Underrepresented Groups in the Economics Major
The Impact of Using Team-Based Learning in the Principles Course

Marcelo Clerici-Arias
Stanford University
marcelo@stanford.edu

Does the implementation of team-based learning in principles of economics attract more women as well as racial and ethnic minorities into the classroom? For the past two years we have experimented with using TBL in some principles courses at Stanford University, and we have observed a marked increase in the fraction of women and racial/ethnic minorities. Using difference-in-difference methods to attempt to identify causality, we find that TBL is associated with an increase of about 9 points in female participation and about 18 points in non-white student participation in principles of economics.

ASSA 2020 - American Economic Association
Team-Based Learning in Economics
Draft - December 27, 2019[1]

[1] This is work in progress. Contact the author at marcelo@stanford.edu for the latest version of this research.

Special thanks to my colleague Mark Tendall for first noticing the change in student composition after TBL, for providing raw student data for those quarters when I did not teach, and for his constant support and encouragement to continue experimenting to improve the teaching and learning of Economics 1 at Stanford University.

Introduction

As instructors, do our choices of class structure and teaching techniques affect who ends up taking our courses? In particular, can we improve the representation of women and racial/ethnic minorities in economics just by changing the way we teach? We attempt to answer these questions in the context of the principles of economics course at Stanford University.

As economists we are very familiar with unintended consequences, and this paper is a great example of such phenomenon. When I started considering using team-based learning in principles of economics back in 2015, our goal was to improve the learning of economics. I did not intend to influence the gender or racial/ethnic makeup of our classes. In fact, I did not even consciously think about this dimension at all. For good or for bad, I am kind of blind to gender, race, or ethnicity. When I see my students, I see people learning, each of them with their own personalities, talents and other strengths, as well as challenges, and I do not define them by their belonging to a certain socioeconomic class, gender, religion, race, or ethnicity.

After the first couple of classes in the first TBL offering of Econ 1 in the spring of 2017-18, my colleague Mark Tendall was the one who first mentioned to me the differences in student population that he casually observed in this cohort versus other Econ 1 courses. That led me to actually count students by gender and race/ethnicity in a few Econ 1 offerings, and after verifying the initial differences, to study the issue further.

In the last decade economists have become more aware and more engaged than ever in the issue of the participation of women and racial/ethnic minorities in economics, and how they are treated first as students, and then as professional economists.

Bayer and Rouse (2016) highlight that only “23.5% of tenured and tenure-track faculty in economics are women,” and only 15% of full professors in economics departments. Bayer and Rouse find that the gender gap is much worse in economics than in other social sciences as well as in other math-intensive fields, when measured in several dimensions, such as tenure rates, promotion, salaries, and job satisfaction. As Bayer and Rouse state, “minority academic economists are even rarer.”

Bayer and Rouse trace some of these disparities back to Ph.D. programs and, more relevant to this paper, to the fraction of undergraduates majoring in economics who are female or racial/ethnic minorities.

In their paper, Bayer and Rouse describe different barriers to diversity, starting at the undergraduate level. Math preparation does not appear to be an issue. On the other hand, Bayer and Rouse cite several other studies on a variety of factors that seem to affect the participation of females and underrepresented minorities in the economics major: self-reported interest, with women reporting twice as frequently as men that they did not take economics in their first year because they did not find it interesting (Dynan and Rouse, 1997); responses to relative grades, where females seem to be more sensitive than men to lower grades in introductory economics courses and select other majors (Rask and Tiefenthaler, 2018); the gender and race/ethnicity of the instructor (Dynan and Rouse, 1997; Fairlie, Hoffman, and Oreopoulos, 2014); implicit bias (Milkman, Akinola, and Chugh, 2015).

For more information on the topic, check the site created by Bayer (2011), dedicated to the study and dissemination of information and best practices on diversity in economics.

Avilova and Goldin (2018) describe their Undergraduate Women in Economics project, designed to learn "why women do not major in economics to the same degree as men and what can be done to change that." After collecting initial data on the male-to-female economics major ratio, Goldin created a randomized control trial where the treated schools have three types of intervention: better information about economics as a field and the career paths open to econ majors; better mentoring, support, and role models; and better instructional content and presentation style "to improve introductory economics courses and to make them more relevant to a wider range of students." Using team-based learning with relevant real-life applications in principles in economics fits in this third goal of Goldin's project, though Stanford is not participating in it. UWE is a long-term project and it will take several years to measure its impact on how many women major in economics.

Bayer and Rouse also cite Freeman et al (2014) on the positive impact of active learning (compared to traditional lecturing) on exam scores and failure rates, particularly for "students from disadvantaged backgrounds and for women in male-dominated fields." However, none of these authors study the impact of the use of active learning techniques on the decision of a student to take a course in the first place, and that is the thrust of our paper.

Background

This paper does not focus on team-based learning as a teaching methodology. For a thorough description of team-based learning in a general context, I recommend Michaelsen et al (2004). For team-based learning in the social sciences, Sweet and Michaelsen (2012) is a great resource, in particular Espey's chapter on team-based learning in economics. Imazeki (2015) is a wonderful summary of TBL in economics.

Previous papers have focused on the impact of active learning techniques (and team-based learning in particular) on learning outcomes. A few of those papers have focused on the specific impact on the learning outcomes of minority students relative to the overall student population. For example, Haak et al (2011) find that a highly structured course in introductory biology with a high active learning component improved performance for students in general, and it did so disproportionately for students for disadvantaged backgrounds, most of them underrepresented minorities. Also see the previously mentioned Freeman et al (2014) for a meta-analysis of papers comparing a variety of active learning approaches against traditional lecturing.

McNeil et al (2019) study the use of team-based learning and SCALE-UP (a classroom environment designed for collaborative learning, similar to the physical used in the TBL version of Econ 1 at Stanford). Their preliminary results show a positive impact on all students, and the narrowing of the gap for racial/ethnic minorities (not the gender gap, though).

