A Personal Reflection on Teaching Effectiveness and Feedback

Steven V. Miller

Institutionen för ekonomisk historia och internationella relationer, Stockholms Universitet

steven.miller@ekohist.su.se

Last updated: October 25, 2022

This document offers an assessment of my teaching effectiveness for my tenure at Clemson University.[1] The assessment is two-fold after a discussion of the course evaluation system at Clemson University. First, I offer an empirical assessment of my teaching evaluations from August 2013 to Fall 2020, pooling teaching evaluations by each unique class I have taught, and comparing them to the department average for all classes.[2] Second, whereas I prepared much of these analyses in the summer of 2018 for a tenure packet, I reflect on these data, highlighting areas where I score well while also discussing changes I have made to my in-class instruction in light of some negative evaluations. The rest of this document outlines how I proceeded with this empirical analysis, starting with a discussion of Clemson University's course evaluation system.

Clemson University's Course Evaluation System

Clemson University has been using online-only course evaluations for longer than I have been employed at the university. This online-only evaluation system has instructors activate the online assessment tool shortly before the end of the semester and then inform the class to fill out these forms. The inferential problem of selection is unavoidable in these evaluations. We, as instructors, try our best to encourage students to fill out these forms, but the evaluations we receive come from those most motivated to complete them. Evaluations will always be a non-random sample of all students, barring a more effective sampling method or outright compulsion (c.f. Nulty, 2008). No matter, the completion rate for these evaluations is around 70% for all my classes.

The evaluations are two-fold and work at two-levels. The "two-levels" indicate the university has a standard catalog of questions for all instructors in all departments and that each department can craft additional questions to include on these assessments to gauge metrics germane to their own interests. The "two-fold" element comes from the varying nature of questions. Most questions are Likert-like responses in which the evaluation gives a prompt, either in the form of a question or a statement, and then the student says on a five-point scale the degree to which the student agrees with it. Higher values indicate more agreement. There are additional open-ended responses though these do not lend themselves to an easy quantification and, for the most part, do not manifest in the more exhaustive feedback we receive about our performance in the classroom. This results in 13 university-level metrics that share this five-point scale and an additional 10 prompts that assess metrics of interest to the Department of Political Science at Clemson University.[3] The prompts at the university-level (with corollary variables codes in parentheses) are:

- "The instructor clearly communicated what I was expected to learn." (G1)

---

[1] Data limitations preclude a consideration of my classes before August 2013. I taught these at the University of Illinois and University of Alabama, where evaluations were written and I no longer have access to these records.

[2] The spring 2020 semester was abruptly moved online because of the COVID-19 pandemic and the university communicated that it did not expect teaching evaluations to be active as a result of this.

[3] This analysis will exclude one university-level metric ("Was the course a requirement for you?") and one department-level metric ("I had a strong desire to take this course.") because neither prompt measures teaching effectiveness in any meaningful way even after considering for the measurement bias inherent in these student assessments of teaching (e.g. McKeachie, 1997).

- "The instructor made the relevance of the course material clear." (`G2`)
- "The course was well organized." (`G3`)
- "There was a positive interaction between the class and the instructor." (`G4`)
- "The instructor's teaching methods helped me understand the course material." (`G5`)
- "The instructor's verbal communication skills helped me understand the course material." (`G6`)
- "The instructor clearly explained what was expected on assignments and tests." (`G7`)
- "The instructor kept me informed about my progress in the course." (`G8`)
- "The feedback I received on assignments and tests gave me the opportunity to improve my performance." (`G9`)
- "Overall, the instructor is an effective teacher." (`G10`)[4]
- "The instructor's grading procedures gave a fair evaluation of my understanding of the material." (`G11`)
- "How much work did you put into this course relative to your other courses?" (`G12`)
- "How difficult was this course for you relative to your other courses?" (`G13`)

The prompts at the department-level (with corollary variable codes in parentheses) are:

