Utrecht University

School of Economics

The Effect of Class Size on Grades and Course Evaluations

Evidence from Multi-section Courses

Alexei Karas

U.S.E. Working Paper Series nr: 19-03

Utrecht University School of Economics

U.S.E. Research Institute
Kriekenpitplein 21-22, 3584 EC Utrecht, The Netherlands Tel: +31 30 253 9800, e-mail: use.ri@uu.nl www.uu.nl/use/research

U.S.E. Research Institute
Working Paper Series 19-03

The Effect of Class Size on Grades and Course Evaluations: Evidence from Multi-section Courses

Alexei Karas
Department of Social Science
University College Roosevelt
Utrecht University School of Economics
Utrecht University

January 2019

Abstract

Using rich administrative data from a small Dutch liberal arts college, I study how the number of students enrolled in a course affects student grades and course evaluations. To identify the causal effect I exploit variation in class size across parallel sections of the same course taught by the same instructor in the same semester. I show that class size has a stronger negative effect on student grades in mandatory courses compared to electives. I show similar results for various components of student course evaluations: perceived overall course quality, perceived amount learned, student participation and engagement. I interpret these findings to be consistent with class size affecting educational outcomes through student engagement.

Keywords: Class size, Grades, Course evaluations, Student engagement

JEL classification: I21

Acknowledgements: I thank Remy Rikers for useful comments.

Comments welcomed to: a.karas@ucr.nl

1 Introduction

In this paper, I address the question of why college class size matters for educational outcomes. Using rich administrative data from a small Dutch liberal arts college, I study how the number of students enrolled in a course affects student grades and course evaluations. It is already well established that, on average, bigger classes lead to lower grades (Kokkelenberg et al., 2008, Bandiera et al., 2010, De Giorgi et al., 2012) and lower evaluations (Bedard and Kuhn, 2008, Mandel and Sussmuth, 2011, Sapelli and Illanes, 2016). What I seek to learn here is whether this effect systematically varies across courses with different characteristics.

In particular, I explore whether class size matters more in mandatory courses compared to electives. I find that the answer is "yes". In mandatory courses, class size has a stronger negative effect on student grades. In addition, it has a stronger effect on several components of student course evaluations: perceived overall course quality, perceived amount learned, student participation and engagement.

I motivate this empirical exercise as a test of a specific channel through which class size affects educational outcomes. I refer to it as "engagement channel". It works through student behavior: in smaller classes, students become more visible to the teacher and thus experience a stronger pressure to participate; alternatively, students enjoy a stronger sense of belonging to the group; both visibility and sense of belonging drive their academic and/or social engagement up (Finn et al., 2003). This increased engagement, in turn, leads to better educational outcomes.

To test this engagement channel I compare class size effects in mandatory courses and electives. My rationale is as follows. Students enroll in electives by choice; engagement in those courses must therefore be high regardless of class size.[1] In contrast, students often attend mandatory courses against their will, and instructors have to work hard to engage them. If class size affects student achievement through engagement, its effects should be particularly pronounced in mandatory courses: in those courses engagement starts low, and smaller class size helps instructors to build it.

To identify the causal effect of class size I use a fixed effects regression. Specifically, I regress final grades (or student evaluation scores) on the number of students enrolled in a section of a course controlling for various combinations of time, course, instructor and student fixed effects. In the most demanding specification, I control for time-course-instructor fixed effects, which amounts to only exploiting variation in class size across parallel sections of the *same* course taught by the *same* instructor in the *same* semester. Including these fixed effects allows me to rule out many alternative interpretations of the observed correlation.

[1] For evidence that the provision of choice raises student engagement see, for example, a review by Stroet et al. (2013). The explanation is offered by the Self-Determination Theory (Deci and Ryan, 2000): provision of choice helps to satisfy one of the students' basic psychological needs, the need for autonomy, and thereby, raises motivation and engagement.

This paper contributes to the longstanding debate on the effects of class size on educational outcomes (Mishel et al., 2002, Hattie, 2005). It is most closely related to recent econometric studies in higher education: Kokkelenberg et al. (2008), Bandiera et al. (2010), De Giorgi et al. (2012), De Paola et al. (2013) estimate a negative effect of class size on student performance; Bedard and Kuhn (2008), Westerlund (2008), Cheng (2011), Mandel and Sussmuth (2011), Monks and Schmidt (2011), Sapelli and Illanes (2016) do the same for student course evaluations. My main added value to this literature is cleaner identification of the class size effect: by exploiting the richness of the data and comparing parallel sections of the *same* course taught by the *same* instructor in the *same* semester I control for sources of bias not accounted for in prior literature.

My second contribution is to the literature on "why class size matters". I show that the heterogeneity of class size effects across college courses is consistent with class size affecting educational outcomes through student engagement. The engagement hypothesis is proposed by Finn et al. (2003); supporting evidence is presented in, for example, Babcock and Betts (2009), Blatchford et al. (2011). This literature, however, focuses on primary and secondary education. I believe mine is the first paper to show that the same mechanisms that explain class size effects at school operate at the university.

My findings are also consistent with Lazear (2001). Lazear builds a theoretical model of educational production with a focus on student disruptive behavior. The model illustrates that schools enrolling students with a higher propensity to disrupt realize a larger benefit from small classes through a greater reduction in the time lost to disruptive behavior. Viewing disruptive behavior as an extreme form of disengagement, the model shows that optimal class size varies with student engagement. If engagement systematically varies across courses (i.e. mandatory vs electives), so does optimal class size. The heterogeneity of class size effects I find across courses is indeed consistent with this proposition.

The paper is organized as follows. Section 2 describes the institutional background and data. Section 3 presents the results for student grades, while section 4 for course evaluations. Section 5 considers a number of alternative interpretations. Section 6 concludes.

2 Institutional Background and Data

I use administrative data from a small liberal arts and sciences college in the Netherlands. The college offers a broad English-taught bachelor program with Majors in Science, Social Science and Arts and Humanities. The total number of students is a little under 600; more than a third come from outside of the Netherlands. The data spans 28 semesters from fall 2004 to spring 2018.

Several features of the college play an important role in the empirical analysis that follows. First, attendance in all courses is compulsory. This rule assures that the number of students enrolled in a section of a course, which I use to measure class size, closely approximates the number of students effectively present in the same class.

Second, students enjoy much freedom in picking courses they would like to take. Several factors affect their choice: preference to work with a certain instructor, interest in the subject matter, need to take a course to qualify for a specific master, desire to spread known-to-be-difficult courses evenly over their study program, desire to spread classes evenly over the week, preference for classes scheduled in certain parts of the day. In addition, students face graduation requirements: they *must* take certain courses at various points in their program. Finally, the chosen courses must not clash, that is, they must be offered in different time slots.

