# THE INFLUENCE OF INSTRUCTOR NATIVE LANGUAGE ON STUDENT LEARNING AND INSTRUCTOR RATINGS

**Kent T. Saunders**
*Le Moyne College*

## INTRODUCTION

The percentage of doctorates in economics awarded in the United States to U.S. citizens has declined from 67.3 percent in 1977 to 55.7 percent in 1986 to 42.9 percent in 1996 [Siegfried and Stock, 1999]. This clearly implies that the number of economics instructors whose native language is not English has increased significantly. In a study using survey data from the mid-1990s at a large public university, Borjas [2000] finds that foreign-born graduate teaching assistants negatively impact the final grade of undergraduate students. In a study with data collected from a large public university in 1984-85, Watts and Lynch [1989] find that students score significantly lower on a nationally normed exam when they have a graduate teaching assistant or graduate instructor whose native language is not English. In an unpublished follow-up study with data collected from the same school in the fall of 1989, Watts and Bosshardt [1992] find that, after the university implemented instructor screening measures, the negative effect that instructors whose native language is not English had on student learning no longer existed. The follow-up study also finds that students gave significantly lower instructor ratings to instructors whose native language is not English. Using the TUCE III data set [Saunders, 1994], Finegan and Siegfried [2000] find that instructors for whom English is a second language receive lower student ratings for teaching effectiveness, holding what students learn in the course constant.

The unpublished Watts and Bossardt findings [1992] and the Finegan and Siegfried [2000] findings indicate that instructors whose native language is not English receive lower ratings even when students learn as much from them as from instructors whose native language is English. This study uses two large databases to seek further answers to two key questions:

(1) Is student learning significantly different in introductory economics classes taught by graduate instructors whose native language is English compared to learning in classes taught by graduate instructors whose native language is not English?

(2) Do students give significantly different instructor ratings to graduate instructors whose native language is English compared to graduate instructors whose native language is not English?

---

**Kent T. Saunders:** LeMoyne College, 1419 Salt Springs Road, Syracuse, NY 13214. E-mail: Kent8484@AOL.com

The evidence discussed below indicates that the answer to the first question is "no" and the answer to the second question is "yes." These findings are potentially troubling in that the "vast majority of all institutions use some form of student evaluations of teaching" to evaluate teaching effectiveness [Becker and Watts, 1999, 346].

## DATA AND VARIABLES

Two large data sets were collected from introductory economics courses at Indiana University at Bloomington, from the fall semester of 1984 through the spring semester of 1990 (a six-year period including 12 semesters). The source of these data sets was first described in Saunders [1975]. They have since been used to study how gender influences student learning and instructor ratings [Saunders and Saunders, 1999]. Since the regular faculty who have taught the introductory courses at Indiana University do not vary much with regard to native language (virtually all are native English speakers), this study will use data collected from the students of different graduate student "Associate Instructors" who taught their own separate sections of introductory microeconomics or macroeconomics during the period under investigation. All of the Associate Instructors (AIs) in this study had completed a semester-long, three-credit seminar (E502, Teaching Undergraduate Economics) before they began teaching. They were given complete responsibility for their own sections of either micro or macro: they choose their own texts, prepared their own assignment sheets, and constructed their own exams, except for a common part of the final exam in each course. Normally the AIs teach for three years (6 semesters) and teach both micro and macro at some time in their careers. Prior to the fall of 1987, AIs were assigned two "small" sections with an enrollment limit of 40 students each or one "large" section with an enrollment limit of 80 students. Beginning in the fall of 1987 all AIs were assigned to teach only one section each semester with an enrollment limit of 50 students per section.

Although these data sets are becoming somewhat dated, they do allow for a controlled comparison between native and non-native English speaking instructors with roughly the same level of experience and the same teaching environment over an extended period of time. These data sets are better suited to assess the specific questions this study addressed than those in Borjas [2000] and Finegan and Siegfried [2000]. Borjas [2000] examines graduate teaching assistants who did not teach the entire course. Finegan and Seigfried [2000] use data that is also rather old and collected in only one year with a much more hetereogeneous group of instructors and teaching environments.