In economics and using team-based learning in particular, Hettler (2015) finds "a small, significant improvement in learning outcomes for low-income and minority students when compared to others" in a population that included students in principles of economics courses as well as undergraduate and MBA students in quantitative methods courses.

Espey (2018) focuses, among other things, on the impact of team heterogeneity on individual and team performance in the context of multiple courses of introductory microeconomics taught using team-based learning. One of the aspects of team heterogeneity that Espey analyzes is what she calls surface-level variables, which include gender, college class (frosh, sophomore, junior, senior), geographic diversity (in- vs. out-of-state students), and race/ethnicity (though Espey does not include the latter in her analysis). Espey's results are mixed, but they do highlight the importance of being aware of the composition of our student populations, and in particular diversity at the team and class levels.

Again, these studies focus on learning outcomes across groups, and none of them analyze the impact of using team-based learning or active learning more broadly on the type of students who are attracted to these courses versus others based on lecturing.

Study design

Had the goal of this paper been to study the impact of the use of team-based learning in the principles of economics course on the learning of economics, we would be concerned about how selection effects (students choosing among different offerings of Econ 1 throughout the year) impacted the characteristics of the student population in each course, thus biasing the results in some way.

However, in this paper we are studying how TBL vs. non-TBL classes may attract some types of students disproportionately, so we are actually interested in highlighting and studying that self-selection, not correcting for it.

Since Autumn 2013, principles of micro and macroeconomics have been taught in a single, one-quarter (10 weeks) course called Econ 1. Almost 4,000 students participated in Econ 1 during the six academic years starting in Autumn 2013 and ending in Spring 2019. I do not include the summer version of Econ 1 since most students taking summer courses are not Stanford students, the class is not taught by Stanford faculty, and even the content of the course can be significantly different.

Table 1 shows the basic descriptive statistics for the 21 Econ 1 courses that were taught during those six years, only two of them in team-based learning format, and that was during the last two spring quarters. Appendix A contains a table with detailed statistics for each of the 21 offerings of Econ 1 during this period.

Econ 1 - 2013-2019	Non-TBL courses	TBL courses	All courses
Number of courses	19	2	21
Number of students	3801	179	3980
Female %	45.0	59.2	45.6
Non-white %	47.8	70.9	48.8
Non-white, non-asian %	21.9	36.3	22.6

Table 1. Summary statistics for Econ 1, 2013-2019

A quick comparison of some key statistics immediately reveals a stark contrast in female participation (+14.2 points for TBL versions of Econ 1), non-white student participation (+23.1 points for TBL courses), and non-white/non-asian student participation (+14.4 points for TBL courses). That dramatic change that was clearly visible from just looking at the students in the classroom is what led to this research.

Of course this gross comparison can disguise other reasons for these observations, and this section attempts to determine as well as possible whether team-based learning in Econ 1 at Stanford actually had a positive impact on the participation of females and racial/ethnic minorities.

Part of the observed differences may be due to students choosing to take a TBL version of Econ 1 over the non-TBL offering, or vice versa. Furthermore, the changes may also capture students who would have never taken Econ 1 under the original lecture format, and now choose the TBL version of Econ 1 over not taking Econ 1 at all (note that the opposite cannot be true—if a student does not like the TBL version of Econ 1, they can always go to a non-TBL class, whereas the TBL offering has only been available since 2017-18, and then only in spring quarters). This is the kind of information we want to capture in our analysis.

Note that changes in the gender, race/ethnicity, and college class composition of Econ 1 can be due to other reasons. For example, different instructors may attract different students, whether it is because of the personality, gender, race/ethnicity, or other characteristics of the instructor. Or students' schedules may push them to take Econ 1 in one quarter or time of day over another. For example, students who are taking the Human Biology core at Stanford (usually

in their sophomore year) have their mornings busy all quarters and can only take afternoon offerings of Econ 1. In our analysis we will want to filter out changes in female and racial/ethnic participation that may be due to the instructor or the timing of the course.

As background of how things work at Stanford, in April the teaching matrix for the following year is created, and by early August students can see what courses are being taught and by whom, and plan accordingly. The current structure of Econ 1, combining principles of micro and macroeconomics in one quarter, has been taught every quarter since 2013-14, and each quarter is taught by a different instructor (in Autumn 2015-16 two instructors co-taught Econ 1). Once a year during the first half of this six-year period, Econ 1 was taught twice at different times of the day (one in the morning, one in the afternoon), but even then the same instructor (or pair of instructors in Autumn 2015-16) taught the two offerings in the same quarter.

Faculty assignments are not random in at least two ways:

(1) Faculty express a preference on what quarter (and what courses) to teach. After quarter assignments are made, day and time of day are chosen by the instructor, with few significant constraints.

(2) Faculty choose what teaching techniques to use, the type of assessments, grading criteria, and even the content to teach (up to a certain extent), with few significant constraints.

The same three instructors taught Econ 1 in the last three years (2019-20 will be the fourth year with the same lineup), always in the same quarter, and only Clerici-Arias expressed an interest in teaching using team-based learning.

When team-based learning was introduced in 2017-18, almost a year in advance the department of Economics publicized that the class format would be different in the spring of 2017-18, the first time TBL was offered, so students could make an informed choice of whether and when to take Econ 1. The two class formats differed in several significant ways, as described in Table 2.

	Non-TBL traditional format	TBL format
Class hours per week	4 hours of lecture with instructor plus 1 hour with teaching assistant.	6 hours of class in TBL format led by instructor with the assistance of two TAs.
Teaching technique	Mostly lecture, though in some quarters clickers are used to encourage student thinking and problem-solving in class. The TA sections are limited to 15 students each, and tend to focus on active learning techniques, with occasional reviews and mini-lectures.	Rigorous use of team-based learning, with all its embedded structure.
Attendance	Lecture attendance is not required. Section attendance is mandatory, with a small part of the grade dedicated to class participation.	Mandatory.
Classroom format	Auditorium with 600 seats (400 in lower floor, 200 in an unused upper deck).	Flexible classroom with 150 seats.
Timing	Offered in fall and winter quarters, in the mornings.	Offered in spring quarters, in the afternoons.
Instructors	Taught by several instructors, including Clerici-Arias.	Only taught by Marcelo Clerici-Arias.