This constrained freedom of choice generates substantial variation in class size. This variation is shown in Figure 1, which depicts class size distribution across all course sections offered in the sample period. While most sections have enrollments between 20 and 26 students, about a third run at 10 to 19, and some have fewer than 10 students.

Figure 1 also shows the third important institutional feature - the upper limit on class size. The college imposes this limit in order to facilitate student-centred teaching. This upper limit applies to nearly all courses.[2] It was changed only once in my sample period, from 25 to 26 students in 2017 (with two exceptions allowed to run at 27). To abide by this limit, the college offers courses with large enrollments in multiple parallel sections. These sections differ little: they use the same course outline, the same assignments and assessments, and often are taught by the same instructor. They do however often differ in class size.

These differences are shown in Figure 2. For each combination of semester t course c and instructor i with multiple parallel sections (each denoted by g) I compute average class size, $\overline{N_{tci}}$, and then subtract it from the original class size, N_{tcig}. Figure 2 plots the distribution of this demeaned class size, $N_{tcig} - \overline{N_{tci}}$. We can see that the class size differences between parallel sections of the same course are often small to moderate, ranging between 0 and 4 students. An example would be two sections running at 22 and 26 students, which translates in -2 and $+2$ students in Figure 2. But sometimes the differences can be as high as 6 students or more.

Two main forces drive this within-semester-course-instructor variation in class size. First, the time table configuration may, accidentally, favor one section over another. For example, one section may be scheduled in the same time slot as some popular courses, and for that reason be avoided by a large number of students. Second, students may prefer to take classes in certain times of the day (or week) and for that reason prefer one section over the other.

[2] Courses deviating from this general rule are excluded from the analysis. These are performance courses, in which teaching happens one on one, as well as labs and capstone courses, in which the upper limit is set at 15.

The modest, within-semester-course-instructor variation in class size shown in Figure 2 may make it hard to identify the effect of interest. However, the range of class sizes I study (see Figure 1) likely helps: it is precisely this range that has been found to matter in pre-university literature. For example, the famous Project STAR compared small classes of 13-17 students with regular classes of 22-26 students, and found substantial effects (Finn and Achilles, 1999, Krueger and Whitmore, 2001). Similarly, Blatchford et al. (2003) mention the number 25 as a threshold below which class size effects become particularly visible.

The last important institutional feature is the presence of two types of courses: mandatory courses and electives. Students can choose the former but not the latter (see Table 1 for an overview). By definition, *all* students have to take mandatory courses, including those who would prefer not to. I assume that these latter students drive average engagement in mandatory courses below that in electives. If this assumption is true, and holding all else constant, I can attribute differences in class size effect between these two groups of courses to varying student engagement.

I study the effect of class size on final grades and student course evaluations. Final grades in courses are assigned as letters, from F to $A+$. The Student Office then converts these letters to grade points following Table 2. In the regressions I use grade points, but make one adjustment: to distinguish $A+$ from A I change all grade points that correspond to $A+$ from 4.0 to 4.3.

Student course evaluations, available from fall 2009 onward, are collected as follows. At the end of every semester, in class, all (attending) students anonymously fill in a questionnaire with several multiple choice and open questions. A student volunteer then brings all questionnaires to the Student Office. The forms are scanned and the responses are digitized.

I use student responses to the following four questions (response options are in parentheses):

1. In your opinion the overall quality of this course is: (Very bad, Bad, Neutral, Good, Very good)
2. I learned a great deal. (Strongly disagree, Disagree, Neither agree nor disagree, Agree, Strongly agree)
3. Active student participation was encouraged. (Same as above)
4. How actively engaged were you in this course? (Very little, Little, Average, Much, Very much)

For the regressions I convert these responses into numerical scores from 1 to 5.[3]

I focus on these four questions for the following reasons. The first question on overall quality

[3]The questionnaire changed in fall 2012. Tables 13-14 in Appendix A detail both the old and the new version. The four questions used in this paper are available, with a slight change of language, in both the old and the new form (see items 21, 2, 14, 19 in Table 13 and items 11, 1, 2, 8 in Table 14).

reflects student's general satisfaction with the course, and can be viewed as an educational outcome interesting in its own right; it has been studied by, for example, Bedard and Kuhn (2008) and Cheng (2011). The second question on perceived amount learned can be viewed as an alternative measure of study achievement, complementary to grades. The last two questions most directly relate to student engagement - the focus of this paper.

In addition to class size, grades and course evaluations, I use information on student personal characteristics. Specifically, for every course section I compute the percent of female students, $\%Female$, the percent of Dutch students, $\%Dutch$, average student age and its standard deviation.

Table 3 compares summary statistics for electives (columns 1-3) and mandatory courses (columns 4-6). On average, mandatory courses report lower final grades and lower scores on course evaluations. Some of those differences are substantial: for example, average scores on Q1 "I learned a great deal" are 3.77 vs 4.15. Other differences are trivial: for example, average scores on Q2 "Active student participation was encouraged" are 4.03 vs 4.06. The average class size is a little higher in mandatory courses: 21.5 vs 20.9. Controls for gender and age distribution are, on average, very similar. Finally, the somewhat lower % of Dutch students in mandatory courses has likely to do with non-Dutch students being obliged to take Dutch language as an additional required course.

3 Class Size and Student Grades

To compare the effect of class size between mandatory courses and electives I estimate the following equation:

$$Y_{tcigs} = \beta_0^E D^E + \beta_1^E D^E N_{tcig} + \beta_0^M D^M + \beta_1^M D^M N_{tcig} + FE + e_{tcigs} \quad (1)$$

Y_{tcigs} is the final grade of student s enrolled in group g of course c in semester t taught by instructor i. N_{tcig} is the number of students enrolled in group g of course c in semester t taught by instructor i. Dummy D^E equals 1 for electives. Dummy D^M equals 1 for mandatory courses. β_1^E measures the ceteris paribus effect of a one-student increase in class size on the final grade in electives. β_1^M does the same for mandatory courses. FE stands for various combinations of time, course, instructor and student fixed effects. All estimations are performed using Stata routine $reghdfe$ (Correia, 2016). Standard errors are three-way clustered at the level of instructor, course and student.

Table 4 reports the results. In column 1, equation 1 includes time dummies λ_t as the only control. The estimates of both β_1^E and β_1^M are negative; the latter is significant at 5%. The magnitude of β_1^M implies that a one-student increase in class size reduces the average grade in a mandatory course by about 0.02. To compute the corresponding effect size note that the standard deviation of class size in mandatory courses is 4 students, while the standard deviation of grade point is 0.95 (see Table 3). Therefore, a one standard deviation increase in class size reduces the average grade in a mandatory course by $0.021 * 4/0.95 \approx 0.09$ standard deviations. This effect size is very close to the ones reported by Bandiera et al. (2010) and De Giorgi et al. (2012), and is somewhat lower than the average effect size reported in Hattie (2005).