The micro data set has 59 different instructors of whom 37 are native English speakers and 22 are non-native English speakers. The macro data set has 38 different instructors of whom 25 are native English speakers and 13 are non-native English speakers. These data sets contain information about students, instructors, and sections. The student variables are STUGENDER, FINALT, EXPGRADE, SAT, FRESHMAN, and JUNIOR+. The Instructor variables are ENGLISH, INSTGENDER, OVRLRATE, COMMUNICATE, E502, EXPERIENCE, EXPERSQRD, and WARMUP. The section variable is TIME. Table 1 presents the definitions for these variables.

### TABLE 1
### Variable Definitions

| Variable | Definition |
|---|---|
| **Student** | |
| FINALT | Common portion final exam t-score with a mean of 50 and standard deviation of 10 |
| EXPGRADE | Student's expected course grade with F = 0 to A = 4 |
| SAT | Student's total SAT score divided by 10 |
| STUGENDER | Dummy variable that takes on a value of 1 if the student is female and 0 if the student is male |
| FRESHMAN | Dummy variable that takes on a value of 1 if the student is a freshman and 0 otherwise |
| JUNIOR+ | Dummy variable that takes on a value of 1 if the student is a junior or senior and 0 otherwise |
| **Instructor** | |
| ENGLISH | Dummy variable that takes on a value of 1 if the instructor is a native English speaker and 0 otherwise |
| OVRLRATE | Student rating of the instructor's overall teaching effectiveness with 4 being the highest and 0 the lowest |
| COMMUNICATE | Student rating of the instructor's speaking ability with 4 being the highest and 0 the lowest |
| E502 | Instructor's final grade in the E502 course with 9 being an A+ and 0 being a C- |
| EXPERIENCE | Instructor's number of semesters teaching experience for all courses |
| EXPERSQRD | EXPERIENCE squared |
| INSTGENDER | Dummy variable that takes on a value of 1 if the instructor is a female and 0 otherwise |
| WARMUP | Dummy variable that takes on a value of 1 if the instructor is teaching their second section of the same class that day and 0 otherwise |
| **Section** | |
| TIME | Dummy variable which takes on a value of 1 if the class meets before 9:00 am or after 4:00 pm |

Official data from the university registrar provided information for the STUGENDER, SAT, FRESHMAN and JUNIOR+ variables.

The unit of observation in this study is the individual student rather than the class average of individuals in the section. Since some instructors teach more than one class and class sizes vary, using un-weighted instructor class means could be misleading. Using individual students takes care of the weighting problem and also acknowledges the fact that it is individual students who provide the data used in the analyses.

We use registrar's data to determine if a student dropped the course after a one-week drop and add period. No information is available for drop and add activity during the one-week drop and add period. In the micro data set, 1.4 percent of students with non-native English speaking instructors dropped the course compared to 0.6 percent of students with native English speaking instructors, a statistically significant ($t = -4.21$) difference. In the macro data set 0.9 percent of students with non-native English speaking instructors dropped the course compared to 0.5 percent of students with native English speaking instructors, but this difference is not statisti-