- "The instructor seemed interested in my progress as a student." (`D1`)
- "The instructor used handouts, audio-visual equipment, and computer applications where appropriate." (`D2`)
- "I was encouraged to visit with my instructor during office hours if I had difficulty with the course material." (`D3`)
- "The instructor encouraged students to use multiple resources (library holdings, etc.) to improve understanding." (`D4`)
- "The instructor asked students to help one another understand ideas and concepts." (`D5`)
- "I felt I gained factual knowledge in this course." (`D6`)
- "I learned important fundamental principles, generalizations, and theories about politics in this course." (`D7`)
- "The instructor stimulated critical and creative thinking about the subject." (`D8`)
- "This course enhanced my ability to write critically." (`D9`)
- "The instructor presented the course material at the appropriate level of difficulty." (`D10`)

An Empirical Assessment of My Teaching Effectiveness

I outline here the approach that I will use to gauge my teaching effectiveness while I have been employed at Clemson University. This approach leverages teaching evaluations data for the Department of Political Science at Clemson University since Fall 2013, complementing it with grade distribution data since Fall 2006, for an analysis that is fundamentally descriptive. As such, the analysis I present here is faithful to the nature of the data but the patterns I describe and the explanations I offer for them are ultimately illustrative.

I logged into Clemson University's Student Assessment of Instructors portal and downloaded the individual-level raw response data for every class I taught since I arrived at Clemson University in the Fall of 2013. Thereafter, I pooled all observations by the unique course I taught, creating four clusters for all teaching evaluations for my classes on introduction to international relations (POSC 1020), quantitative methods in political science (POSC 3410), international conflict (POSC 3610), and U.S. foreign policy (POSC 3630). I calculate semester averages for the entire department using the available information that the Student Assessment of Instructors

---

[4]Clemson University dropped this prompt after the 2019-2020 academic year when it transitioned to another provider for course evaluations. The university also instructed department chairs to not consider this metric in faculty evaluations for the aforementioned concerns of bias. Fall 2019 is the last semester for which this item appears in my evaluations data.
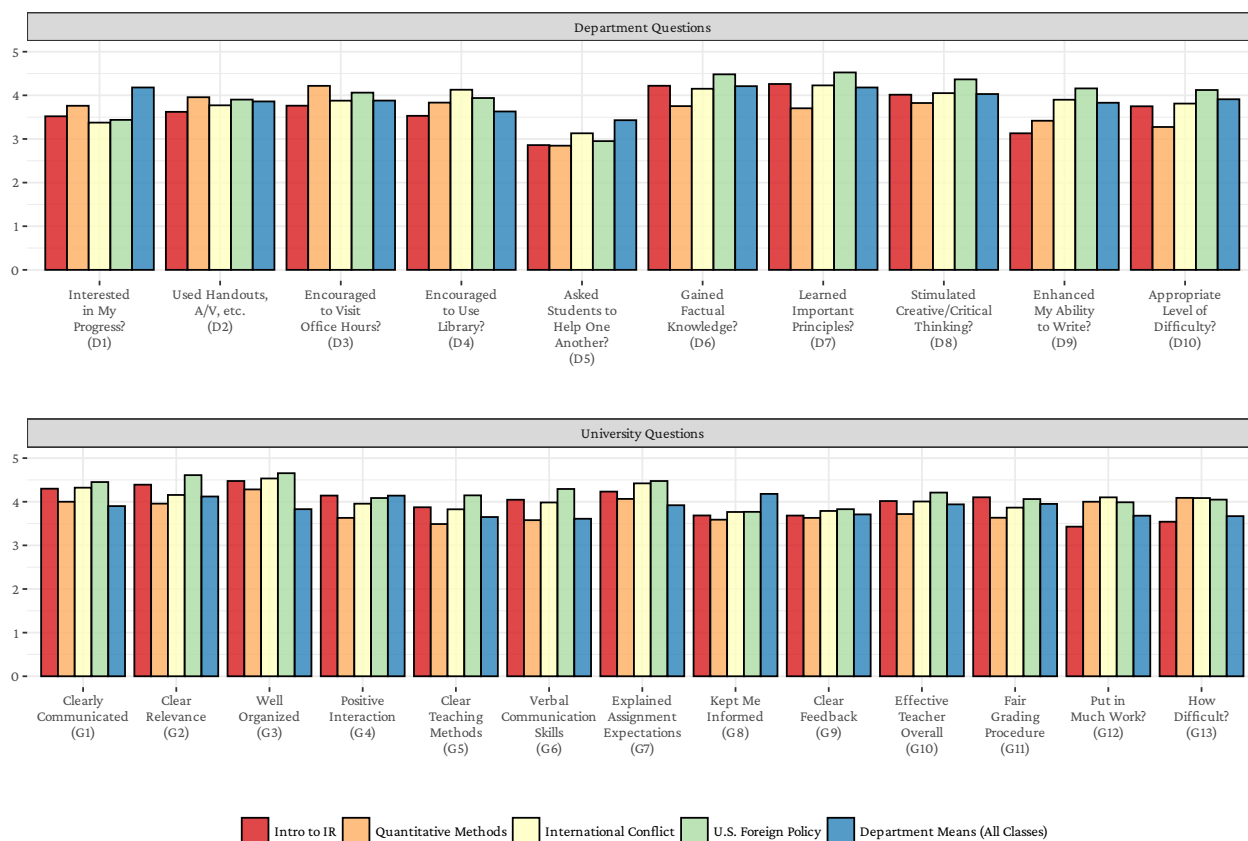
Figure 1: My University-Level and Department-Level Teaching Evaluations, Fall 2013-Fall 2019