The coefficients reported in column 1 may suffer from omitted variable bias: the same factors that drive student course choice (and by implication class size) may correlate with student performance. One such potential factor is instructor characteristics. For example, some instructors may be particularly effective teachers, both enjoying popularity among students and allowing most of them to excel. This mechanism introduces a positive correlation between grades and class size, and thus biases β_1's upward. Alternatively, popular instructors may apply higher standards and stricter grading, without running the risk of scaring too many students away. In this case, big classes go together with strict grading, and thus introduce a negative bias. In line with prior studies (Bandiera et al., 2010), I control for this source of bias by adding to the regression instructor fixed effects, μ_i. The results are in Table 4 column 2. The estimates of β_1^E and β_1^M are again negative. Unlike in column 1, both are highly statistically significant ($p < 0.01$).

Another potential contaminating factor is student characteristics. For example, top performers may seek smaller classes where they can enjoy more quality time one-on-one with the instructor. This mechanism biases β_1's downward. Following Bandiera et al. (2010) and De Giorgi et al. (2012), I control for this source of bias by adding student fixed effects, α_s. The results are in Table 4 column 3. The estimates of β_1^E and β_1^M are still negative and significant at 5%. Both, however, shrink in magnitude compared to column 2. This shrinkage suggests that the downward bias discussed above, is indeed alleviated by the addition of student fixed effects.

Columns 4 and 5 progressively recognize that factors inducing a spurious correlation between grades and class size *across* instructors, may also operate *within* instructors, both across courses and over time. For example, an instructor may be particularly skilled at teaching a subset of her courses. That subset would then likely be both more popular among students and result in their better performance. Alternatively, if teaching skills improve with experience, the same course-instructor may enjoy growing popularity and improving student performance over time. To address these concerns, I add course-instructor and semester-course-instructor fixed effects in, respectively, columns 4 and 5. In both cases I find a negative and highly statistically significant β_1 ($p < 0.01$) for mandatory courses, but not for electives. In column 5 the magnitude of β_1^M is very close to that reported in column 1: a one-student increase in class size reduces the average grade by 0.016; effect size ≈ 0.07.

The regression reported in column 5 represents an important value added of this paper. Inclusion of semester-course-instructor fixed effects amounts to only exploiting variation in class size across parallel sections of the *same* course taught by the *same* instructor in the *same* semester. It allows me to control for sources of bias operating across semesters within the same course-instructor, and thus to achieve a cleaner identification of the class size effect compared to existing studies. The latter (Kokkelenberg et al., 2008, Bandiera et al., 2010, De Giorgi et al., 2012) employ course, instructor, and time fixed effects, but not the interaction of the three, because their data does not have parallel sections of the same course taught by the same instructor.

There still remains one concern with the results presented in column 5. Parallel course sections are scheduled in different time slots, some potentially more favored by students than others. For example, students tend to avoid classes in the early morning, as well as on Friday afternoon (see Table 12 in Appendix A for the list of time slots). If these time slot preferences correlate with student performance, the estimated β_1's in column 5 would be biased. To alleviate this concern, I include time slot fixed effects, μ_h, in column 6. The results change little. β_1^M remains negative and highly statistically significant.

Finally, column 7 controls for differences in student body characteristics across course sections: the percent of female students, $\%Female$, the percent of Dutch students, $\%Dutch$, average student age and its standard deviation. These are standard controls employed in prior literature. Again, adding these controls affects the results little.

Let me explore the source of variation in class size that drives the results in columns 6 and 7. This variation is clearly unrelated to semester, course, instructor or time slot characteristics as those are controlled for using fixed effects. There are two remaining sources of variation.

First, students must enroll in a course section that does not clash with other chosen courses; in addition, they often try to spread classes evenly over the week. In other words, student's choice depends on the college-wide time table configuration. Because the latter represents an outcome of dozens of random factors (e.g. availability and time preferences of instructors, year-on-year variation in the number of students, strategic changes in course offerings imposed by management, etc.), I argue that this source of variation in class size is as good as exogenous.

Second, some students literally don't care and pick the course section at random. Clearly, this too represents an exogenous source of variation in class size, appropriate to study its causal effect on student performance.

Having made the case for clean identification of the class size effect, let me now turn to the hypothesis of interest: is the effect stronger in mandatory courses compared to electives? First, in line with expectations, in all specifications of Table 4 β_1^M is negative and smaller than β_1^E. Second, the last row reports the p-value of an F-test of the null hypothesis: $\beta_1^E = \beta_1^M$. In the last three specifications, which most cleanly identify the class size effect, the equality is rejected at 1%. These results support the engagement channel: small class size helps the instructors to raise student engagement (and consequently performance); but the effects are only visible in mandatory courses, in which engagement is (relatively) low to start with.

The last point I would like to discuss is the positive and significant estimate of β_1^E in columns 5-7. The significance is not very strong, yet the positive sign is unexpected: it suggests bigger classes raise performance in electives. Such positive estimates of the class size coefficient have been reported in prior literature (Dobbelsteen et al., 2002, Denny and Oppedisano, 2013). The proposed explanation builds on social cognitive learning theory: students learn from each other, in particular, from classmates with similar levels of competence; bigger class size raises the (expected) number of classmates with a similar level of competence and thus benefits learning of a particular student.

Taken together, the opposite signs of β_1^M and β_1^E clearly suggest optimal class size differs between mandatory courses and electives. In some sense, this conclusion is not surprising. As shown by Lazear (2001), optimal class size varies with student engagement.[4] If engagement systematically varies across courses, so should optimal class size.

4 Class Size and Course Evaluations

To study the effect of class size on student course evaluations, I re-estimate equation 1 replacing grades on the left-hand side with each of the numerical scores reported in Table 3. Those scores represent student answers to questions on overall course quality, amount learned, participation and engagement. Tables 5-8 present the results. They follow the same structure as Table 4 with one exception: because course evaluations are anonymous, regressions do not include student fixed effects nor clustering at the student level.

Table 5 reports the results for perceived overall course quality. The pattern of findings is very similar to that reported for final grades in Table 4: β_1^M is consistently negative, highly statistically significant, and smaller than β_1^E; the latter difference is significant at 5% in four out of six specifications (see last row), including columns 4-6 that include semester-course-instructor fixed effects. The magnitude of β_1^M in column 6 implies that a one-student increase in class size reduces the overall quality score in a mandatory course by 0.023; effect size ≈ 0.11. These results are very close to those reported in prior literature (Cheng, 2011).