## TABLE 2
## Comparison of Means

| | | | MICRO ENGLISH | | | NON-ENGLISH | | | Two-tailed |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Range | N | Mean | StDev | N | Mean | StDev | t-test | p-value |
| **Student Data** | | | | | | | | | |
| FINALT | (15-73) | 5881 | 50.24 | 9.79 | 4495 | 50.19 | 9.85 | 0.24 | 0.81 |
| EXPGRADE | (0-4) | 4034 | 2.85 | 0.82 | 3011 | 2.95 | 0.77 | −4.92 | 0.00[a] |
| SAT | (38-158) | 5432 | 98.84 | 15.90 | 4167 | 99.47 | 15.37 | −1.96 | 0.05[b] |
| STUGENDER | (0-1) | 5947 | 0.46 | 0.50 | 4571 | 0.43 | 0.50 | 2.67 | 0.01[a] |
| FRESHMAN | (0-1) | 5945 | 0.48 | 0.50 | 4571 | 0.61 | 0.49 | −13.14 | 0.00[a] |
| JUNIOR+ | (0-1) | 5945 | 0.11 | 0.31 | 4571 | 0.08 | 0.27 | 4.82 | 0.00[a] |
| **Instructor Data** | | | | | | | | | |
| OVRLRATE | (0-4) | 4020 | 2.89 | 0.99 | 2997 | 2.82 | 1.00 | 2.70 | 0.01[a] |
| COMMUNICATE | (0-4) | 4041 | 3.24 | 0.94 | 3012 | 2.40 | 1.17 | 32.77 | 0.00[a] |
| E502 | (4-8) | 5522 | 7.54 | 0.75 | 4571 | 7.38 | 0.63 | 11.45 | 0.00[a] |
| EXPERIENCE | (1-9) | 5947 | 3.43 | 2.32 | 4571 | 3.49 | 2.45 | −1.30 | 0.19 |
| EXPERSQRD | (1-81) | 5947 | 17.16 | 19.95 | 4571 | 18.18 | 22.81 | −2.40 | 0.02[b] |
| INSTGENDER | (0-1) | 5947 | 0.34 | 0.47 | 4571 | 0.38 | 0.49 | −3.99 | 0.00[b] |
| WARMUP | (0-1) | 5947 | 0.16 | 0.37 | 4571 | 0.08 | 0.28 | 12.23 | 0.00[a] |
| **Section Data** | | | | | | | | | |
| TIME | (0-1) | 5947 | 0.11 | 0.31 | 4571 | 0.10 | 0.30 | 1.88 | 0.06 |

| | | | MACRO ENGLISH | | | NON-ENGLISH | | | Two-tailed |
|---|---|---|---|---|---|---|---|---|---|
| Variable | Range | N | Mean | StDev | N | Mean | StDev | t-test | p-value |
| **Student Data** | | | | | | | | | |
| FINALT | (4-79) | 3234 | 49.85 | 9.84 | 1782 | 50.36 | 9.62 | −1.78 | 0.08 |
| EXPGRADE | (0-4) | 2557 | 2.86 | 0.79 | 1374 | 2.91 | 0.77 | −1.93 | 0.05 |
| SAT | (43-151) | 3381 | 99.08 | 15.50 | 1908 | 99.67 | 15.77 | −1.33 | 0.18 |
| STUGENDER | (0-1) | 3707 | 0.43 | 0.49 | 2086 | 0.39 | 0.49 | 2.61 | 0.01[a] |
| FRESHMAN | (0-1) | 3706 | 0.11 | 0.31 | 2086 | 0.11 | 0.31 | −0.35 | 0.72 |
| JUNIOR+ | (0-1) | 3706 | 0.20 | 0.40 | 2086 | 0.20 | 0.40 | −0.17 | 0.86 |
| **Instructor Data** | | | | | | | | | |
| OVRLRATE | (0-4) | 2557 | 2.84 | 0.99 | 1370 | 2.57 | 1.07 | 7.66 | 0.00[a] |
| COMMUNICATE | (0-4) | 2564 | 3.29 | 0.81 | 1376 | 2.23 | 1.21 | 29.32 | 0.00[a] |
| E502 | (6-8) | 3131 | 7.40 | 0.66 | 1784 | 7.08 | 0.47 | 19.82 | 0.00[a] |
| EXPERIENCE | (1-9) | 3707 | 4.31 | 2.20 | 2086 | 5.64 | 2.46 | −20.48 | 0.00[a] |
| EXPERSQRD | (1-81) | 3707 | 23.45 | 22.44 | 2086 | 37.88 | 27.16 | −20.63 | 0.00[a] |
| INSTGENDER | (0-1) | 3707 | 0.09 | 0.29 | 2086 | 0.12 | 0.33 | −3.24 | 0.00[a] |
| WARMUP | (0-1) | 3707 | 0.23 | 0.42 | 2086 | 0.11 | 0.31 | 12.56 | 0.00[a] |
| **Section Data** | | | | | | | | | |
| TIME | (0-1) | 3707 | 0.24 | 0.42 | 2086 | 0.15 | 0.36 | 8.41 | 0.00[a] |

a. Significant at the 1 percent level.
b. Significicnat at 5 percent.

cally significant. It should be noted that the decision to drop a course is also influenced by other factors; thus, these results may not be conclusive.

