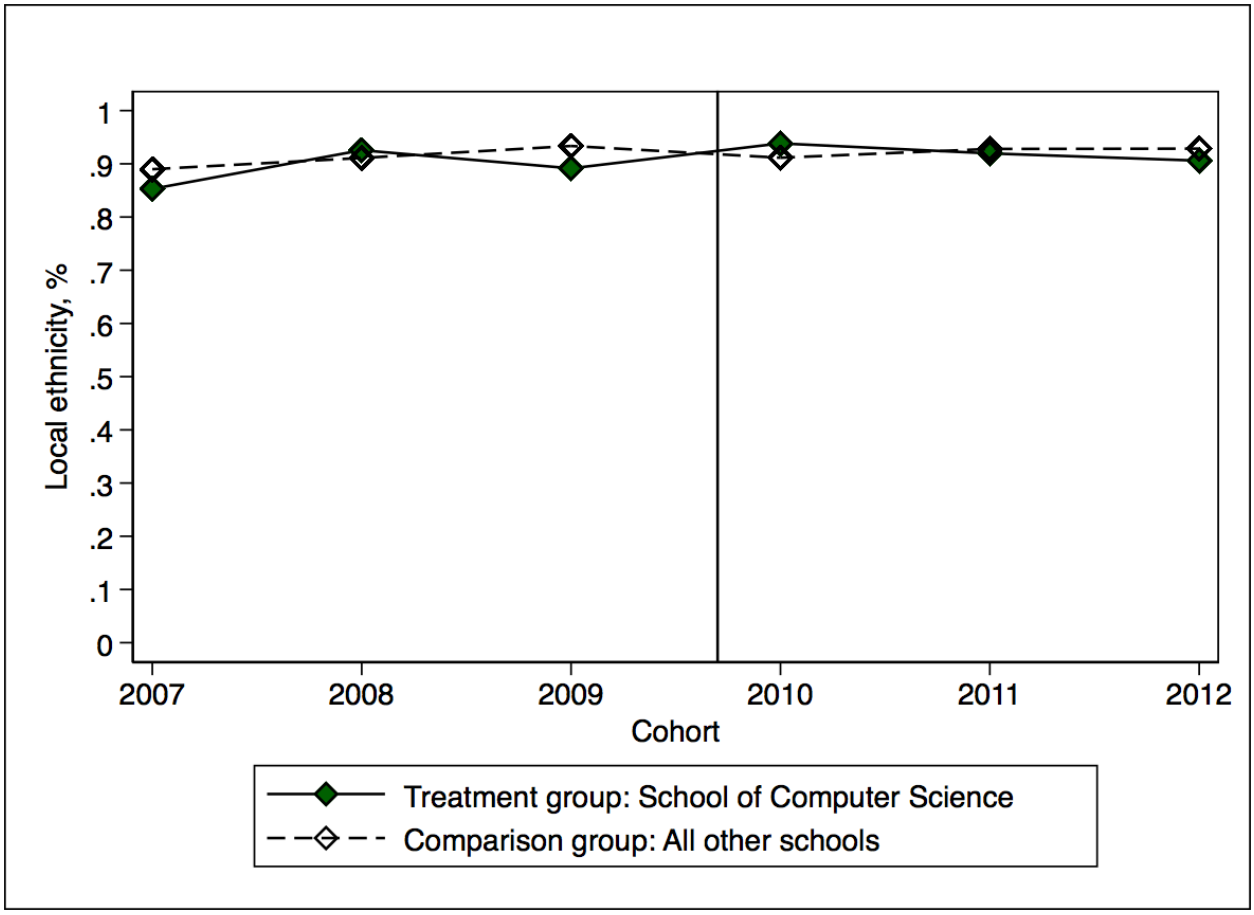

Figure 4. *Testing parallel trends assumption: school-leaving exam score in the treatment and comparison groups by cohort.*