Table 2. Key differences between non-TBL and TBL Econ 1 offerings

To attempt to establish a causal link between the introduction of team-based learning in Econ 1 and the increase in female and racial/ethnic minorities participation rates, we have chosen to use differences-in-differences, a statistical technique popularized in economics in the last three decades or so, though its origins go back to the nineteenth century (Angrist and Pischke, 2009).

We will first focus on female participation rates in Econ 1, using two different treatment groups. First, all the Econ 1 courses taught by Marcelo Clerici-Arias, without TBL for the first four years and with TBL for the last two, attempting to account for differences across instructors. Then, we run the analysis again using spring quarters as the treated group, without TBL for the first four years and with TBL for the last two.

Finally, we will focus on the non-white student participation rate, again using the two treatment groups described above.

Female participation rate - Clerici-Arias as treatment group

A key assumption of differences-in-differences is that the trends of treatment and control groups are parallel prior to the intervention, so we can then make inferences about the differences in changes in trends after the intervention. Figure 1 shows that the trends in female participation rate in Econ 1 prior to the introduction of team-based learning in spring 2018 were roughly the same in courses taught by either Clerici-Arias or other instructors, hinting that using differences-in-differences is appropriate. Female participation rate increased for both treatment and control groups in the post-treatment period, and this is exactly what differences-in-differences is designed to capture.

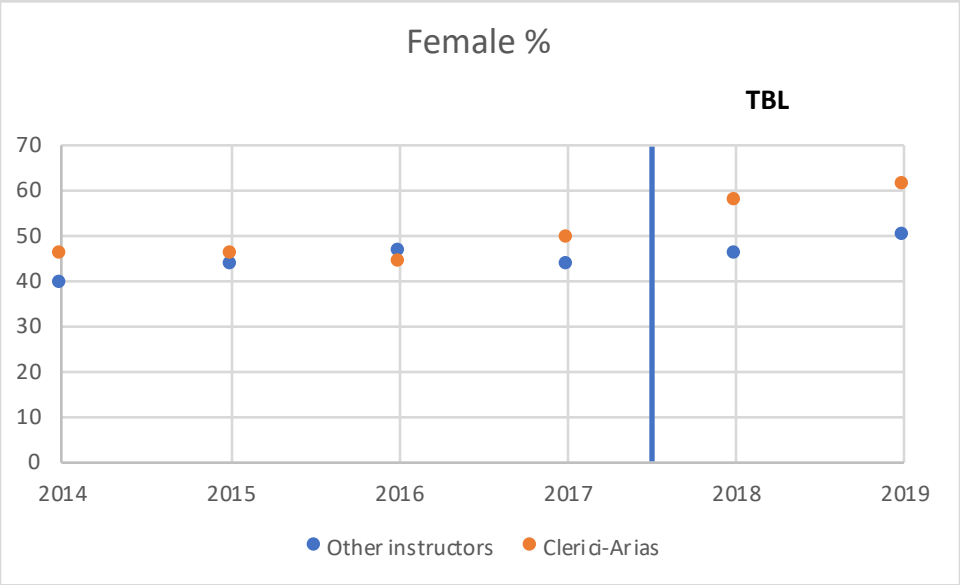


Figure 1. Female participation rate in Econ 1 by instructor

The following equation describes the differences-in-differences regression. Our main interest is the coefficient δ that shows the increase in the female participation rate due to the implementation of team-based learning.

$$Y_{Ct} = \alpha + \beta * Treat_{TBL} + \gamma * Post_t + \delta * (Treat_{TBL} * Post_t) + \varepsilon_{Ct}$$

Running diff in Stata with robust standard errors tells us that team-based learning in Econ 1 **increased the female participation ratio by 8.9 points** by itself, after accounting for other instructors and changes in the general trend starting in 2017-18 (Table 3).

```
. diff female, t(mca) p(post) robust

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS
Number of observations in the DIFF-IN-DIFF: 21
            Before         After
   Control: 8              4           12
   Treated: 7              2           9
            15             6
------------------------------------------------------------
 Outcome var.   | female  | S. Err. |   |t|   |  P>|t|
----------------+---------+---------+---------+---------
Before          |         |         |         |
   Control      | 43.713  |         |         |
   Treated      | 45.813  |         |         |
   Diff (T-C)   | 2.100   | 2.208   | 0.95    | 0.355
After           |         |         |         |
   Control      | 48.131  |         |         |
   Treated      | 59.170  |         |         |
   Diff (T-C)   | 11.039  | 2.869   | 3.85    | 0.001***
                |         |         |         |
Diff-in-Diff    | 8.939   | 3.620   | 2.47    | 0.024**
------------------------------------------------------------
R-square:    0.56
* Means and Standard Errors are estimated by linear regression
**Robust Std. Errors
**Inference: *** p<0.01; ** p<0.05; * p<0.1
```

Table 3. Stata results for differences-in-differences with robust standard errors. Treatment group: courses taught by Clerici-Arias

Bertrand et al (2004) point out that conventional standard errors can "severely underestimate the standard deviation of the estimators" and suggest bootstrapping as a way of obtaining more accurate standard errors. Rokicki et al (2018) suggest that when the number of groups is small (they define it as less than 20; our number of groups is larger than that, but very close to that number), then the standard errors can also be underestimated. Rokicki et al suggest a number of possible corrections, bootstrapping among them. For both reasons, we also run diff-in-diff with bootstrapping, as shown on Table 4. Fortunately, the results hold, even more strongly than without bootstrapping.

```
. diff female, t(mca) p(post) bs rep(50)
(running regress on estimation sample)

Bootstrap replications (50)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
.....x.....................................xx....    50

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS
Number of observations in the DIFF-IN-DIFF: 21
            Before         After
   Control: 8              4           12
   Treated: 7              2           9
            15             6
Bootstrapped Standard Errors

------------------------------------------------------------
 Outcome var.   | female  | S. Err. |   |t|   |  P>|t|
----------------+---------+---------+---------+---------
Before          |         |         |         |
   Control      | 43.713  |         |         |
   Treated      | 45.813  |         |         |
   Diff (T-C)   | 2.100   | 2.647   | 0.79    | 0.428
After           |         |         |         |
   Control      | 48.131  |         |         |
   Treated      | 59.170  |         |         |
   Diff (T-C)   | 11.039  | 2.718   | 4.06    | 0.000***
                |         |         |         |
Diff-in-Diff    | 8.939   | 3.764   | 2.38    | 0.018**
------------------------------------------------------------
R-square:    0.56
* Means and Standard Errors are estimated by linear regression
**Inference: *** p<0.01; ** p<0.05; * p<0.1
```