portal provides in its written summary report. I then compare the department averages to the averages for each of the unique clusters for the separate courses I teach.[5]

I use these data to create Figure 1, which shows my average evaluation scores by each unique course I have taught relative to the overall department average since Fall 2013.

My Strengths as an Instructor

My teaching record at Clemson University since I arrived in Fall 2013 suggests I have numerous strengths in the classroom that cluster on three categories. First, my evaluations suggest I am a great classroom organizer, excelling at structuring material for a semester and explaining what I expect on class assignments . Second, my evaluations also suggest I rank well as a classroom communicator and that my teaching methods and verbal communication skills ultimately lead to a positive learning environment. Third, my evaluations generally show that students take more from my class and learn more useful skills than they acquire in the typical class. I discusses these in detail below.

I score high across the board as a class organizer. Students evaluate me above the mean on how well I clearly communicate to them what I expect them to learn in the classroom (G1). All my classes score above the average in how well I explain what I expect from them on assignments and tests (G7). A review of my course

---

[5]The use of arithmetic means for data that are fundamentally ordinal is technically an incorrect measure of central tendency but this approach is ubiquitous in administrative assessments of instructors in the classroom.

websites will illustrate that my midterm and final reviews are detailed and make the details of the assignments transparent. I provide documentation for exam-grading policies and I even provide the rubrics that I will use to evaluate their written assignments due at the end of the semester. My classes also all score above the department average for how positively students evaluate the organization of the course (G3). Generally, students across all my classes respect the depth of information I offer them and how well I structure the course to meet the benchmarks I outline at the beginning of the syllabus.

I also score above the department average as a classroom communicator as well. I score above the department average in how well I explain the relevance of the course material (G2) in three of the four different classes I teach. I show similar results in how well my teaching methods (G5) and my verbal communication skills (G6) help students understand the course material. This manifests in an overall assessment of myself as a teacher (G10) that scores above the department average in three of the four unique classes I teach at Clemson University. The only course exception to this trend are the evaluations for my quantitative methods class. These evaluations score just below the department average for all classes on these four metrics.