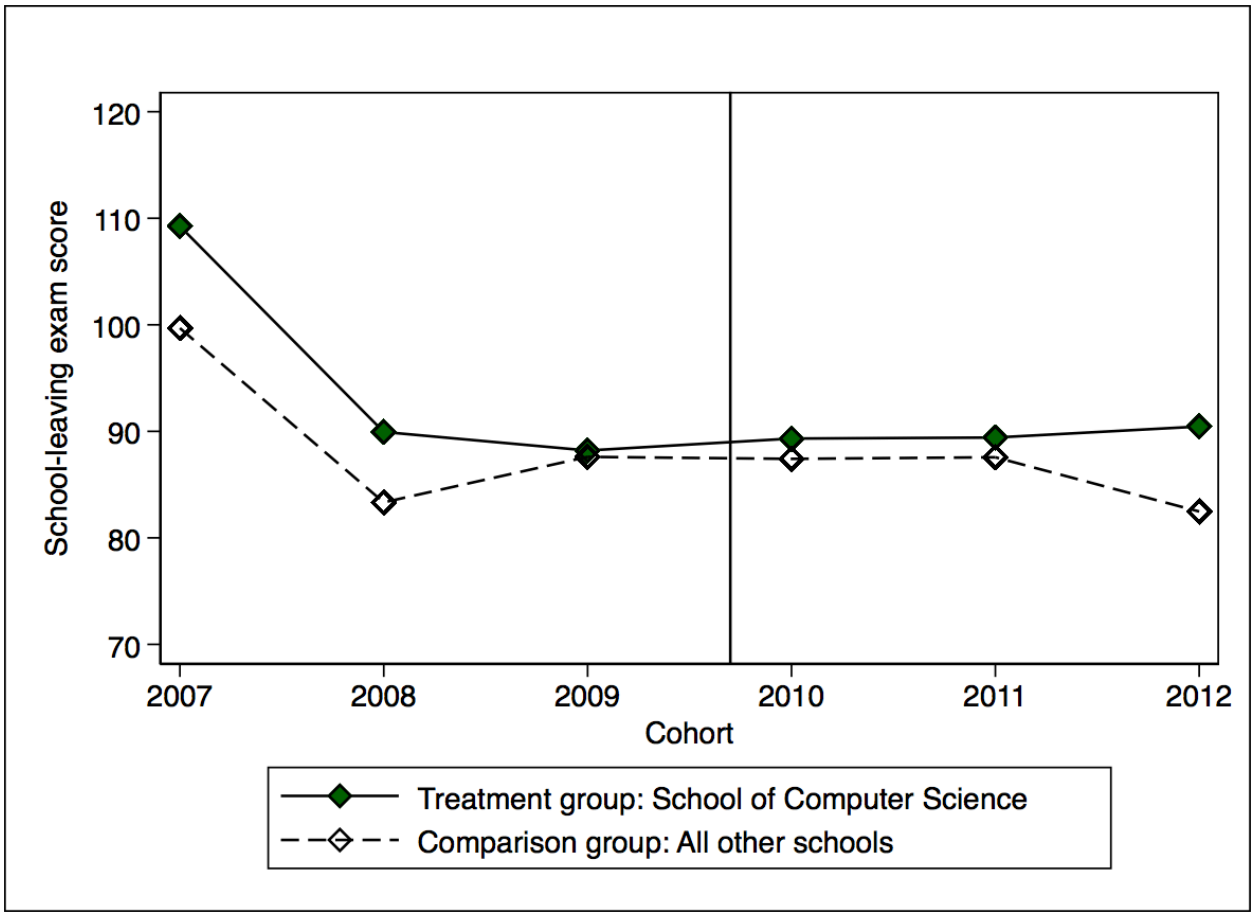

Figure 5. *Testing parallel trends assumption: proportion of Anon U grant recipients in the treatment and comparison groups by cohort.*

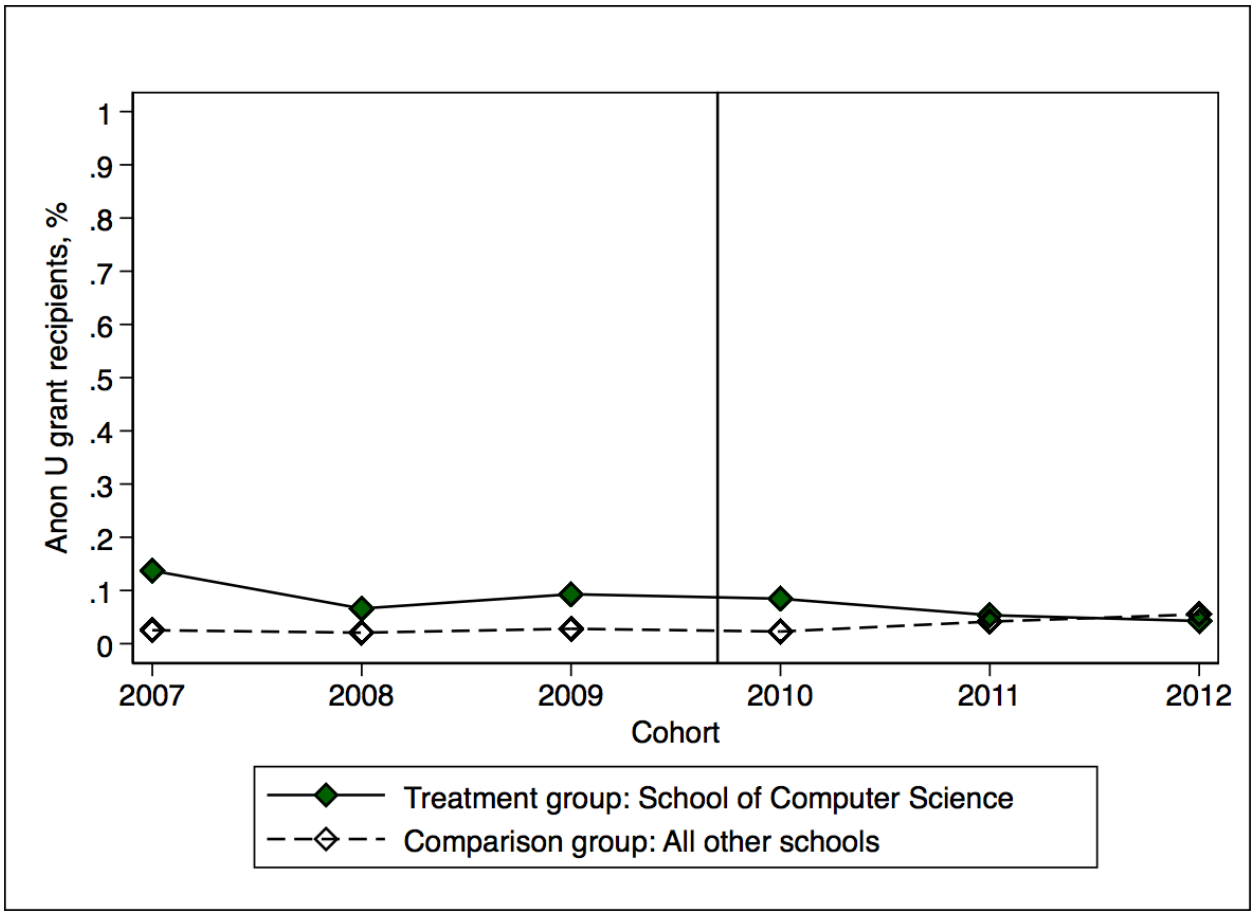

Figure 6. *Testing parallel trends assumption: proportion of students who transferred in from other colleges in the treatment and comparison groups by cohort.*