Table 6 reports the results for perceived amount learned. The pattern of findings is similar to that reported for overall course quality[5]: bigger classes reduce perceived amount learned; with semester-course-instructor fixed effects included in the regression, the effect is present in mandatory courses, but not in electives; the difference is significant at 5%.

Tables 7-8 report the results for questions on, respectively, student participation and engagement. In both tables we again see β_1^M being negative, significant, and typically smaller than β_1^E; with

[4] Lazear (2001) focuses on disruptive behavior, which can be viewed as an extreme form of disengagement (Finn et al., 2003).

[5] This outcome is not surprising as the two scores are strongly correlated ($r \approx 0.66$).

semester-course-instructor fixed effects included in the regression, the latter difference is often (but not always) significant. In other words, bigger classes result in students feeling less encouraged to participate and less engaged in mandatory courses, but not in electives. These findings provide further support for student engagement being the channel through which class size affects educational outcomes.

5 Further Considerations

Any findings based on observational data are subject to alternative interpretations. I now turn to those.

5.1 Endogeneity of Class Size

The first concern I would like to address is potential endogeneity of class size. While I have made the case for exploiting its exogenous variation across multiple sections of the same course, ultimately students *choose* which section of the course to enrol in. Because of this presence of student choice, there remains a concern that some unobserved variable drives both class size and student performance. To address this concern, I need to disable student choice.

Fortunately, a sub-sample of my data is "free" from student choice. Specifically, first semester students do not get a chance to choose a specific section within a course. For them the procedure runs as follows. When they apply to college, they indicate their preferred courses on the application form. Before semester starts, the Student Office enrolls them in courses trying to accommodate their preferences. If a course has multiple sections, the Student Office picks one without consulting the student. That is, student choice plays no role in her assignment to a specific section within a course.

I exploit this assignment of first semesters in Table 9. It reproduces the results from Table 4 but for a sub-sample of first semester students only. Of particular interest are regressions with semester-course-instructor fixed effects reported in columns 5-7. The results are similar to those in Table 4: β_1^M is consistently negative and significant; β_1^E is positive and sometimes marginally significant; equality $\beta_1^E = \beta_1^M$ is rejected at least at 5%. Arguably, these regressions provide the cleanest identification of the class size effect. Importantly, the results remain consistent with the engagement channel. Unfortunately, I can't replicate this exercise for course evaluations because the latter are anonymous.

5.2 Student versus Teacher Behavior

My tests of engagement channel focus on the difference $\beta_1^E - \beta_1^M$. Yet, it is interesting to consider what drives the levels of β_1^E and β_1^M. Similarly to affecting the behavior of students, class size may affect the behavior of teachers. For example, Blatchford et al. (2002) and Blatchford et al. (2005) report extensive evidence on changes in teacher behavior as a potential explanation for the benefits of smaller classes.

Two items from the course evaluation most directly address teacher behavior: item 6 "I was provided with feedback on my individual work" and item 7 "The instructor was available for individual guidance" (see Table 14). Using the numerical score on each of those two items as a dependent variable, I re-estimate the most demanding specifications with semester-course-instructor fixed effects of Table 5. For both items I find very small and highly insignificant β_1's (see Table 10). That is, across parallel sections of the same course, class size has little effect on teacher behavior. This result holds for both mandatory courses and electives. It suggests that, for parallel sections of the same course, student behavior may not be just one mechanism through which class size affects educational outcomes; it may be the most important one. Clearly, this conclusion is very tentative: teacher behavior includes much more than feedback and availability; those aspects have not been considered here.

5.3 Placebo Regressions

Finally, I run some placebo regressions. These are regressions in which class size is expected to *not affect* the dependent variable. Finding that it does, would cast doubt on student engagement being the story behind my earlier results.

Assuming clarity of (oral or written) explanation does not depend on the size of the audience, two items from the course evaluations should be unrelated to class size: item 4 "The requirements were clear" and item 5 "The grading criteria were clear" (see Table 14). Using the numerical score on each of those two items as a dependent variable, I re-estimate the most demanding specifications with semester-course-instructor fixed effects of Table 5. For both items I find mostly insignificant β_1's (see Table 11). These results are in line with expectations.

6 Conclusion

In this paper, I use administrative data from a small Dutch liberal arts college to study the effect of class size on student grades and course evaluations. My identification strategy exploits variation in class size across parallel sections of the *same* course taught by the *same* instructor in the *same* semester. I show that class size has a stronger negative effect on student grades in mandatory courses compared to electives. I show similar results for various components of student course evaluations: perceived overall course quality, perceived amount learned, student participation and engagement.

I interpret these findings to support the engagement hypothesis (Finn et al., 2003). Because class size works as a substitute for engagement, its effects should be particularly pronounced in courses with low engagement. That's exactly what I find: mandatory courses, which students often take because they have to, enjoy a stronger class size effect compared to electives. The takeaway for policy is clear: holding all else constant, class size reductions should prioritize mandatory courses (or other courses with low student engagement). And vice versa, class size increases would likely do less harm in electives.

References

Babcock, P. and Betts, J. R. (2009). Reduced-class distinctions: Effort, ability, and the education production function. *Journal of Urban Economics*, 65(3):314–322.

Bandiera, O., Larcinese, V., and Rasul, I. (2010). Heterogeneous Class Size Effects: New Evidence from a Panel of University Students. *The Economic Journal*, 120(549):1365–1398.

Bedard, K. and Kuhn, P. (2008). Where class size really matters: Class size and student ratings of instructor effectiveness. *Economics of Education Review*, 27(3):253–265.

Blatchford, P., Bassett, P., and Brown, P. (2005). Teachers' and Pupils' Behavior in Large and Small Classes: A Systematic Observation Study of Pupils Aged 10 and 11 Years. *Journal of Educational Psychology*, 97(3):454–467.

Blatchford, P., Bassett, P., and Brown, P. (2011). Examining the effect of class size on classroom engagement and teacherpupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction*, 21(6):715–730.

Blatchford, P., Bassett, P., Goldstein, H., and Martin, C. (2003). Are class size differences related to pupils' educational progress and classroom processes? findings from the institute of education class size study of children aged 57 years. *British Educational Research Journal*, 29(5):709–730.

Blatchford, P., Moriarty, V., Edmonds, S., and Martin, C. (2002). Relationships Between Class Size and Teaching: A Multimethod Analysis of English Infant Schools. *American Educational Research Journal*, 39(1):101–132.

Cheng, D. A. (2011). Effects of class size on alternative educational outcomes across disciplines. *Economics of Education Review*, 30(5):980–990.