Third, my record shows that students ultimately acquire meaningful skills and knowledge that university instructors should strive to impart on their students. I am pleased that my upper-division international conflict and U.S. foreign policy courses stimulate an ability to write (D9). Further, my intro-level course scores at the department mean as well. Good writing is an essential skill for all college graduates and especially political science graduates. I take care to communicate tips on my class websites toward that end. Three of my four classes also score above the department average on stimulating creative/critical thinking (D8), a result that follows how much I get students to think about political problems strategically and force them to read and critically evaluate regressions at the upper-division (especially in my international conflict course). Finally, I am pleased at how much students appreciate my efforts to get them into the library and to use library resources (D4). This is a hobby horse of mine. I want students to take advantage of library resources and I routinely fill out class readings with journal articles toward that end. In fact, my international conflict course is taught exclusively with journal articles. Students have remarked to me via email that no other professor they had at Clemson University used the library's resources like I had and that few even encouraged them to go into the campus' library. I take considerable pride in that.

Documenting My Shortcomings and Where I Have Improved on Them

My teaching evaluations since I arrived at Clemson University suggest the following pitfalls in my teaching record. This section outlines these limitations and offers contextualizations of them.

The first trend that is evident from Figure 1 is how my evaluations for the quantitative methods lag behind the department mean and my other classes on almost every metric. There are a few factors, some unique to our department at Clemson University, that can partially account for this trend, though I will highlight one factor in particular. Students tend to do poorly in this class no matter who teaches it. Consider Figure 2, which leverages Clemson University's Grade Distribution Reports data from Fall 2006 to the present to highlight this. Figure 2 shows that the grade distribution for quantitative methods skews much more negative than the grade distribution for any other upper-division course our department has offered since Fall 2006. It is unsurprising that more negative teaching evaluations may follow if student assessments of teaching are in part functions of expected grades and that higher grades generally lead to higher evaluations (e.g. Germaine and Scandura, 2005).[6]

---

[6]This comment implies that our department's quantitative methods class generally receives lower evaluations than other course offerings, which follows informal conversation with colleagues in my department. However, I do not have access to my colleagues' assessments of teaching effectiveness to illustrate this.
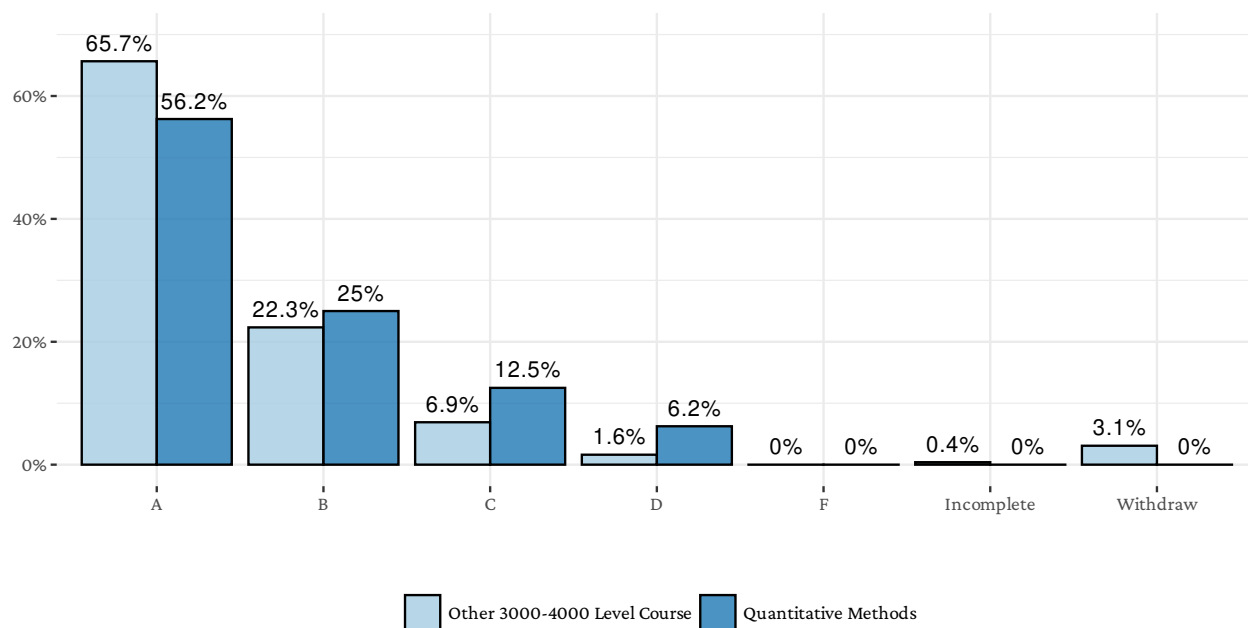
Figure 2: Grade Distribution for Quantitative Methods Relative to Other Department 3000-4000 Level Classes, Fall 2006-Present