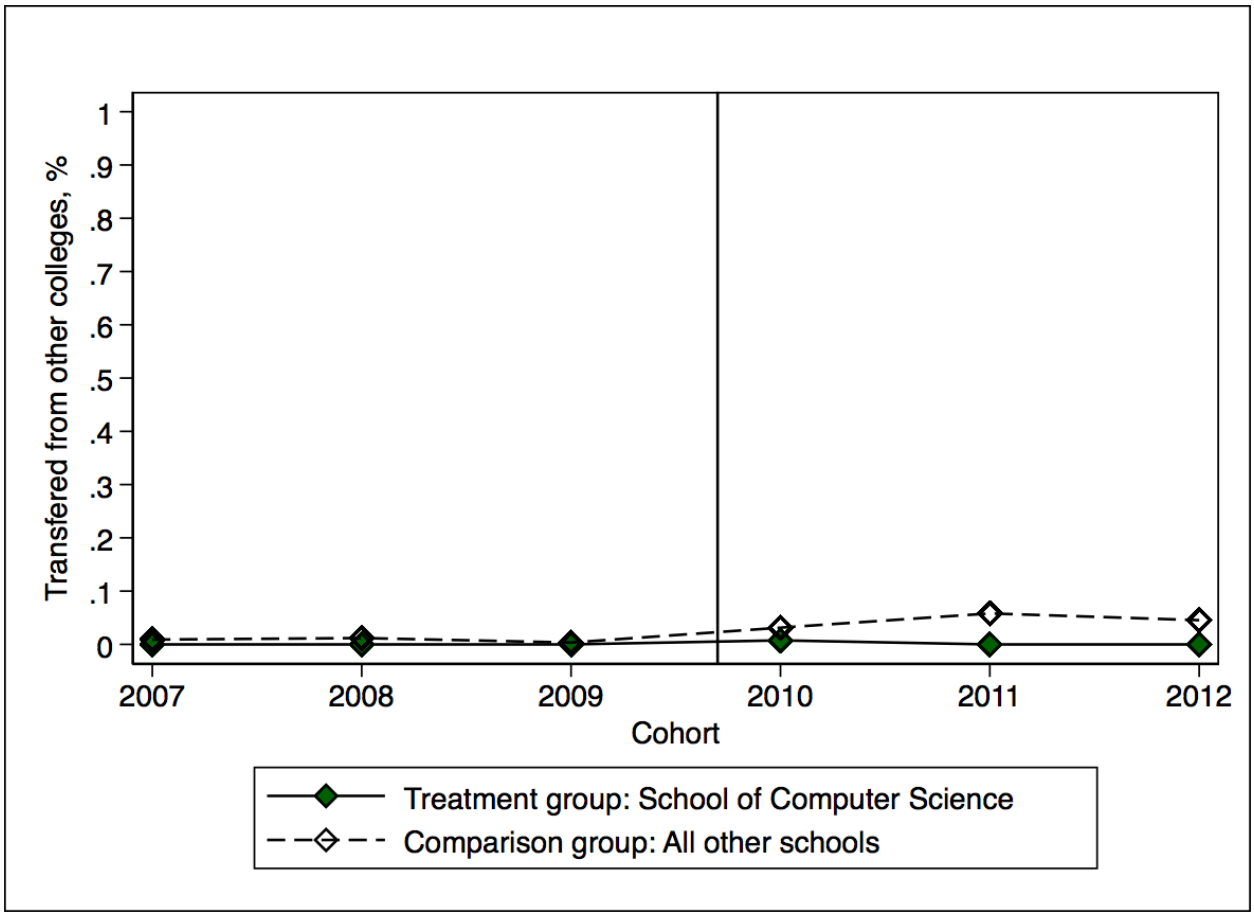

Figure 7. *Testing parallel trends assumption: proportion of students from local language high schools in the treatment and comparison groups by cohort.*

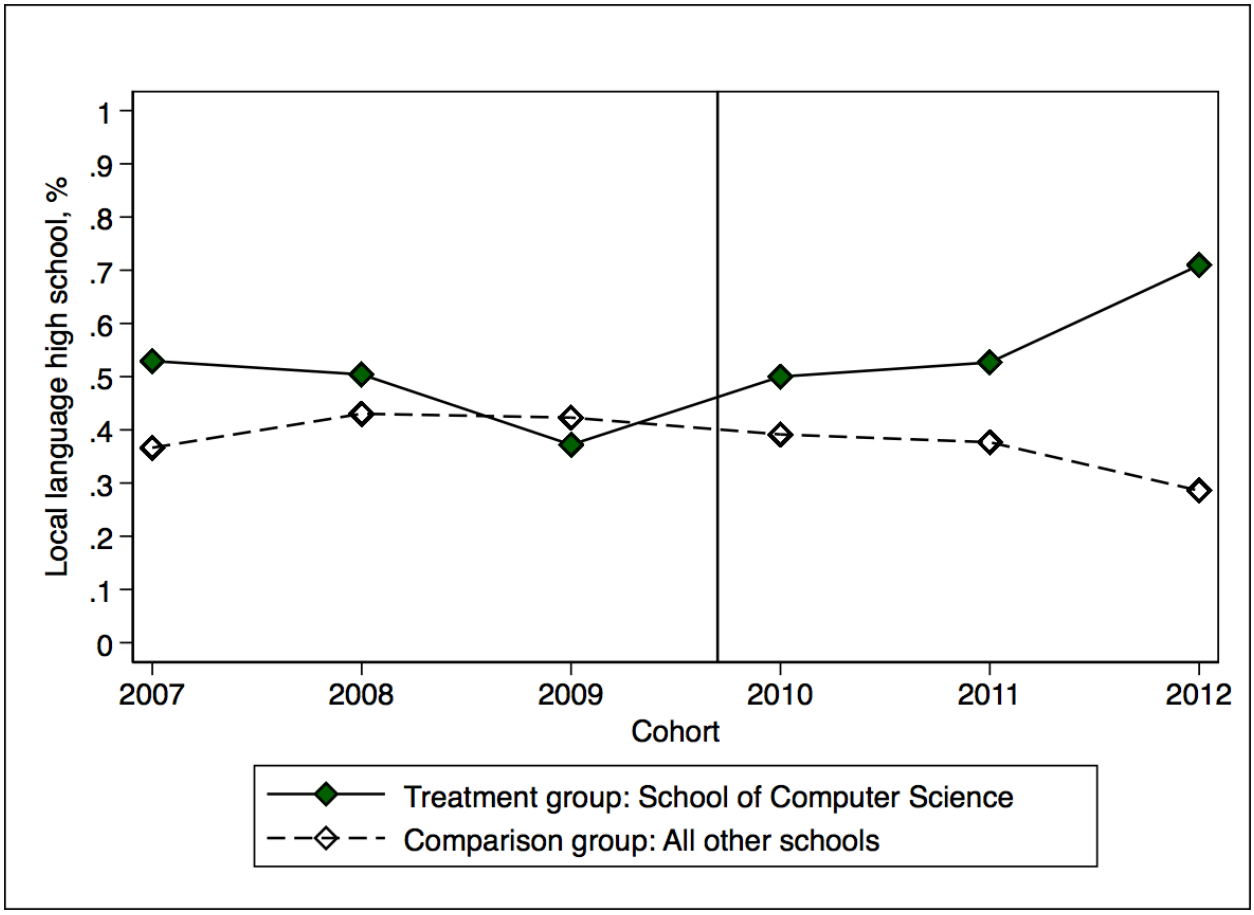

Figure 8. *Testing parallel trends assumption: proportion of state grant recipients in the treatment and comparison groups by cohort.*