Table 4. Stata results for differences-in-differences with robust standard errors and bootstrapping. Treatment group: courses taught by Clerici-Arias

Female participation rate - Spring quarter as treatment group

The results are similar when we use spring quarter as the treatment group, to account for differences in the timing of the offering of the team-based learning version of Econ 1. Figure 2 shows that the trends in female participation rate in Econ 1 prior to the introduction of team-based learning in spring 2018 were roughly similar in courses taught in the spring versus other quarters, despite an uptick in spring 2016. Running diff in Stata shows that the **female participation rate increased by 9.3 points** (Table 5), consistent with our results when controlling by instructor in the previous section. The results continue to be statistically significant even when using bootstrapping (Table 6).

The slightly different results are due to Clerici-Arias having higher female participation rates 2.1 points than other instructors before team-based learning was implemented, while spring quarters had female participation rates only 1.8 points higher than other quarters before TBL implementation.

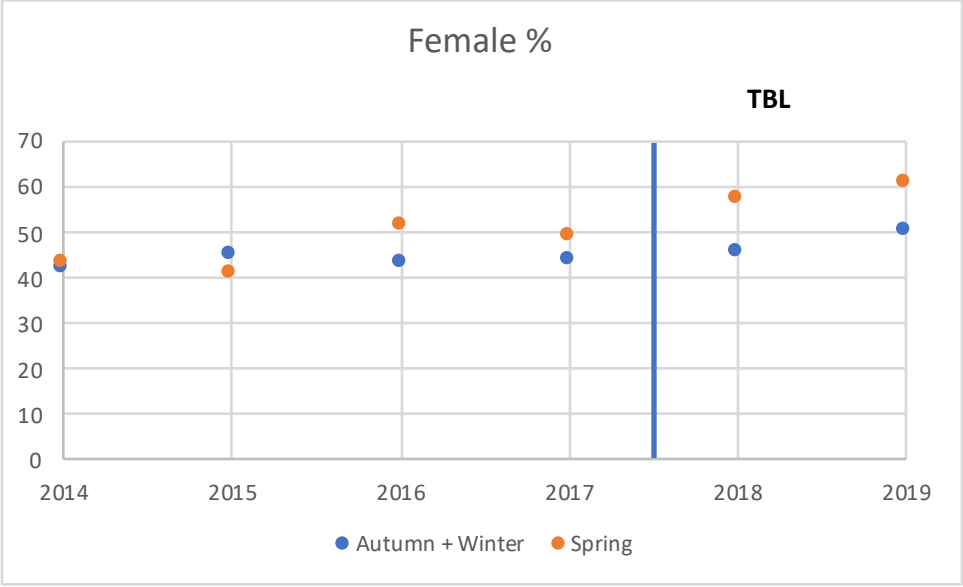


Figure 2. Female participation rate in Econ 1 by quarter

```
. diff female, t(spring) p(post) robust

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS
Number of observations in the DIFF-IN-DIFF: 21
            Before         After
   Control: 11             4           15
   Treated: 4              2           6
            15             6
------------------------------------------------------------
 Outcome var.   | female  | S. Err. |   |t|   |  P>|t|
----------------+---------+---------+---------+---------
Before          |         |         |         |
   Control      | 44.219  |         |         |
   Treated      | 45.997  |         |         |
   Diff (T-C)   | 1.778   | 2.701   | 0.66    | 0.519
After           |         |         |         |
   Control      | 48.131  |         |         |
   Treated      | 59.170  |         |         |
   Diff (T-C)   | 11.039  | 2.869   | 3.85    | 0.001***
                |         |         |         |
Diff-in-Diff    | 9.262   | 3.940   | 2.35    | 0.031**
------------------------------------------------------------
R-square:    0.55
* Means and Standard Errors are estimated by linear regression
**Robust Std. Errors
**Inference: *** p<0.01; ** p<0.05; * p<0.1
```

Table 5. Stata results for differences-in-differences with robust standard errors. Treatment group: courses taught in spring quarter

```
. diff female, t(spring) p(post) bs rep(50)
(running regress on estimation sample)

Bootstrap replications (50)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................x..x...    50

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS
Number of observations in the DIFF-IN-DIFF: 21
            Before         After
   Control: 11             4           15
   Treated: 4              2           6
            15             6
Bootstrapped Standard Errors

------------------------------------------------------------
 Outcome var.   | female  | S. Err. |   |t|   |  P>|t|
----------------+---------+---------+---------+---------
Before          |         |         |         |
   Control      | 44.219  |         |         |
   Treated      | 45.997  |         |         |
   Diff (T-C)   | 1.778   | 2.578   | 0.69    | 0.490
After           |         |         |         |
   Control      | 48.131  |         |         |
   Treated      | 59.170  |         |         |
   Diff (T-C)   | 11.039  | 2.736   | 4.03    | 0.000***
                |         |         |         |
Diff-in-Diff    | 9.262   | 3.842   | 2.41    | 0.016**
------------------------------------------------------------
R-square:    0.55
* Means and Standard Errors are estimated by linear regression
**Inference: *** p<0.01; ** p<0.05; * p<0.1
```

Table 6. Stata results for differences-in-differences with robust standard errors and bootstrapping. Treatment group: courses taught in spring quarter

Non-white student participation rate - Clerici-Arias as treatment group

Now we switch to non-white student participation rate as the dependent variable. In a similar way to the previous section, we first use the courses taught by Clerici-Arias as the treatment group, and then we explore the courses taught in spring quarter as the treatment group.