The second class of negatives that appear in my teaching evaluations concern how "positive" or approachable I seem. I score below the department average in the department-level question about how interested I seem in the progress students make during the course (D1). I also score below the mean on the university-level question about the positive interaction I offer in class (G4). These scores are concerning and made more curious by the fact that I generally score well on the department-level question about how often I encourage students to come to my office hours (D3) and I score around the mean on how clear my feedback is on assignments (G9). I took inventory of how students interpreted my presence in the class as less than positive and viewed me as less interested in their progress and made important corrections to my classroom approach starting in the Fall 2018 semester.

Table 1 shows the improvements I have made beginning with the Fall 2018 semester relative to the semesters before it on three items of interest. These are the means for how well I showed interest in the student's progress (D1), the positive interaction with the class and the instructor (G4), and how well students evaluated me as an effective teacher overall (G10). I have made a little progress increasing the mean of evaluations for how much interest I show in the student's progress in more recent semesters at Clemson University relative to previous semesters. There is much greater improvement on the other two items, though. I have made important strides in increasing the positive interaction I have with the class in more recent semesters relative to previous semesters. That beneficial change may also explain why students think of me as a more effective teacher overall as well.

Table 1: A Comparison of Means for Three Items of Interest from Before and After Fall 2018

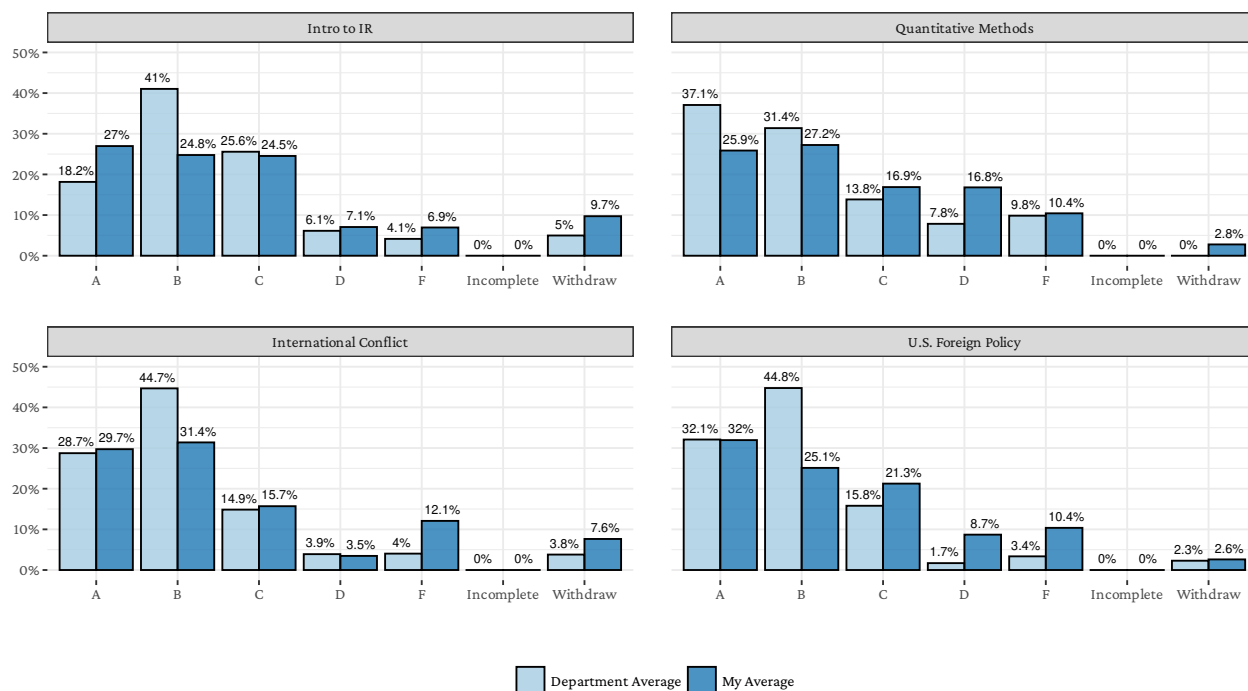|	**Interested in My Progress?**	**Positive Interaction?**	**Effective Teacher Overall?**
Before Fall 2018	3.43	3.86	3.91
Fall 2018 and After	3.61	4.22	4.15