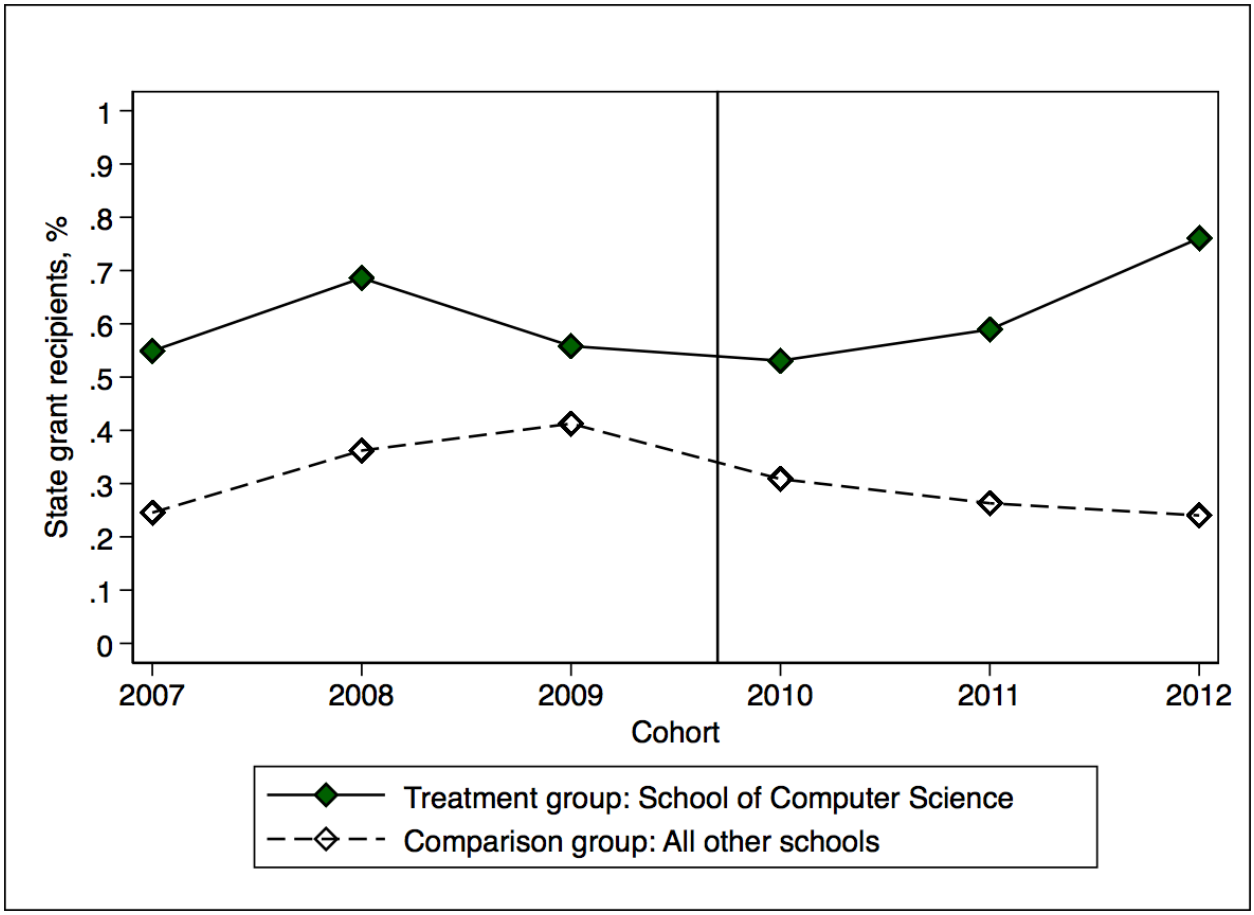

Figure 9. *Testing parallel trends assumption: proportion of self-supported students in the treatment and comparison groups by cohort.*