Correia, S. (2016). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects. *Statistical Software Components*.

De Giorgi, G., Pellizzari, M., and Woolston, W. G. (2012). CLASS SIZE AND CLASS HETEROGENEITY. *Journal of the European Economic Association*, 10(4):795–830.

De Paola, M., Ponzo, M., and Scoppa, V. (2013). Class size effects on student achievement: heterogeneity across abilities and fields. *Education Economics*, 21(2):135–153.

Deci, E. L. and Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11:227–268.

Denny, K. and Oppedisano, V. (2013). The surprising effect of larger class sizes: Evidence using two identification strategies. *Labour Economics*, 23:57–65.

Dobbelsteen, S., Levin, J., and Oosterbeek, H. (2002). The causal effect of class size on scholastic achievement: distinguishing the pure class size effect from the effect of changes in class composition. *Oxford Bulletin of Economics and Statistics*, 64(1):17–38.

Finn, J. D. and Achilles, C. M. (1999). Tennessee's Class Size Study: Findings, Implications, Misconceptions. *Educational Evaluation and Policy Analysis*, 21(2):97–109.

Finn, J. D., Pannozzo, G. M., and Achilles, C. M. (2003). The Why's of Class Size: Student Behavior in Small Classes. *Review of Educational Research*, 73(3):321–368.

Hattie, J. (2005). The paradox of reducing class size and improving learning outcomes. *International Journal of Educational Research*, 43(6):387–425.

Kokkelenberg, E. C., Dillon, M., and Christy, S. M. (2008). The effects of class size on student grades at a public university. *Economics of Education Review*, 27(2):221–233.

Krueger, A. B. and Whitmore, D. M. (2001). The Effect of Attending a Small Class in the Early Grades on College-test Taking and Middle School Test Results: Evidence from Project Star. *The Economic Journal*, 111(468):1–28.

Lazear, E. P. (2001). Educational Production. *The Quarterly Journal of Economics*, 116:777–803.

Mandel, P. and Sussmuth, B. (2011). Size matters. The relevance and Hicksian surplus of preferred college class size. *Economics of Education Review*, 30(5):1073–1084.

Mishel, L. R., Rothstein, R., Krueger, A. B., Hanushek, E. A. E. A., and Rice, J. K. (2002). *The class size debate*. Economic Policy Institute.

Monks, J. and Schmidt, R. M. (2011). The Impact of Class Size on Outcomes in Higher Education. *The B.E. Journal of Economic Analysis & Policy*, 11(1).

Sapelli, C. and Illanes, G. (2016). Class size and teacher effects in higher education. *Economics of Education Review*, 52:19–28.

Stroet, K., Opdenakker, M.-C., and Minnaert, A. (2013). Effects of need supportive teaching on early adolescents motivation and engagement: A review of the literature. *Educational Research Review*, 9:65–87.

Westerlund, J. (2008). Class size and student evaluations in Sweden. *Education Economics*, 16(1):19–28.

Tables and Figures

Table 1. Overview of Mandatory Courses

Course	Mandatory for
Methods and Statistics I	All students
Methods and Statistics II	Social Science majors
Qualitative Methods	Social Science majors
English for Academic Purposes (abolished in 2015)	Students with insufficient English
Academic Writing and Presenting	All students
Writing Across the Disciplines	All students
Introduction to Rhetoric and Argumentation	Arts and Humanities majors
Mathematical Ideas and Methods in Context	Science majors
Foreign language: Dutch, French, German or Spanish	All students

Table 2. Equivalence between Letter Grades and Grade Points

Letter	F	D-	D	D+	C-	C	C+	B-	B	B+	A-	A	A+
Grade Point	0	.7	1	1.3	1.7	2	2.3	2.7	3	3.3	3.7	4	4

Table 3. Summary Statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	Electives			Mandatory		
VARIABLES	N	mean	sd	N	mean	sd
Grade Point	33,342	3.28	0.90	14,721	3.12	0.95
Score on Q1	22,236	4.15	0.81	9,391	3.77	0.90
Score on Q2	22,195	4.06	0.85	9,383	4.03	0.82
Score on Q8	22,217	3.58	0.87	9,395	3.46	0.88
Score on Q11	22,188	4.08	0.82	9,370	3.77	0.85
Class Size	33,782	20.9	4.71	14,943	21.5	4.01
% Female	33,782	65.2	16.3	14,943	65.2	11.6
% Dutch	33,782	69.6	13.2	14,943	67.3	23.4
Mean of Age	33,782	20.5	0.79	14,943	20.0	0.89
SD of Age	33,780	1.39	0.48	14,930	1.36	0.55

Table 4. Results for Student Grade

$$Y_{tcigs} = \beta_0^E D^E + \beta_1^E D^E N_{tcig} + \beta_0^M D^M + \beta_1^M D^M N_{tcig} + FE + e_{tcigs}$$

$tcigs$ identifies student s enrolled in section g of course c in semester t taught by instructor i. The dependent variable, Y_{tcigs}, is reported in table heading. N_{tcig} is the number of students enrolled in section g of course c in semester t taught by instructor i. Dummy D^E equals 1 for electives. Dummy D^M equals 1 for mandatory courses. FE stands for various combinations of time, course, instructor and student fixed effects reported in column headings. The last column regression includes the following (unreported) controls: $\%Female$, $\%Dutch$, Mean and SD of Age. Clustered standard errors are in parentheses. The last row reports the p-value of an F-test of the following null hypothesis: $\beta_1^E = \beta_1^M$. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	λ_t	$\lambda_t + \mu_i$	$\lambda_t + \mu_i + \alpha_s$	$\lambda_t + \mu_{ci} + \alpha_s$	$\lambda_{tci} + \alpha_s$	$\lambda_{tci} + \mu_h + \alpha_s$	$\lambda_{tci} + \mu_h + \alpha_s$
β_1^E	-0.004	-0.012***	-0.004**	-0.004**	0.017*	0.020**	0.021**
	(0.003)	(0.002)	(0.002)	(0.002)	(0.009)	(0.010)	(0.010)
β_1^M	-0.021**	-0.013***	-0.010**	-0.007**	-0.016***	-0.018***	-0.017***
	(0.008)	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)	(0.005)
Observations	48,063	48,063	48,045	48,044	48,042	48,042	48,042
R^2	0.016	0.090	0.564	0.586	0.625	0.625	0.625
Pval: $\beta_1^E = \beta_1^M$	0.043	0.866	0.174	0.369	0.002	0.002	0.002

Table 5. Results for Score on Q11: In your opinion the overall quality of this course is

$$Y_{tcigs} = \beta_0^E D^E + \beta_1^E D^E N_{tcig} + \beta_0^M D^M + \beta_1^M D^M N_{tcig} + FE + e_{tcigs}$$