## UNIVARIATE ANALYSIS

Table 2 compares the means of the student, instructor and section variables for native English speaking instructors with those of instructors whose native language is not English (non-English) for both micro and macro courses. As measured by student performance on the portion of the final common to all sections (FINALT), student learning did not differ significantly in either the micro or the macro courses.

Students rate the overall teaching effectiveness (OVRLRATE) of non-native English speaking instructors significantly lower than native English speaking instructors in both micro and macro courses. In the micro course, students rate non-native English speaking instructors 0.07 points lower on a 4 point scale. In the macro course, students rate non-native English speaking instructors 0.27 points lower on a 4 point scale.

Students in both micro and macro courses rate the ability of instructors whose native language is not English to communicate (COMMUNICATE) significantly lower. Students in micro rate non-native English speaking instructors 0.84 points lower on a 4 point scale. Students in macro rate non-native English speaking 1.06 points lower on a 4 point scale.

Since OVRLRATE and COMMUNICATE are both measured on a discrete ordinal scale, it is probably more appropriate to conduct non-parametric tests. We reach the same conclusions with respect to OVRLRATE and COMMUNICATE when using a Mann-Whitney test rather than a t-test.

Table 2 also identifies several significant differences among student, instructor, and section data between native English and non-native English speaking instructors. Thus, multivariate analysis must be conducted to control for student, instructor, and section variables other than instructor's native language.

## MULTIVARIATE ANALYSIS

Equation (1) represents the ordinary least squares regression model used when FINALT is the dependent variable. It should be noted that students do not take a pretest in these classes, and the posttest minus pretest regression model has been criticized for showing bias [Becker, 1983; Becker, Greene and Rosen, 1990].

(1) FINALT = f(constant, student variables, instructor variables, section variable)

OVRLRATE and COMMUNICATE are both discrete ordinal dependent variables. Equation (2) and Equation (3) represent the ordered probit regression technique that was used when OVRLRATE and COMMUNICATE are the dependent variables.

(2) OVRLRATE = f(constant, student variables, instructor variables, section variable)

### TABLE 3
### Micro: Multivariate Analysis

| Dependent Variable: | FINALT (OLS) | | OVRLRATE (Ordered Probit) | | COMMUNICATE (Ordered Probit) | |
|---|---|---|---|---|---|---|
| | Coef. | t-test | Coef. | t-test | Coef. | t-test |
| Constant | 18.56 | 15.02[a] | 0.75 | 3.88[a] | −0.38 | −1.86 |
| **Student Variables** | | | | | | |
| SAT | 0.27 | 45.95[a] | −0.10 | −9.82[a] | −0.01 | −7.29[a] |
| STUGENDER | −1.44 | −7.79[a] | 0.02 | 0.65 | 0.12 | 4.13[a] |
| FRESHMAN | −0.35 | −1.70 | 0.08 | 2.56[b] | 0.04 | 1.28 |
| JUNIOR+ | 0.86 | 2.51[b] | 0.15 | 2.82[a] | 0.24 | 4.39[a] |
| EXPGRADE | — | — | 0.46 | 23.86[a] | 0.31 | 15.94[a] |
| **Instructor Variables** | | | | | | |
| ENGLISH | 0.18 | 0.93 | 0.05 | 1.73 | 0.89 | 30.01[a] |
| E502 | 0.47 | 3.33[a] | 0.07 | 3.42[a] | 0.18 | 8.03[a] |
| EXPERIENCE | 0.72 | 4.44[a] | 0.24 | 9.29[a] | 0.08 | 3.10[a] |
| EXPERSQRD | −0.04 | −2.20[b] | −0.01 | −4.74[a] | 0.00 | 1.48 |
| INSTGENDER | −0.08 | −0.38 | −0.40 | −12.54[a] | −0.41 | −12.80[a] |
| WARMUP | 0.74 | 2.62[a] | 0.22 | 5.27[a] | 0.14 | 3.38[a] |
| **Section Variable** | | | | | | |
| TIME | −0.48 | −1.61 | -0.11 | −2.33[b] | −0.14 | −2.99[a] |
| N | 9080 | | 6104 | | 6132 | |
| R² | 0.21 | | — | | — | |
| Adjusted R² | 0.21 | | — | | — | |
| Percent Concordant | — | | 42% | | 45% | |

a. Significant at the 1 percent level
b. Significant at 5 percent.