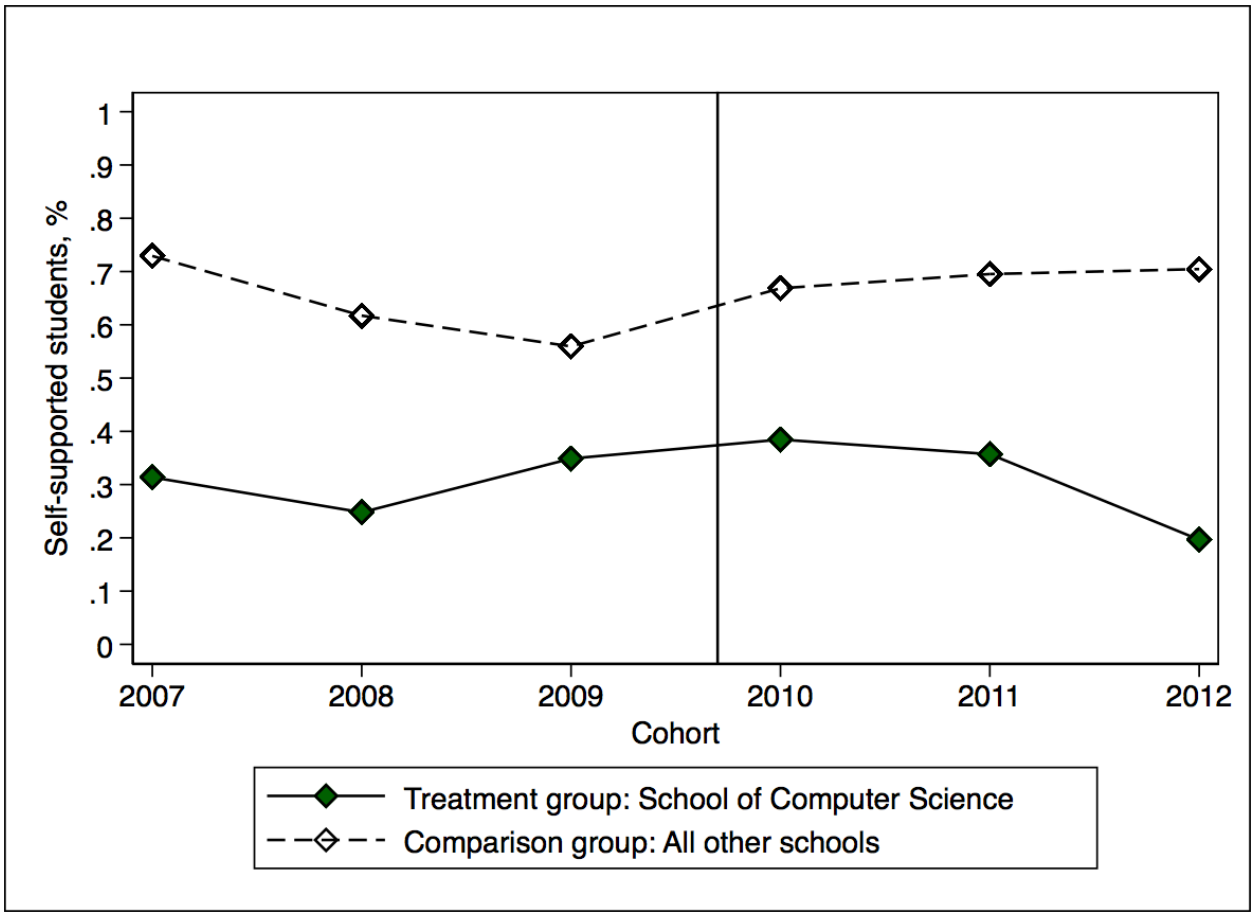

Figure 10. *Third year retention rate of the treatment and comparison groups by cohort.*