$tcigs$ identifies student s enrolled in section g of course c in semester t taught by instructor i. The dependent variable, Y_{tcigs}, is reported in table heading. N_{tcig} is the number of students enrolled in section g of course c in semester t taught by instructor i. Dummy D^E equals 1 for electives. Dummy D^M equals 1 for mandatory courses. FE stands for various combinations of time, course, instructor and student fixed effects reported in column headings. The last column regression includes the following (unreported) controls: $\%Female$, $\%Dutch$, Mean and SD of Age. Clustered standard errors are in parentheses. The last row reports the p-value of an F-test of the following null hypothesis: $\beta_1^E = \beta_1^M$. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	λ_t	$\lambda_t + \mu_i$	$\lambda_t + \mu_{ci}$	λ_{tci}	$\lambda_{tci} + \mu_h$	$\lambda_{tci} + \mu_h$
β_1^E	0.004	-0.009***	-0.010***	0.012	0.012	0.008
	(0.005)	(0.003)	(0.003)	(0.012)	(0.013)	(0.011)
β_1^M	-0.014**	-0.015***	-0.016***	-0.020***	-0.023***	-0.023***
	(0.006)	(0.004)	(0.005)	(0.006)	(0.007)	(0.008)
Observations	31,558	31,558	31,558	31,558	31,558	31,545
R^2	0.032	0.211	0.254	0.337	0.337	0.338
Pval: $\beta_1^E = \beta_1^M$	0.021	0.195	0.264	0.018	0.009	0.015

Table 6. Results for Score on Q1: I learned a great deal

$$Y_{tcigs} = \beta_0^E D^E + \beta_1^E D^E N_{tcig} + \beta_0^M D^M + \beta_1^M D^M N_{tcig} + FE + e_{tcigs}$$

$tcigs$ identifies student s enrolled in section g of course c in semester t taught by instructor i. The dependent variable, Y_{tcigs}, is reported in table heading. N_{tcig} is the number of students enrolled in section g of course c in semester t taught by instructor i. Dummy D^E equals 1 for electives. Dummy D^M equals 1 for mandatory courses. FE stands for various combinations of time, course, instructor and student fixed effects reported in column headings. The last column regression includes the following (unreported) controls: $\%Female$, $\%Dutch$, Mean and SD of Age. Clustered standard errors are in parentheses. The last row reports the p-value of an F-test of the following null hypothesis: $\beta_1^E = \beta_1^M$. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	λ_t	$\lambda_t + \mu_i$	$\lambda_t + \mu_{ci}$	λ_{tci}	$\lambda_{tci} + \mu_h$	$\lambda_{tci} + \mu_h$
β_1^E	-0.001	-0.009***	-0.008***	0.014*	0.016	0.014
	(0.004)	(0.003)	(0.003)	(0.008)	(0.010)	(0.010)
β_1^M	-0.012**	-0.015***	-0.019***	-0.020*	-0.021*	-0.022*
	(0.005)	(0.005)	(0.004)	(0.012)	(0.012)	(0.012)
Observations	31,627	31,627	31,627	31,627	31,627	31,614
R^2	0.045	0.161	0.208	0.274	0.274	0.274
Pval: $\beta_1^E = \beta_1^M$	0.106	0.336	0.044	0.015	0.026	0.027

Table 7. Results for Score on Q2: Active student participation was encouraged

$$Y_{tcigs} = \beta_0^E D^E + \beta_1^E D^E N_{tcig} + \beta_0^M D^M + \beta_1^M D^M N_{tcig} + FE + e_{tcigs}$$

$tcigs$ identifies student s enrolled in section g of course c in semester t taught by instructor i. The dependent variable, Y_{tcigs}, is reported in table heading. N_{tcig} is the number of students enrolled in section g of course c in semester t taught by instructor i. Dummy D^E equals 1 for electives. Dummy D^M equals 1 for mandatory courses. FE stands for various combinations of time, course, instructor and student fixed effects reported in column headings. The last column regression includes the following (unreported) controls: $\%Female$, $\%Dutch$, Mean and SD of Age. Clustered standard errors are in parentheses. The last row reports the p-value of an F-test of the following null hypothesis: $\beta_1^E = \beta_1^M$. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	λ_t	$\lambda_t + \mu_i$	$\lambda_t + \mu_{ci}$	λ_{tci}	$\lambda_{tci} + \mu_h$	$\lambda_{tci} + \mu_h$
β_1^E	-0.007	-0.017***	-0.016***	0.028***	0.029**	0.025**
	(0.005)	(0.002)	(0.002)	(0.010)	(0.011)	(0.012)
β_1^M	-0.023***	-0.017***	-0.018***	-0.022***	-0.023***	-0.022***
	(0.008)	(0.003)	(0.004)	(0.005)	(0.005)	(0.005)
Observations	31,578	31,578	31,578	31,578	31,578	31,566
R^2	0.005	0.183	0.213	0.275	0.275	0.276
Pval: $\beta_1^E = \beta_1^M$	0.110	0.889	0.618	0.000	0.000	0.000

Table 8. Results for Score on Q8: How actively engaged were you in this course

$$Y_{tcigs} = \beta_0^E D^E + \beta_1^E D^E N_{tcig} + \beta_0^M D^M + \beta_1^M D^M N_{tcig} + FE + e_{tcigs}$$

$tcigs$ identifies student s enrolled in section g of course c in semester t taught by instructor i. The dependent variable, Y_{tcigs}, is reported in table heading. N_{tcig} is the number of students enrolled in section g of course c in semester t taught by instructor i. Dummy D^E equals 1 for electives. Dummy D^M equals 1 for mandatory courses. FE stands for various combinations of time, course, instructor and student fixed effects reported in column headings. The last column regression includes the following (unreported) controls: $\%Female$, $\%Dutch$, Mean and SD of Age. Clustered standard errors are in parentheses. The last row reports the p-value of an F-test of the following null hypothesis: $\beta_1^E = \beta_1^M$. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	λ_t	$\lambda_t + \mu_i$	$\lambda_t + \mu_{ci}$	λ_{tci}	$\lambda_{tci} + \mu_h$	$\lambda_{tci} + \mu_h$
β_1^E	-0.014***	-0.022***	-0.020***	0.003	-0.008	-0.010
	(0.003)	(0.002)	(0.003)	(0.010)	(0.012)	(0.012)
β_1^M	-0.020***	-0.017***	-0.014***	-0.021***	-0.024***	-0.025***
	(0.004)	(0.003)	(0.004)	(0.006)	(0.005)	(0.005)
Observations	31,612	31,612	31,612	31,612	31,612	31,599
R^2	0.011	0.062	0.082	0.125	0.125	0.125
Pval: $\beta_1^E = \beta_1^M$	0.293	0.263	0.216	0.049	0.172	0.172