(3) *COMMUNICATE* =
 f(constant, student variables, instructor variables, section variable)

Table 3 displays the regression results for the micro course. In terms of overall model fit, approximately 21 percent of the variation in *FINALT* is explained by variation in the independent variables. For *OVRLRATE* and *COMMUNICATE* 42 and 45 percent, respectively, of the observations are concordant with the model estimates. Instructor native language does not have a significant effect on student learning (*FINALT*) or the overall rating of teaching effectiveness (*OVRLRATE*). Instructors whose native language is not English are rated significantly lower than instructors whose native language is English in terms of ability to communicate (*COMMUNICATE*).

Table 4 displays the regression results for the macro data set. In terms of overall model fit, approximately 15 percent of the variation in *FINALT* is explained by variation in the independent variables. For *OVRLRATE* and *COMMUNICATE* 40 and 44

### TABLE 4
### Macro: Multivariate Analysis

| Dependent Variable: | FINALT (OLS) | | OVRLRATE (Ordered Probit) | | COMMUNICATE (Ordered Probit) | |
|---|---|---|---|---|---|---|
| | Coef. | t-test | Coef. | t-test | Coef. | t-test |
| Constant | 22.24 | 10.42[a] | −0.45 | −1.46 | −0.41 | −1.34 |
| **Student Variables** | | | | | | |
| SAT | 0.22 | 22.97[a] | −0.01 | −6.43[a] | −0.01 | −3.87[a] |
| STUGENDER | −2.17 | −7.23[a] | 0.06 | 1.38 | 0.19 | 4.68[a] |
| FRESHMAN | −1.05 | −2.03[b] | 0.03 | 0.51 | −0.05 | −0.77 |
| JUNIOR+ | 0.00 | −0.01 | 0.11 | 2.19[b] | 0.15 | 3.01[a] |
| EXPGRADE | — | — | 0.41 | 15.06[a] | 0.24 | 8.69[a] |
| **Instructor Variables** | | | | | | |
| ENGLISH | 0.01 | 0.24 | 0.11 | 2.27[b] | 1.04 | 20.25[a] |
| E502 | 0.92 | 3.48[a] | 0.22 | 6.08[a] | 0.19 | 5.21[a] |
| EXPERIENCE | 0.19 | 0.60 | 0.23 | 5.40[a] | 0.10 | 2.23[b] |
| EXPERSQRD | −0.01 | −0.44 | −0.03 | −6.03[a] | −0.01 | −2.72[a] |
| INSTGENDER | −1.08 | −2.34[b] | −0.57 | −8.30[a] | −0.39 | −5.55[a] |
| WARMUP | −1.08 | −2.66[a] | 0.03 | 0.58 | 0.07 | 1.27 |
| **Section Variable** | | | | | | |
| TIME | −0.42 | −1.09 | 0.01 | 0.20 | −0.01 | 0.00 |
| N | 3858 | | 3033 | | 3039 | |
| R² | 0.15 | | — | | — | |
| Adjusted R² | 0.14 | | — | | — | |
| Percent Concordant | — | | 40% | | 44% | |

a. Significant at the 1 percent level.
b. Significant at 5 percent.

percent, respectively, of the observations are concordant with the model estimates. Instructor native language does not significantly affect student learning (*FINALT*). Students rate instructors whose native language is not English significantly lower than instructors whose native language is English, in terms of both their overall rating of teaching effectiveness (*OVRLRATE*) and their ability to communicate (*COMMUNICATE*).

Some of the other independent variables in the multivariate regressions are worthy of comment. [Saunders and Saunders, 1999] studies the effect of gender on learning and instructor ratings. An interesting finding consistent for both micro and macro courses is the significant positive effect that teaching experience has on overall ratings and ratings of communication ability. Teaching experience also positively affects learning in both data sets, a significant result for both courses. Lastly, performance in the E502 course positively and significantly affects learning, overall ratings and ratings of communication ability in both the micro and macro courses.