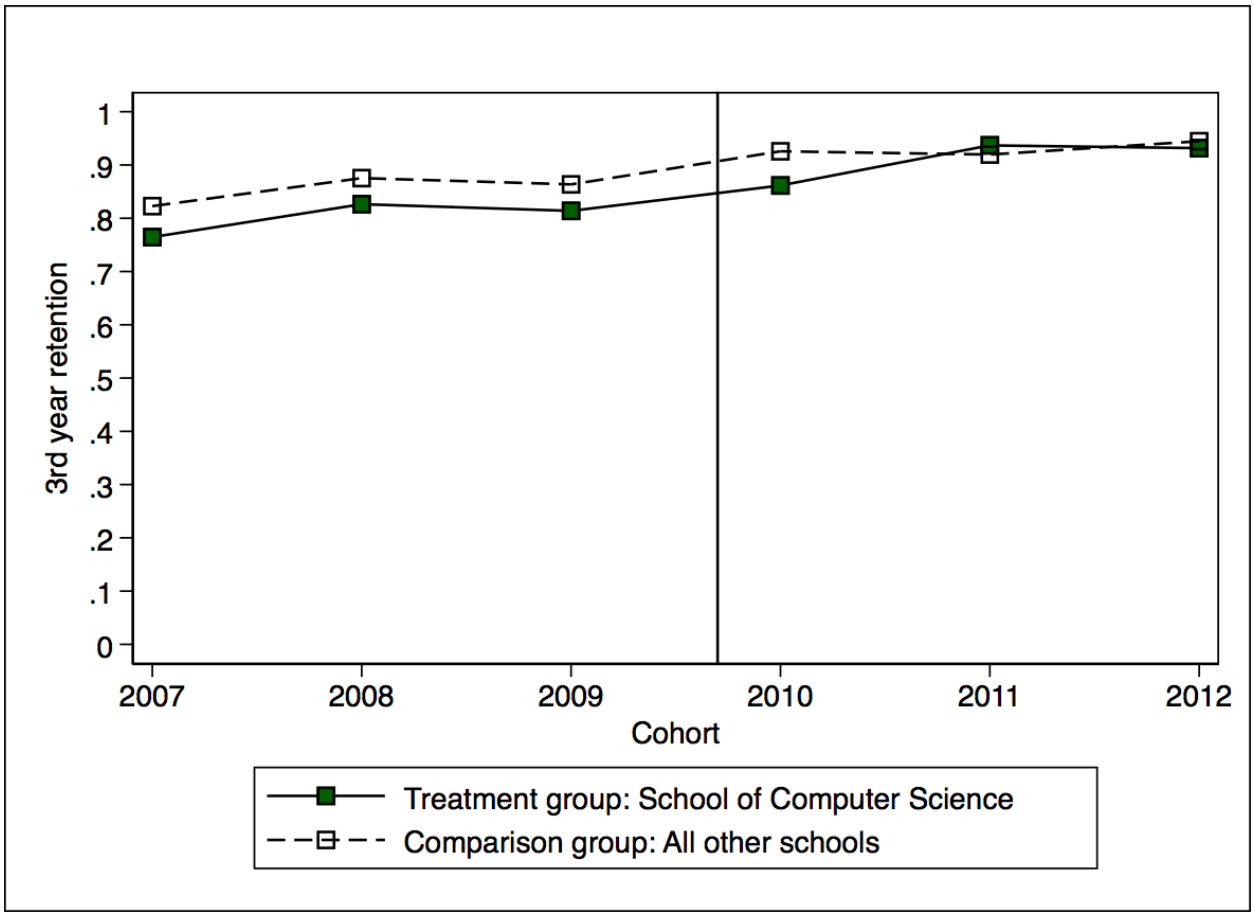

Figure 11. *Dropout rate of the treatment and comparison groups by cohort.*

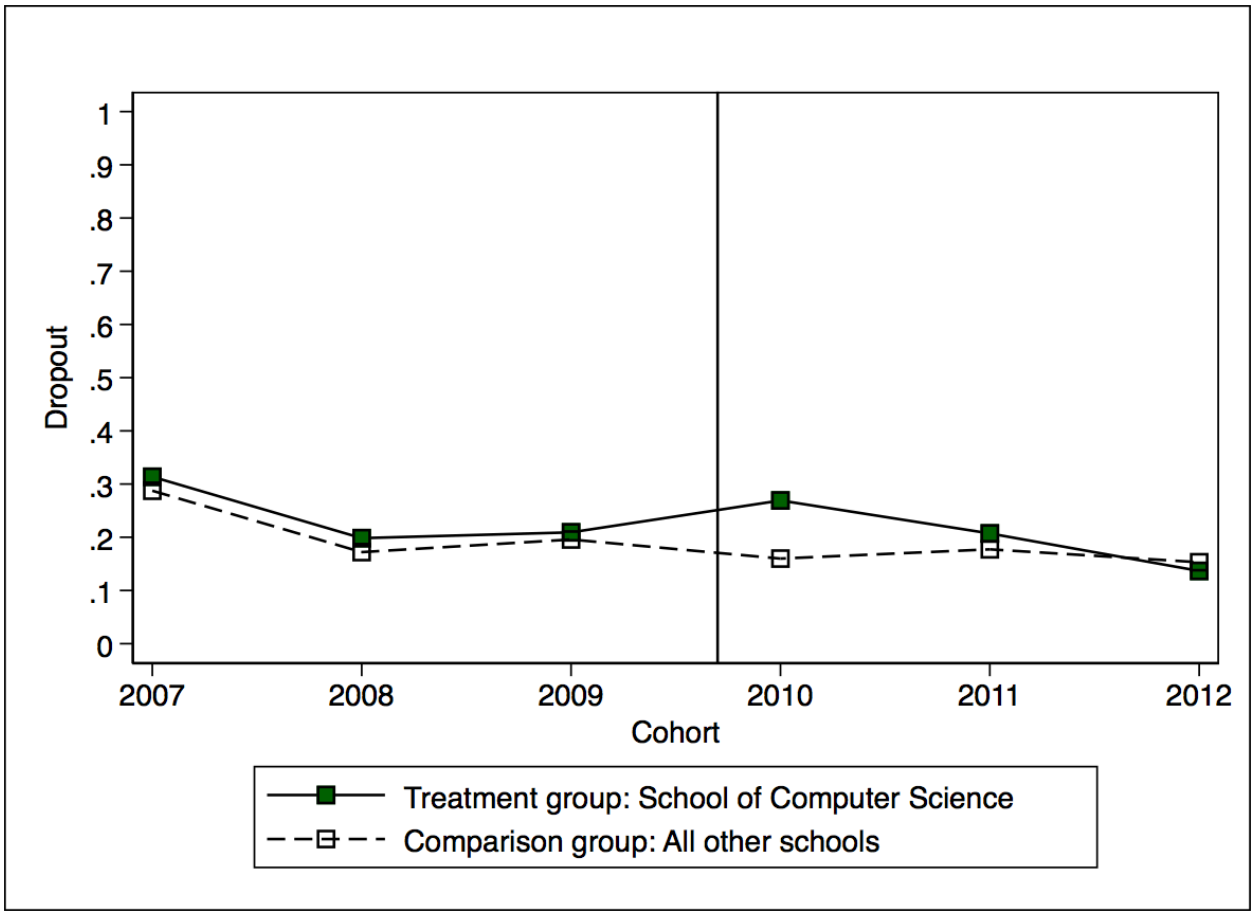

Figure 12. *Course credits failed in the treatment and comparison groups by cohort.*