Figure 3 shows that the trends in courses taught by Clerici-Arias versus other instructors prior to the implementation of team-based learning were roughly similar.

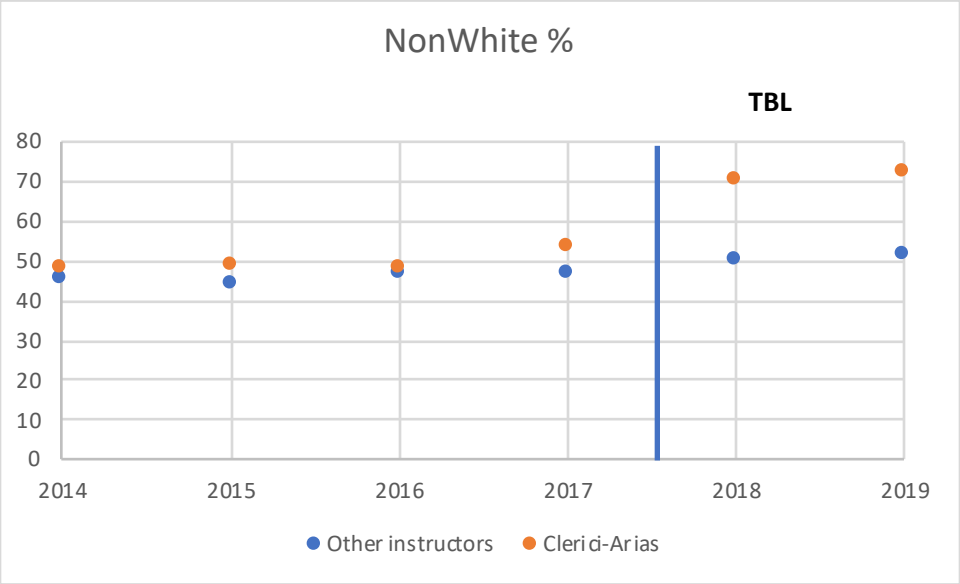


Figure 3. Non-white student participation rate by instructor

We ran differences-in-differences in Stata, first with robust standard errors (Table 7), then adding bootstrapping (Table 8), and we find that the **participation rate of non-white students in Econ 1 increased by 17.8 points** due to the implementation of team-based learning.

```
. diff nw, t(mca) p(post) robust

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS
Number of observations in the DIFF-IN-DIFF: 21
            Before         After
   Control: 8              4           12
   Treated: 7              2           9
            15             6
------------------------------------------------------------
 Outcome var.   | nw      | S. Err. |   |t|   |  P>|t|
----------------+---------+---------+---------+---------
Before          |         |         |         |
   Control      | 45.754  |         |         |
   Treated      | 48.194  |         |         |
   Diff (T-C)   | 2.440   | 3.400   | 0.72    | 0.483
After           |         |         |         |
   Control      | 50.658  |         |         |
   Treated      | 70.927  |         |         |
   Diff (T-C)   | 20.269  | 2.131   | 9.51    | 0.000***
                |         |         |         |
Diff-in-Diff    | 17.829  | 4.013   | 4.44    | 0.000***
------------------------------------------------------------
R-square:    0.67
* Means and Standard Errors are estimated by linear regression
**Robust Std. Errors
**Inference: *** p<0.01; ** p<0.05; * p<0.1
```

Table 7. Stata results for differences-in-differences with robust standard errors. Treatment group: courses taught by Clerici-Arias

```
. diff nw, t(mca) p(post) bs rep(50)
(running regress on estimation sample)

Bootstrap replications (50)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..x.....xx....x.xx...........xx.x.................    50

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS
Number of observations in the DIFF-IN-DIFF: 21
            Before         After
   Control: 8              4           12
   Treated: 7              2           9
            15             6
Bootstrapped Standard Errors

------------------------------------------------------------
 Outcome var.   | nw      | S. Err. |   |t|   |  P>|t|
----------------+---------+---------+---------+---------
Before          |         |         |         |
   Control      | 45.754  |         |         |
   Treated      | 48.194  |         |         |
   Diff (T-C)   | 2.440   | 3.641   | 0.67    | 0.503
After           |         |         |         |
   Control      | 50.658  |         |         |
   Treated      | 70.927  |         |         |
   Diff (T-C)   | 20.269  | 1.839   | 11.02   | 0.000***
                |         |         |         |
Diff-in-Diff    | 17.829  | 3.951   | 4.51    | 0.000***
------------------------------------------------------------
R-square:    0.67
* Means and Standard Errors are estimated by linear regression
**Inference: *** p<0.01; ** p<0.05; * p<0.1
```

Table 8. Stata results for differences-in-differences with robust standard errors and bootstrapping. Treatment group: courses taught by Clerici-Arias

Non-white student participation rate - Spring quarter as treatment group

The same analysis, now using spring quarter as treatment group, yields similar results. Figure 4 shows the trends before and after team-based learning was implemented. Again similar, though with a bit of an uptick in spring 2017.

The **participation rate of non-white students in Econ 1 increased by 20.4 points** due to the implementation of team-based learning when using spring quarter as treatment group. Compare to 17.8 points when using Clerici-Arias as treatment group. That difference is due to Clerici-Arias having non-white student participation rates 2.4 points higher than other instructors even before team-based learning was implemented.