Table 9. Results for Student Grade: First semester students only

$$Y_{tcigs} = \beta_0^E D^E + \beta_1^E D^E N_{tcig} + \beta_0^M D^M + \beta_1^M D^M N_{tcig} + FE + e_{tcigs}$$

$tcigs$ identifies student s enrolled in section g of course c in semester t taught by instructor i. The dependent variable, Y_{tcigs}, is reported in table heading. N_{tcig} is the number of students enrolled in section g of course c in semester t taught by instructor i. Dummy D^E equals 1 for electives. Dummy D^M equals 1 for mandatory courses. FE stands for various combinations of time, course, instructor and student fixed effects reported in column headings. The last column regression includes the following (unreported) controls: $\%Female$, $\%Dutch$, Mean and SD of Age. Clustered standard errors are in parentheses. The last row reports the p-value of an F-test of the following null hypothesis: $\beta_1^E = \beta_1^M$. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	λ_t	$\lambda_t + \mu_i$	$\lambda_t + \mu_i + \alpha_s$	$\lambda_t + \mu_{ci} + \alpha_s$	$\lambda_{tci} + \alpha_s$	$\lambda_{tci} + \mu_h + \alpha_s$	$\lambda_{tci} + \mu_h + \alpha_s$
β_1^E	0.002	-0.003	0.002	0.003	0.028*	0.025*	0.017
	(0.007)	(0.006)	(0.004)	(0.005)	(0.014)	(0.014)	(0.014)
β_1^M	-0.030**	-0.013***	-0.013***	-0.017***	-0.030***	-0.028***	-0.026**
	(0.013)	(0.004)	(0.005)	(0.004)	(0.009)	(0.010)	(0.010)
Observations	10,305	10,281	10,261	10,190	9,959	9,959	9,959
R^2	0.022	0.116	0.711	0.726	0.764	0.764	0.765
Pval: $\beta_1^E = \beta_1^M$	0.019	0.170	0.020	0.001	0.001	0.003	0.026

Table 10. Results for Scores on Q6 and Q7

$$Y_{tcigs} = \beta_0^E D^E + \beta_1^E D^E N_{tcig} + \beta_0^M D^M + \beta_1^M D^M N_{tcig} + FE + e_{tcigs}$$

$tcigs$ identifies student s enrolled in section g of course c in semester t taught by instructor i. The dependent variable, Y_{tcigs}, is the numerical score on Q6 "I was provided with feedback on my individual work" in columns 1-3 and on Q7 "The instructor was available for individual guidance" in columns 4-6. N_{tcig} is the number of students enrolled in section g of course c in semester t taught by instructor i. Dummy D^E equals 1 for electives. Dummy D^M equals 1 for mandatory courses. FE stands for various combinations of time, course, instructor and student fixed effects reported in column headings. Regressions in columns 3 and 6 include the following (unreported) controls: $\%Female$, $\%Dutch$, Mean and SD of Age. Clustered standard errors are in parentheses. The last row reports the p-value of an F-test of the following null hypothesis: $\beta_1^E = \beta_1^M$. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	λ_{tci}	$\lambda_{tci} + \mu_h$	$\lambda_{tci} + \mu_h$	λ_{tci}	$\lambda_{tci} + \mu_h$	$\lambda_{tci} + \mu_h$
β_1^E	0.002	0.005	-0.001	0.018	0.005	-0.001
	(0.015)	(0.019)	(0.019)	(0.017)	(0.023)	(0.026)
β_1^M	0.005	0.006	0.006	0.002	-0.006	-0.005
	(0.014)	(0.016)	(0.017)	(0.014)	(0.013)	(0.014)
Observations	21,231	21,231	21,231	20,552	20,552	20,552
R^2	0.247	0.247	0.247	0.213	0.213	0.213
Pval: $\beta_1^E = \beta_1^M$	0.873	0.953	0.749	0.473	0.659	0.875

Table 11. Results for Scores on Q4 and Q5

$$Y_{tcigs} = \beta_0^E D^E + \beta_1^E D^E N_{tcig} + \beta_0^M D^M + \beta_1^M D^M N_{tcig} + FE + e_{tcigs}$$

$tcigs$ identifies student s enrolled in section g of course c in semester t taught by instructor i. The dependent variable, Y_{tcigs}, is the numerical score on Q4 "The requirements were clear" in columns 1-3 and on Q5 "The grading criteria were clear" in columns 4-6. N_{tcig} is the number of students enrolled in section g of course c in semester t taught by instructor i. Dummy D^E equals 1 for electives. Dummy D^M equals 1 for mandatory courses. FE stands for various combinations of time, course, instructor and student fixed effects reported in column headings. Regressions in columns 3 and 6 include the following (unreported) controls: $\%Female$, $\%Dutch$, Mean and SD of Age. Clustered standard errors are in parentheses. The last row reports the p-value of an F-test of the following null hypothesis: $\beta_1^E = \beta_1^M$. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	λ_{tci}	$\lambda_{tci} + \mu_h$	$\lambda_{tci} + \mu_h$	λ_{tci}	$\lambda_{tci} + \mu_h$	$\lambda_{tci} + \mu_h$
β_1^E	0.024*	0.028*	0.018	0.002	0.001	-0.010
	(0.014)	(0.016)	(0.018)	(0.011)	(0.019)	(0.019)
β_1^M	-0.011	-0.009	-0.008	-0.004	-0.007	-0.007
	(0.018)	(0.018)	(0.020)	(0.010)	(0.010)	(0.011)
Observations	21,372	21,372	21,372	21,229	21,229	21,229
R^2	0.281	0.281	0.282	0.247	0.248	0.248
Pval: $\beta_1^E = \beta_1^M$	0.121	0.071	0.251	0.721	0.682	0.864