**TABLE 5**
**Summary of Results for Non-Native English Speaking Instructors**

|  |  | Micro | Macro |
|---|---|---|---|
| Student Learning: | Univariate | Same | Same |
|  | Multivariate | Same | Same |
| Overall Ratings: | Univariate | Less | Less |
|  | Multivariate | Same | Less |
| Ability to Communicate: | Univariate | Less | Less |
|  | Multivariate | Less | Less |

Table 5 summarizes the statistical results from Tables 2 through 4. The results of both univariate and multivariate analysis indicate that the instructor's native language does not significantly affect student learning. Instructors whose native language is not English, however, receive significantly lower overall instructor ratings compared to instructors whose native language is English. This result is statistically significant for the micro and macro courses in a univariate analysis and for the macro course in a multivariate analysis. In addition, instructors whose native language is not English receive significantly lower ratings of ability to communicate compared to instructors whose native language is English. This result is significant in both courses in both a univariate and a multivariate analysis.

## CONCLUSION

The evidence in this study indicates that the student ratings for overall teaching effectiveness of instructors whose native language is not English are usually lower than would be warranted if instructor performance were evaluated on the basis of student learning. It is recommended that economics departments consider using a common final exam or a comparable instrument as an additional tool to evaluate graduate instructors. Since many schools do not have measures of learning comparable to the ones used in this study, one wonders what non-native speakers might do to improve their ratings beyond making efforts to improve their English skills and ability to communicate. Nelson suggests that for non-native instructors "the use of personal examples from their cultures to illustrate concepts in their lectures works to decrease student uncertainty, improve student attitude, and increase student recall" [1991, 433]. Further, incorporating active learning activities in their courses with less reliance on the straight lecture method might be worth trying. Finegan and Siegfried [2000] find that roughly two thirds of the gap in overall instructor ratings between native and non-native speakers of English disappears after controlling for differences in how instructors teach their courses and evaluate student achievement.

A path for future research in this area might consider a comparable study conducted for graduate instructors teaching in New York City, Miami, Texas, or Southern California, where daily interaction with non-native English speakers is more common.

## REFERENCES

**Becker, W.** Economic Education Research: Part III, Statistical Estimation Methods. *Journal of Economic Education,* Summer 1983, 4-15.
**Becker, W., Greene W., and Rosen, S.** Research on High School Economic Education. *Journal of Economic Education,* Summer 1990, 231-45.
**Becker, W. and Watts, M.** How Departments of Economics Evaluate Teaching. *American Economic Review, Papers and Proceedings,* May 1999, 344-9.
**Borjas, G. J.** Foreign-Born Teaching Assistants and the Academic Performance of Undergraduates. *American Economic Review, Papers and Proceedings,* May 2000, 355-9.
**Finegan, T. and Siegfried, J.** Are Student Ratings of Teaching Effectiveness Influenced by Instructors' English Language Proficiency? *The American Economist,* Fall 2000, 17-29.
**Nelson, G.** Effective Teaching Behavior for International Teaching Assistants. In *Preparing the Professorate to Teach: Selected Readings in TA Training,* edited by J.D. Nyquist, R. D. Abbott, D. H. Wulff, and J. Sprague. Iowa: Kendall/Hunt, 1991.
**Saunders, K. and Saunders, P.** The Influence of Gender on Learning and Instructor Ratings. *Atlantic Economic Journal,* December 1999, 460-73.
**Saunders, P.** Experimental Course Development in Introductory Economics at Indiana University. *The Journal of Economic Education,* Fall 1975, Special Issue No. 4.
_____. *The TUCE III Data Set: Background Information and File Codes.* New York: National Council on Economic Education, 1994.
**Siegfried, J. and Stock, W.** The Labor Market for New Ph.D. Economists. *Journal of Economic Perspectives* 13, Summer 1999, 115-34.
**Watts, M. and Bosshardt, W.** International TAs and Student Time Allocations: Impacts on Learning, Grades, and Student Course and Instructor Evaluations. Unpublished paper presented at the January meetings of the Allied Social Science Association, 1992.
**Watts, M., and Lynch, G.** The Principles Courses Revisited. *American Economic Review, Papers and Proceedings,* May 1989, 236-41.