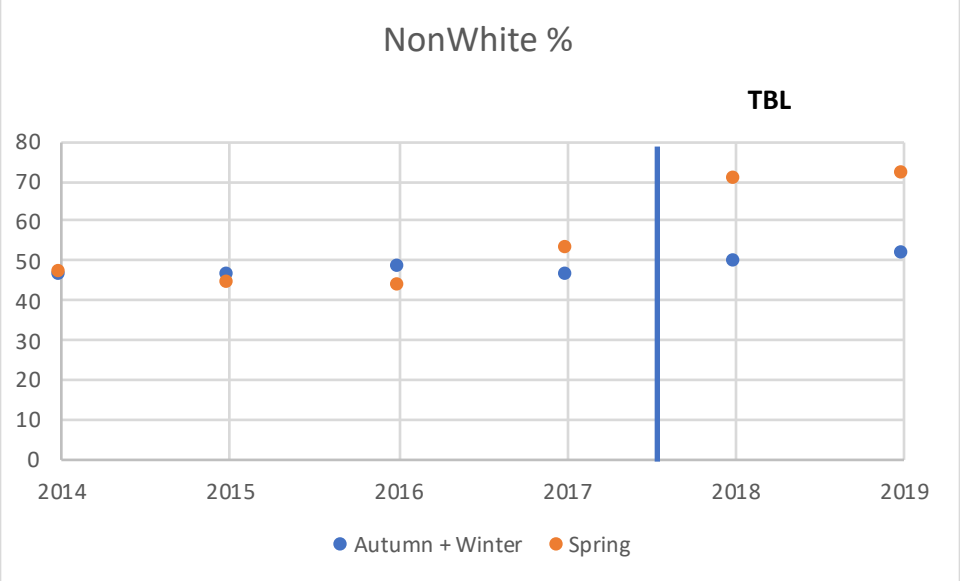


Figure 2. Non-white student participation rate in Econ 1 by quarter

```
. diff nw, t(spring) p(post) robust

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS
Number of observations in the DIFF-IN-DIFF: 21
            Before         After
   Control: 11             4           15
   Treated: 4              2           6
            15             6
------------------------------------------------------------
 Outcome var.   | nw      | S. Err. |   |t|   |  P>|t|
----------------+---------+---------+---------+---------
Before          |         |         |         |
   Control      | 46.918  |         |         |
   Treated      | 46.825  |         |         |
   Diff (T-C)   | -0.093  | 3.002   | -0.03   | 0.976
After           |         |         |         |
   Control      | 50.658  |         |         |
   Treated      | 70.927  |         |         |
   Diff (T-C)   | 20.269  | 2.131   | 9.51    | 0.000***
                |         |         |         |
Diff-in-Diff    | 20.362  | 3.681   | 5.53    | 0.000***
------------------------------------------------------------
R-square:    0.65
* Means and Standard Errors are estimated by linear regression
**Robust Std. Errors
**Inference: *** p<0.01; ** p<0.05; * p<0.1
```

Table 9. Stata results for differences-in-differences with robust standard errors. Treatment group: courses taught in spring quarter

```
. diff nw, t(spring) p(post) bs rep(50)
(running regress on estimation sample)

Bootstrap replications (50)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................x.........x.....x...    50

DIFFERENCE-IN-DIFFERENCES ESTIMATION RESULTS
Number of observations in the DIFF-IN-DIFF: 21
            Before         After
   Control: 11             4           15
   Treated: 4              2           6
            15             6
Bootstrapped Standard Errors

------------------------------------------------------------
 Outcome var.   | nw      | S. Err. |   |t|   |  P>|t|
----------------+---------+---------+---------+---------
Before          |         |         |         |
   Control      | 46.918  |         |         |
   Treated      | 46.825  |         |         |
   Diff (T-C)   | -0.093  | 3.338   | -0.03   | 0.978
After           |         |         |         |
   Control      | 50.658  |         |         |
   Treated      | 70.927  |         |         |
   Diff (T-C)   | 20.269  | 2.206   | 9.19    | 0.000***
                |         |         |         |
Diff-in-Diff    | 20.362  | 4.469   | 4.56    | 0.000***
------------------------------------------------------------
R-square:    0.65
* Means and Standard Errors are estimated by linear regression
**Inference: *** p<0.01; ** p<0.05; * p<0.1
```

Table 10. Stata results for differences-in-differences with robust standard errors and bootstrapping. Treatment group: courses taught in spring quarter

Discussion

These results strongly suggest that the introduction of team-based learning in Stanford's principles of economics course had a very significant impact on the participation of women and racial/ethnic minorities. We are not just talking about statistical significance, but also very sizable effects, with female students going from being less than half to almost 60% of the Econ 1 population, and non-white students going from less than half to 70% of the Econ 1 population.

Note that other course characteristics changed concurrently with the appearance of team-based learning in Econ 1. For example, these two courses were taught in a different classroom with flexible furniture and an audio system that facilitates student communication with teams as well as across the room, very much unlike the large auditorium where Econ 1 is usually taught. Also, the class size (87 and 92 students in the two only TBL offerings) was small relative to the other versions of Econ 1, which ranged from 129 to 375 students. These changes could have had an impact independent from team-based learning.

Also, these are just two years (really only two courses) worth of TBL implementation in Stanford's Econ 1 course. We need more experiences/data to see if these findings hold. As we get more data at Stanford, I will be happy to update the results.

Furthermore, these are the results of just one institution (and not a typical/representative one). I encourage the other researchers who participated in the National Science Foundation team-based learning project to expand their studies to include the impact of TBL on women and racial/ethnic minorities.

As we add more years worth of data, we plan on expanding our research to study the impact on the number and composition of majors, as well as the impact on performance and participation in intermediate micro and macroeconomics.

Conclusions

Addressing the underrepresentation of women and racial/ethnic minorities in economics most certainly requires a multi pronged approach, probably starting at K12, most certainly in the undergrad and Ph.D. periods, and undoubtedly in the professional years, in academia, government, corporations, and non-profit organizations.

The way we teach principles of economics and the way students perceive this course are small but very significant pieces of this puzzle. This paper shows that the use of a collaborative, real-life application/problem-based approach to teaching and learning principles of economics has the potential to open up the world of economics to many more women and racial/ethnic minorities. Prior literature has shown usually better (and at least equal) learning outcomes when using appropriate active learning approaches, in particular collaborative and cooperative learning. When you pair these two results—better learning and a more inclusive classroom with a more diverse student population—implementing these techniques in our courses certainly is a Pareto improvement.