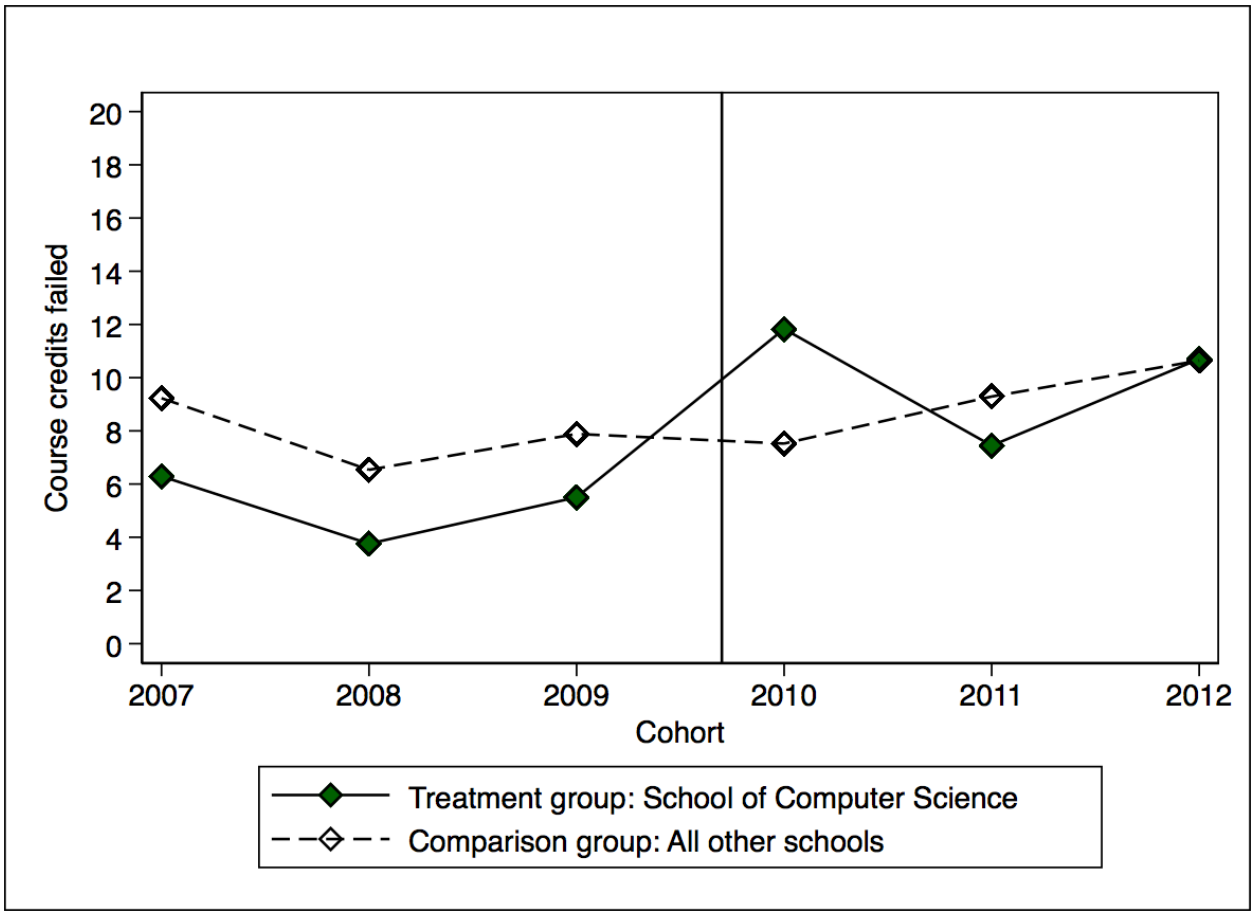

Figure 13. *GPA of the treatment and comparison groups by cohort.*

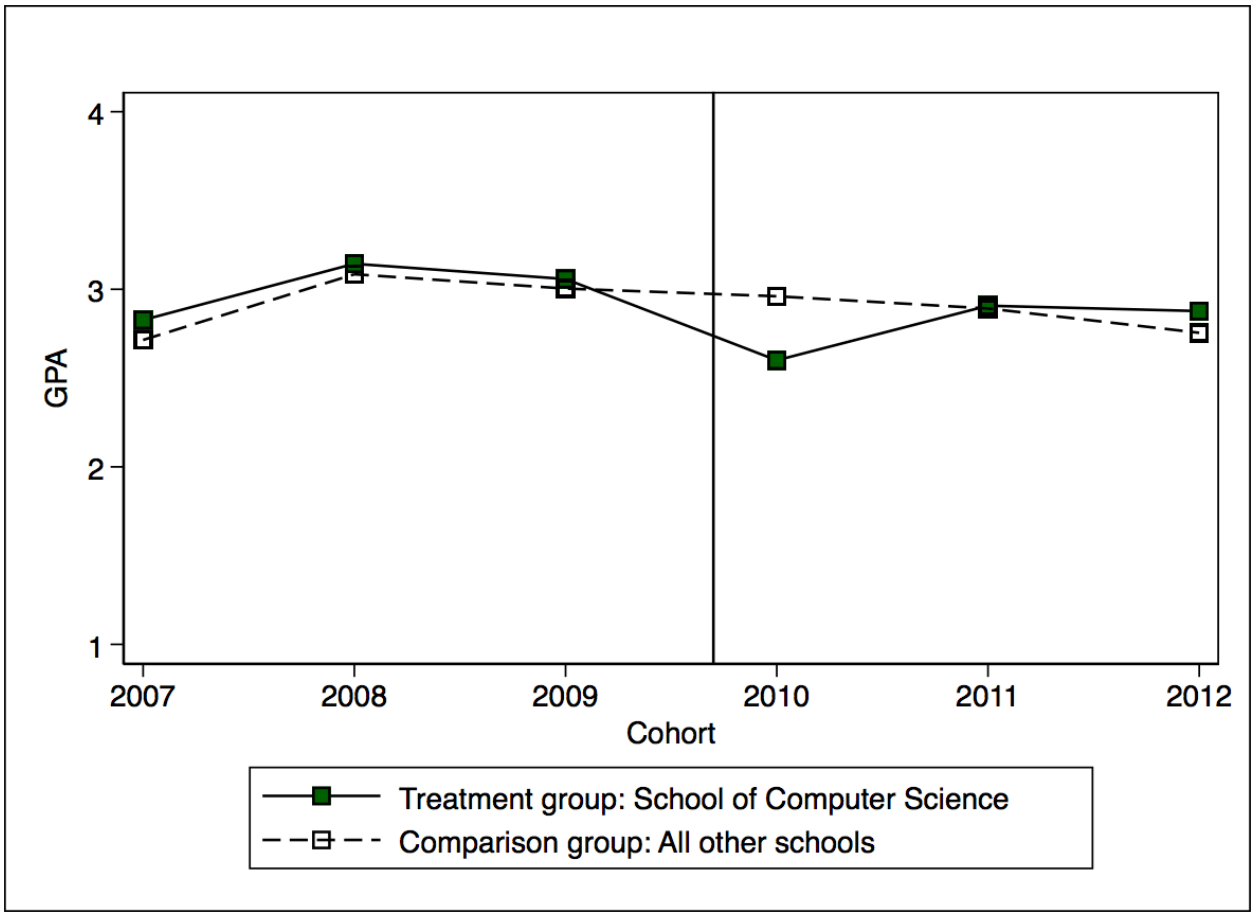

Figure 14. *Graduation rate of the treatment and comparison groups by cohort.*

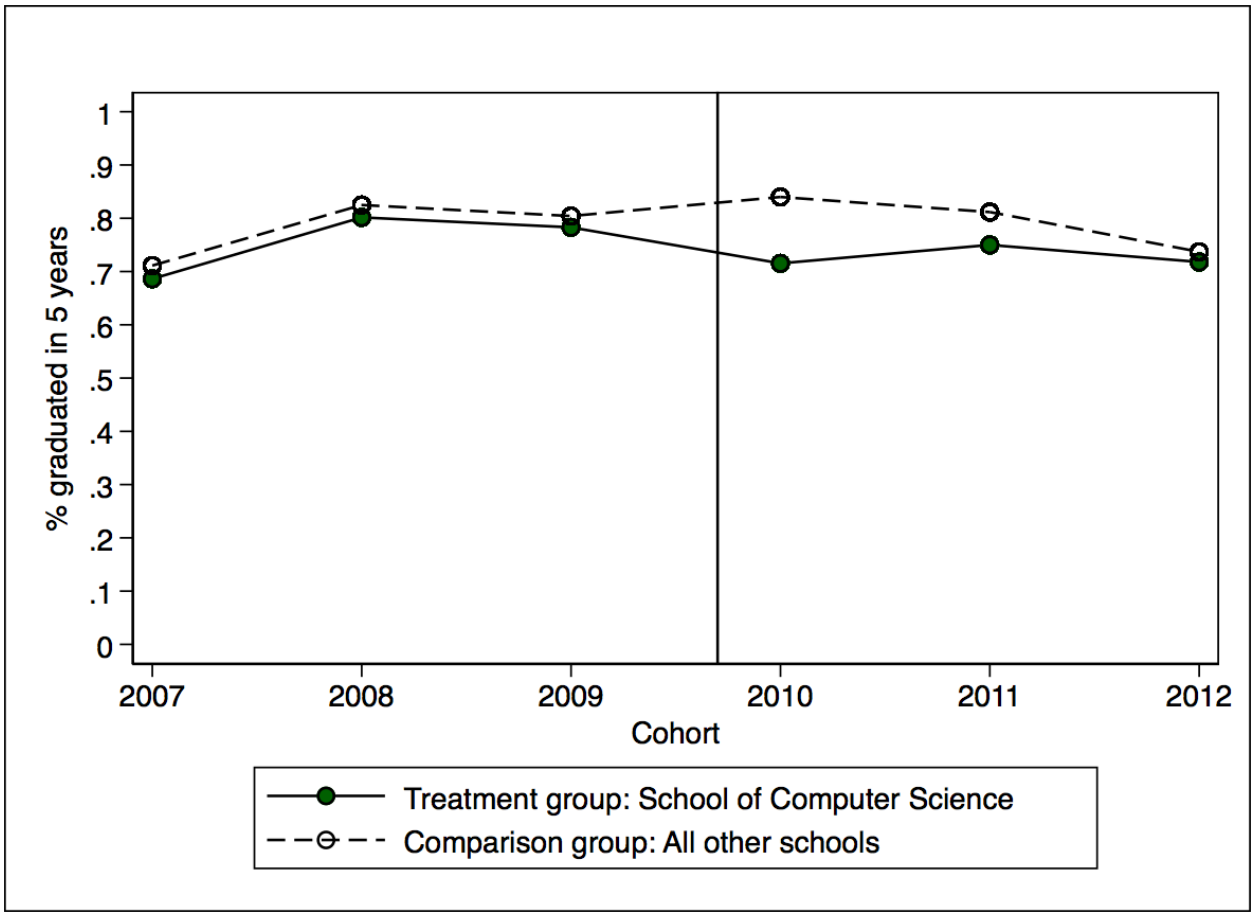