Figure 3: Grade Distribution for My Classes Relative to the Average Political Science Class, Fall 2013-Fall 2017

The third class of negatives from my evaluations concern the level of difficulty and the effort I require from students in my class. Generally, students say they put in much more work relative to their other courses (G12) and that my course is much more difficult than other courses a student takes (G13). This also partly manifests in assessments of whether the difficulty of my classes is appropriate (D10). My intro-level course on international relations and quantitative methods class score particularly low on this metric, coinciding with informal feedback from students that I demand too much in the classroom.

Figure 3 offers illustrative evidence for the extent of my in-class expectations relative to what students may get from other offerings.[7] This graph compares the grade distribution for each unique class type I have taught since I arrived at Clemson University relative to the grade distribution for the class when some other instructor teaches it. On average, I give fewer As and Bs than the typical instructor in the typical class. 73% of students in the typical class taught by some instructor other than myself gets an A or a B while just over 56% of students get an A or B in classes I teach. This can explain why students generally say I am more difficult than other instructors and why, across all classes I teach, they say they have to put in much more work relative to other courses. However, I think there is value in my approach. Harder classes for students born of greater expectations from the instructor will be a more informative experience for the student in the long run. This is one reason why it makes sense that my classes score higher in difficulty relative to other courses (G13) and how much work students feel they must put into my courses (G12), but my approach may also account for why students rate me highly for how much factual knowledge they gain (D6), how well I stimulate creative and critical thinking (D8), and how much my classes enhance the student's ability to write (D9).

---

[7]Per agreement between the Office of the Provost and the Faculty Senate, the grade distribution report was taken off-line as of August 6, 2018 for an indefinite period in order that a broader campus discussion can occur regarding its availability, access details, and wise use by University citizens. My data end with the Fall 2017 semester.

Conclusion

This document served as an assessment of my teaching effectiveness, leveraging course evaluation data from Clemson University to highlight and discuss my strengths and weaknesses in the classroom. I score highly as a class organizer and classroom communicator. My evaluations suggest that students acquire important and meaningful skills from my classes as well and that students learn good writing, critical thinking, and how to use their library's resources. My shortcomings in the classroom cluster on my evaluations for quantitative methods, how "positive" or approachable I seem, and how difficult my classes are relative to other department or university offerings. I offered contextualizations for these shortcomings and even offer evidence of recent improvement in important metrics, especially positive in-class interaction. The findings on the balance demonstrate teaching effectiveness in the classroom, concentrated in multiple and important categories, that signal competence in the classroom.

References

Germaine, Marie-Line and Terri A. Scandura. 2005. "Grade Inflation And Student Individual Differences as Systematic Bias in Faculty Evaluations." *Journal of Instructional Psychology* 32(1):58–67.

McKeachie, Wilbert J. 1997. "Student Ratings: The Validity of Use." *American Psychologist* 52(11):1218–1225.

Nulty, Duncan D. 2008. "The Adequacy of Response Rates to Online and Paper Surveys: What Can Be Done?" *Journal Assessment & Evaluation in Higher Education* 33(3):301–314.