Of course further research is necessary, but in the meantime I think there is enough evidence that indicates that we should redesign our courses and put more emphasis on active learning, collaboration, and real-life applications to open up the bottleneck of the principles of economics course that currently seems to dissuade many students—particularly women and racial/ethnic minorities—from embracing economics and making it a lifelong journey.

References

Angrist, Joshua D., and Jörn-Steffen Pischke. Mostly Harmless Econometrics. An Empiricist's Companion. Princeton University Press, 2009.

Avilova, Tatyana, and Claudia Goldin. "What Can UWE Do for Economics?" AEA Papers and Proceedings 108 (2018): 186–90. https://doi.org/10.1257/pandp.20181103.

Bayer, Amanda. "Diversifying Economic Quality. American Economic Association Committee on the Status of Minority Groups in the Economics Profession (CSMGEP).," 2011. http://diversifyingecon.org.

Bayer, Amanda, and Cecilia Elena Rouse. "Diversity in the Economics Profession: A New Attack on an Old Problem." Journal of Economic Perspectives 30, no. 4 (November 2016): 221–42. https://doi.org/10.1257/jep.30.4.221.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. "How Much Should We Trust Differences-In-Differences Estimates?" The Quarterly Journal of Economics 119, no. 1 (February 1, 2004): 249–75. https://doi.org/10.1162/003355304772839588.

Diversifying Economic Quality. "Undergraduate Participation Data." Accessed November 17, 2019. http://diversifyingecon.org/index.php/Undergraduate_participation_data

Dynan, Karen E., and Cecilia Elena Rouse. "The Underrepresentation of Women in Economics: A Study of Undergraduate Economics Students." The Journal of Economic Education 28, no. 4 (January 1, 1997): 350–68. https://doi.org/10.1080/00220489709597939.

Espey, Molly. Team-based learning in economics: A Pareto-improvement. In M. Sweet & L. K. Michaelson (Eds.), *Team-based learning in the social studies and humanities: Group work that works to generate critical thinking and engagement* (pp. 99–112). Sterling, Va: Stylus Publishing, 2012.

Espey, Molly. "Diversity, Effort, and Cooperation in Team-Based Learning." Journal of Economic Education 49, no. 1 (January 2018): 8–21.

Fairlie, Robert W., Florian Hoffmann, and Philip Oreopoulos. "A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom." American Economic Review 104, no. 8 (August 2014): 2567–91. https://doi.org/10.1257/aer.104.8.2567.

Milkman, Katherine L., Modupe Akinola, and Dolly Chugh. “What Happens before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations.” Journal of Applied Psychology 100, no. 6 (2015): 1678–1712. https://doi.org/10.1037/apl0000022.

Haak, David C., Janneke HilleRisLambers, Emile Pitre, and Scott Freeman. “Increased Structure and Active Learning Reduce the Achievement Gap in Introductory Biology.” Science 332, no. 6034 (2011): 1213–16.

Hettler, Paul L. “Student Demographics and the Impact of Team-Based Learning.” International Advances in Economic Research 21, no. 4 (November 2015): 413–22.

Imazeki, Jennifer. “Getting Students to Do Economics: An Introduction to Team-Based Learning.” International Advances in Economic Research; Edwardsville 21, no. 4 (November 2015): 399–412. http://dx.doi.org.stanford.idm.oclc.org/10.1007/s11294-015-9541-0.

McNeil, Jane, Borg, M., Kerrigan, M., Waller, S., Richter, U., Berkson, R., Tweddell, S. and McCarter, R. “Scaling up Active Collaborative Learning for Student Success,” October 2019. https://aclproject.org.uk/wp-content/uploads/2019/07/NTU-ABSS-Final-Report-revised-Oct-2019.pdf.

Michaelsen, Larry K., Arletta Bauman Knight, and L. Dee Fink, eds. Team-Based Learning: A Transformative Use of Small Groups in College Teaching. 1 edition. Sterling, Va: Stylus Publishing, 2004.

Milkman, Katherine L., Modupe Akinola, and Dolly Chugh. “What Happens before? A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations.” Journal of Applied Psychology 100, no. 6 (2015): 1678–1712. https://doi.org/10.1037/apl0000022.

Rask, Kevin, and Jill Tiefenthaler. “The Role of Grade Sensitivity in Explaining the Gender Imbalance in Undergraduate Economics.” Economics of Education Review 27, no. 6 (December 2008): 676–87. https://doi.org/10.1016/j.econedurev.2007.09.010.