Figure 15.

Average grades in first semester courses in the School of Computer Science, by cohort.

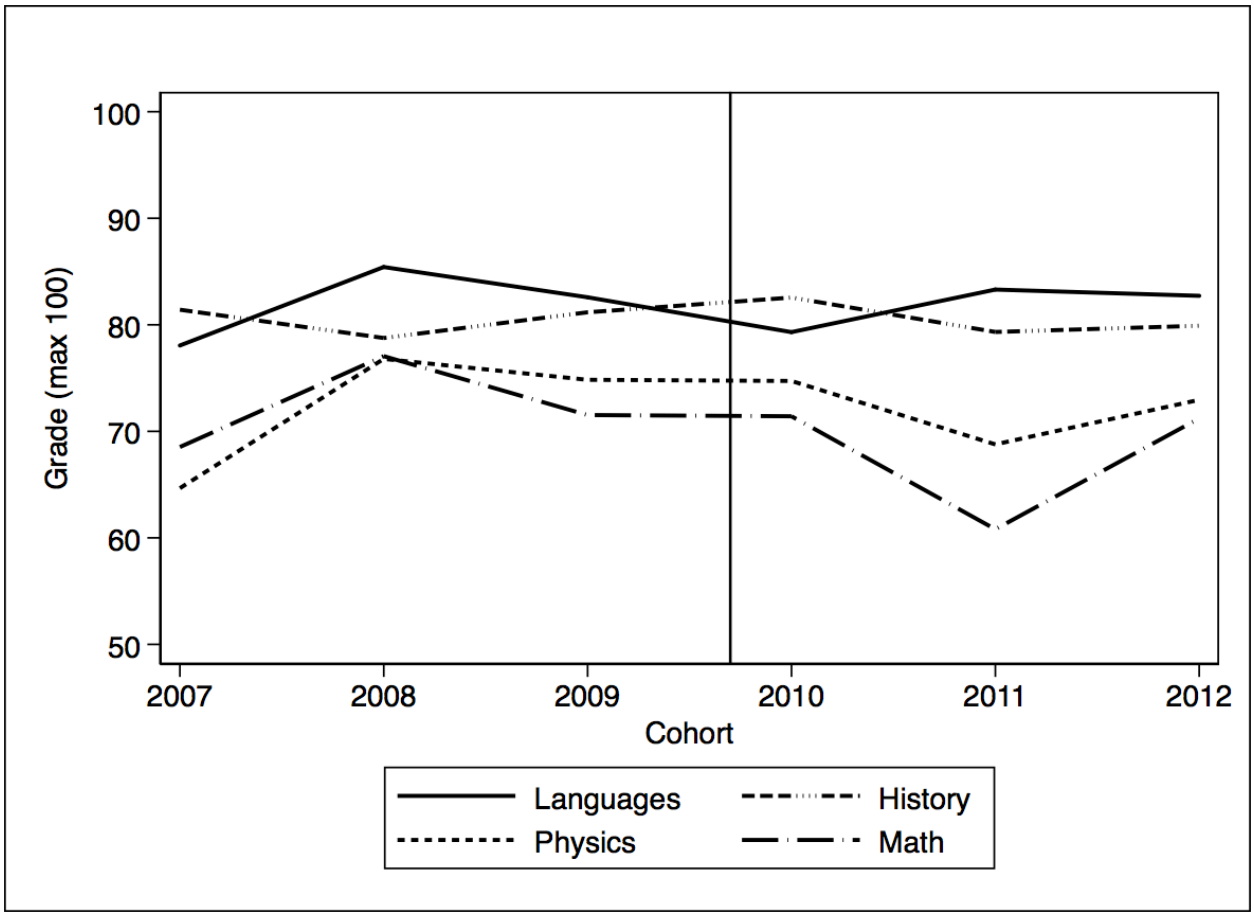

Figure 16. *Proportion of new instructors at Anon U by treatment status and academic year.*