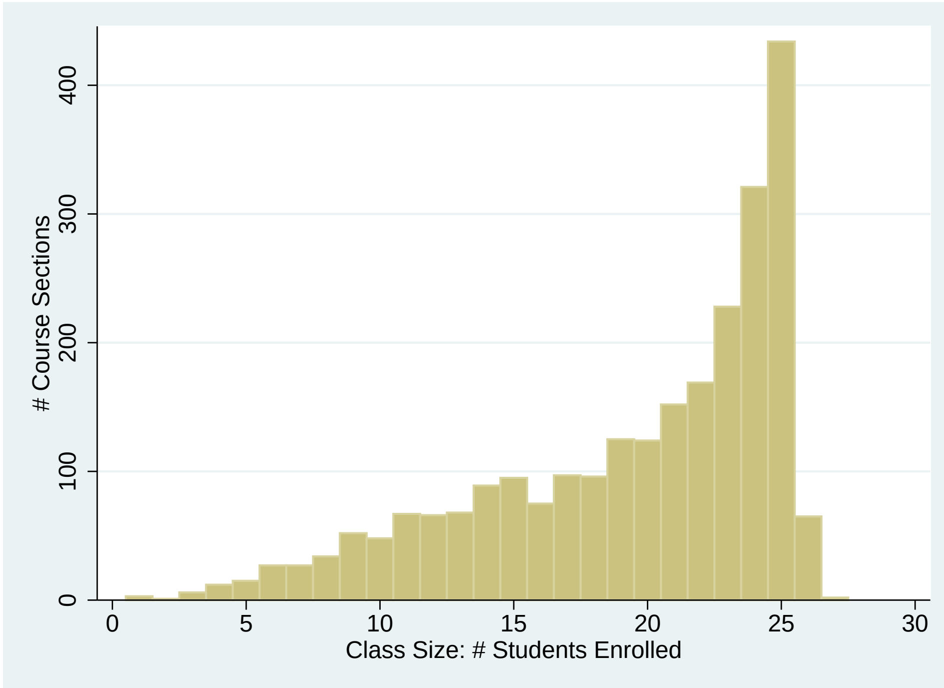

Figure 1. Distribution of Class Size

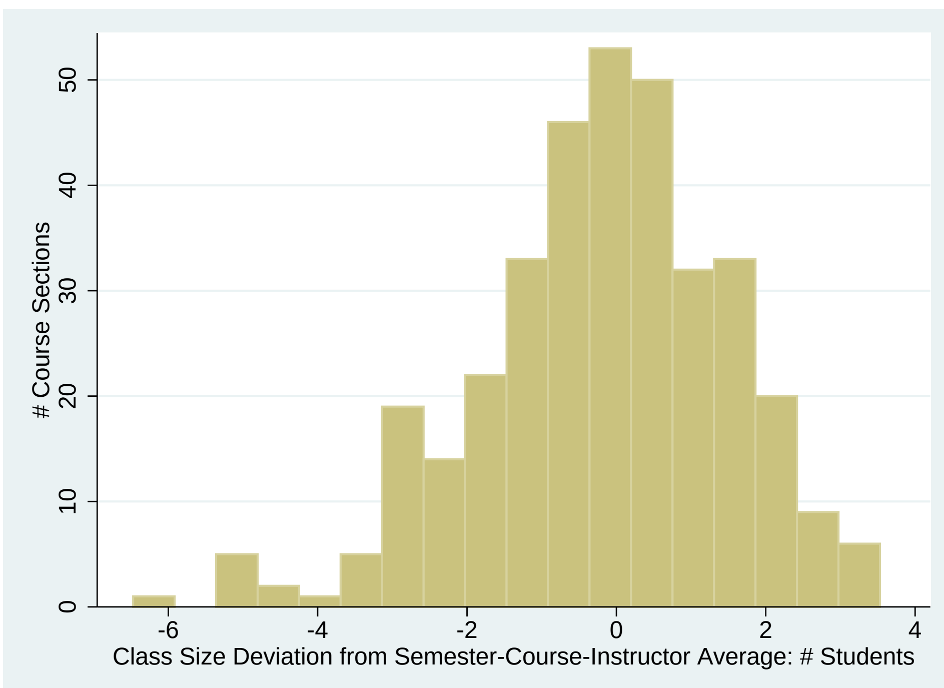

Figure 2. Within Semester-Course-Instructor Variation in Class Size

Appendices

A Appendix

Table 12. Description of Time Slots

Time Slot	Description
T0	Irregular times (typically late afternoons)
T1	Monday 08.45 - 10.45 and Thursday 13.45 - 15.45
T2	Monday 11.00 - 13.00 and Thursday 11.00 - 13.00
T3	Monday 13.45 - 15.45 and Thursday 08.45 - 10.45
T4	Monday 16.00 - 18.00 and Tuesday 11.00 - 13.00
T5	Tuesday 08.45 - 10.45 and Friday 13.45 - 15.45
T6	Tuesday 16.00 - 18.00 and Friday 11.00 - 13.00
T7	Tuesday 13.45 - 15.45 and Friday 16.00 - 18.00
T8	Thursday 16.00 - 18.00 and Friday 08.45 - 10.45
T9	Wednesday all day

Table 13. List of Questions from the Old Course Evaluation Form (2009h2-2012h1)

	Question	Option 1	Option 2	Option 3	Option 4	Option 5
1	My interest in the subject matter has increased as a consequence of this course.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
2	I learned a great deal in this course.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
3	This course improved my understanding of concepts and principles on the subject matter.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
4	I learned a lot from the assignments.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
5	The required reading texts were relevant.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
6	It was clear to me how I would be assessed.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
7	Assessment methods are appropriate.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
8	The instructor keeps me informed on my progress.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
9	The instructor provides personal help when needed.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
10	The feedback on my performance is helpful.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
11	The course was soundly planned.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
12	The course manual was helpful.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
13	The instructor stimulates thinking and my desire to learn.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
14	Active student involvement was encouraged.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
15	The instructor explains clearly.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
16	The instructor was flexible enough to meet individual learning needs.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
17	The instructor has a genuine interest in students.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
18	The instructor has interest in the quality of his / her teaching.	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
19	How actively engaged have you been?	Very little	Little	Some	Much	Very much
20	How many hours did you spend on this course on average per week (including classes)?	< 6 hr	6 - 10 hr	11 - 14 hr	15 - 18 hr	> 18 hr
21	How would you evaluate the overall quality of this course?	Very bad	Bad	Neutral	Good	Very good

Table 14. List of Questions from the New Course Evaluation Form (2012h2-2018h1)

	Question	Option 1	Option 2	Option 3	Option 4	Option 5
1	I learned a great deal.	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
2	Active student participation was encouraged.	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
3	My critical thinking was stimulated.	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
4	The requirements were clear.	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
5	The grading criteria were clear.	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
6	I was provided with feedback on my individual work.	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
7	The instructor was available for individual guidance.	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
8	How actively engaged were you in this course?	Very little	Little	Average	Much	Very much
9	Approx. how many hours per week did you devote to this class (NOT including class time)?	< 5 hr	5 - 8 hr	8 - 12 hr	12 - 15 hr	> 15 hr
10	The final grade you expect in this course is:	A	B	C	D	F
11	In your opinion the overall quality of this course is:	Very good	Good	Neutral	Bad	Very bad