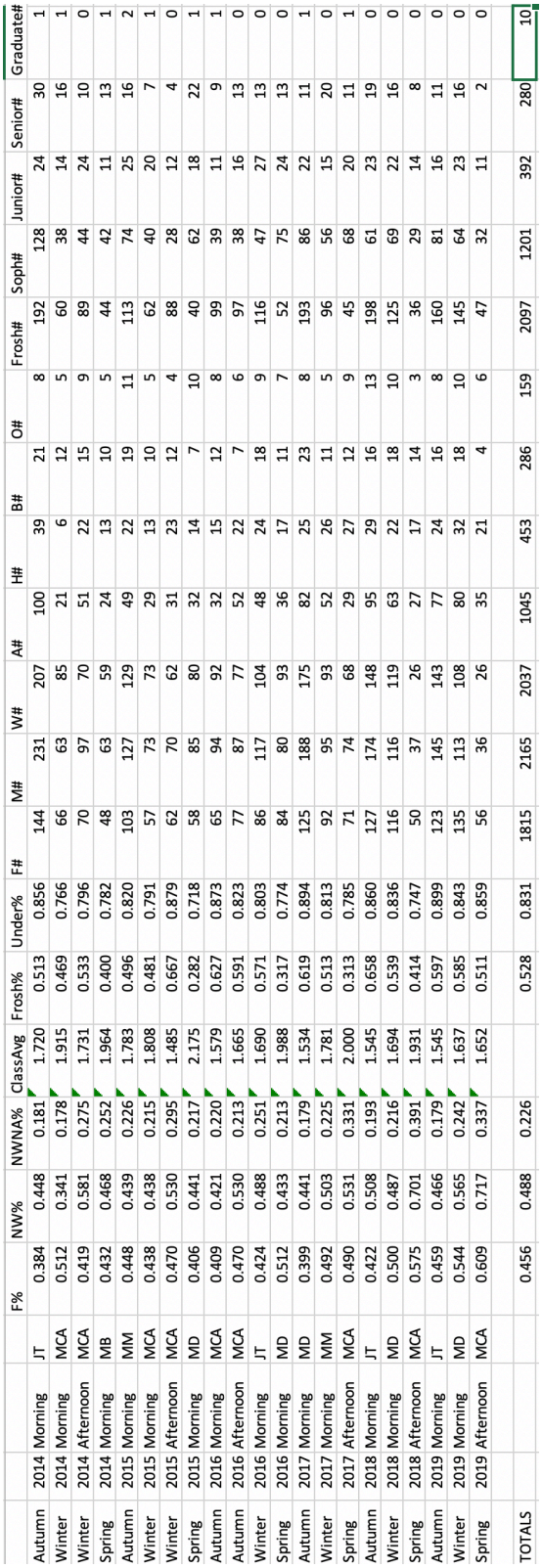
Rokicki, Slawa, Jessica Cohen, Günther Fink, Joshua A. Salomon, and Mary Beth Landrum. “Inference With Difference-in-Differences With a Small Number of Groups: A Review,

Simulation Study, and Empirical Application Using SHARE Data." Medical Care 56, no. 1 (2018): 97–105. https://doi.org/10.1097/MLR.0000000000000830.

Sweet, Michael, and Larry K. Michaelsen, eds. Team-Based Learning in the Social Sciences and Humanities: Group Work That Works to Generate Critical Thinking and Engagement. Sterling, Va: Stylus Publishing, 2012.

Wing, Coady, Kosali Simon, and Ricardo A. Bello-Gomez. "Designing Difference in Difference Studies: Best Practices for Public Health Policy Research." Annual Review of Public Health 39, no. 1 (2018): 453–69. https://doi.org/10.1146/annurev-publhealth-040617-013507.

Appendix A - Detailed data on each Econ 1 course offering



				F%	NW%	NWNA%	ClassAvg	Frosh%	Under%	F#	M#	W#	A#	H#	B#	O#	Frosh#	Soph#	Junior#	Senior#	Graduate#
Autumn	2014	Morning	JT	0.384	0.448	0.181	1.720	0.513	0.856	144	231	207	100	39	21	8	192	128	24	30	1
Winter	2014	Morning	MCA	0.512	0.341	0.178	1.915	0.469	0.766	66	63	85	21	6	12	5	60	38	14	16	1
Winter	2014	Afternoon	MCA	0.419	0.581	0.275	1.731	0.533	0.796	70	97	70	51	22	15	9	89	44	24	10	0
Spring	2014	Morning	MB	0.432	0.468	0.252	1.964	0.400	0.782	48	63	59	24	13	10	5	44	42	11	13	1
Autumn	2015	Morning	MM	0.448	0.439	0.226	1.783	0.496	0.820	103	127	129	49	22	19	11	113	74	25	16	2
Winter	2015	Morning	MCA	0.438	0.438	0.215	1.808	0.481	0.791	57	73	73	29	13	10	5	62	40	20	7	1
Winter	2015	Afternoon	MCA	0.470	0.530	0.295	1.485	0.667	0.879	62	70	62	31	23	12	4	88	28	12	4	0
Spring	2015	Morning	MD	0.406	0.441	0.217	2.175	0.282	0.718	58	85	80	32	14	7	10	40	62	18	22	1
Autumn	2016	Morning	MCA	0.409	0.421	0.220	1.579	0.627	0.873	65	94	92	32	15	12	8	99	39	11	9	1
Autumn	2016	Afternoon	MCA	0.470	0.530	0.213	1.665	0.591	0.823	77	87	77	52	22	7	6	97	38	16	13	0
Winter	2016	Morning	JT	0.424	0.488	0.251	1.690	0.571	0.803	86	117	104	48	24	18	9	116	47	27	13	0
Spring	2016	Morning	MD	0.512	0.433	0.213	1.988	0.317	0.774	84	80	93	36	17	11	7	52	75	24	13	0
Autumn	2017	Morning	MD	0.399	0.441	0.179	1.534	0.619	0.894	125	188	175	82	25	23	8	193	86	22	11	1
Winter	2017	Morning	MM	0.492	0.503	0.225	1.781	0.513	0.813	92	95	93	52	26	11	5	96	56	15	20	0
Spring	2017	Afternoon	MCA	0.490	0.531	0.331	2.000	0.313	0.785	71	74	68	29	27	12	9	45	68	20	11	1
Autumn	2018	Morning	JT	0.422	0.508	0.193	1.545	0.658	0.860	127	174	148	95	29	16	13	198	61	23	19	0
Winter	2018	Morning	MD	0.500	0.487	0.216	1.694	0.539	0.836	116	116	119	63	22	18	10	125	69	22	16	0
Spring	2018	Afternoon	MCA	0.575	0.701	0.391	1.931	0.414	0.747	50	37	26	27	17	14	3	36	29	14	8	0
Autumn	2019	Morning	JT	0.459	0.466	0.179	1.545	0.597	0.899	123	145	143	77	24	16	8	160	81	16	11	0
Winter	2019	Morning	MD	0.544	0.565	0.242	1.637	0.585	0.843	135	113	108	80	32	18	10	145	64	23	16	0
Spring	2019	Afternoon	MCA	0.609	0.717	0.337	1.652	0.511	0.859	56	36	26	35	21	4	6	47	32	11	2	0
TOTALS				0.456	0.488	0.226		0.528	0.831	1815	2165	2037	1045	453	286	159	2097	1201	392	280	10