Lost in Translation? Teacher Training and Outcomes in High School Economics Classes

Robert G. Valletta
K. Jody Hoff
Jane S. Lopus

DISCUSSION PAPER SERIES

IZA

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Lost in Translation?
Teacher Training and Outcomes in
High School Economics Classes

Robert G. Valletta
*Federal Reserve Bank of San Francisco
and IZA*

K. Jody Hoff
Federal Reserve Bank of San Francisco

Jane S. Lopus
California State University, East Bay

ABSTRACT

Lost in Translation? Teacher Training and Outcomes in High School Economics Classes[*]

Using data on 24 teachers and 982 students from a 2006 survey of California high school economics classes, we assess the effects of student and teacher characteristics on student achievement. We estimate value-added models of outcomes on multiple choice and essay exams, with matched classroom pairs for each teacher enabling random effects and fixed-effects estimation. Students' own and peer GPAs and their attitudes towards economics have the largest effects on value-added scores. We also find a substantial impact of specialized teacher experience and college-level coursework in economics, although the effects of the latter are positive for the multiple choice test and negative for the essay test.

JEL Classification: A21, I21

Keywords: high school economics, teacher training

Corresponding author:

Robert G. Valletta
Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
USA
E-mail: rob.valletta@sf.frb.org

**Lost in Translation? Teacher Training and
Outcomes in High School Economics Classes**

I. Introduction

In this paper, we assess the contributions of teacher and student characteristics to students' success in high school economics classes. Much of the voluminous literature on educational outcomes focuses on students in primary school. A key finding from this literature is that the quantitative contribution of teacher quality to student outcomes is large, but the relationship between teacher quality and measurable characteristics such as training and experience is limited and ambiguous (e.g., Rockoff 2004; Rivkin, Hanushek, and Kain 2005; Aaronson, Barrow, and Sander 2007). By contrast with this literature, we focus on subject matter education at the high school level, for which teachers' specialized educational background and experience may play a larger role than it does in primary school settings.

The data are from a special survey of California high school economics teachers conducted by the Federal Reserve Bank of San Francisco (FRBSF) and outside researchers in 2006, originally aimed at assessing the effectiveness of a video curriculum program developed by FRBSF. The survey produced value-added outcome data for nearly 1000 students in 48 matched-pair classes taught by 24 teachers. Our experimental design enables us to explicitly account for observed and unobserved teacher effects in addition to observable student and class characteristics. We obtained pre- and post-test outcomes for two testing modes: multiple choice questions and an essay question. The essay mode represents an additional novel element of our

1

analyses, given the economic education literature's typical reliance on standardized multiple choice questions.[1]

After discussing relevant background literature and our data, we describe our value-added methodological framework and results. We find that student characteristics including own GPA and peer GPA have the largest effects on student achievement. The estimated effects of teacher characteristics such as experience teaching economics and formal education in economics also raise student achievement by statistically significant amounts that are nearly as large as the effects of student characteristics. However, the impact of teachers' formal education in economics varies across the different testing modes, with a college emphasis in economics leading to higher multiple choice scores but lower essay scores; we dub the latter the "lost in translation" effect. By contrast, advanced degrees in fields other than economics (mainly education) enhance student outcomes for both testing modes. These results suggest that content-specific teacher training may have substantial value for high school subject matter, but further research is required to identify the specific settings in which this value can be realized.

II. LITERATURE REVIEW

II.A. Education literature

Teacher quality has been identified in a variety of studies as a significant factor in the determination of student achievement (see e.g. Hanushek and Rivkin 2010; Kane et al. 2010) However, identifying and measuring teacher quality has proved to be challenging, because observable characteristics such as attainment of advanced degrees, years of teaching experience, certifications, and ongoing professional development have not consistently explained variation in

---

[1] Exceptions to this focus on multiple choice questions include an earlier paper that used the same dataset as ours but emphasized other determinants of outcomes (Lopus and Hoff 2009) and a much earlier paper focusing on gender differences at the university level (Ferber, Birnbaum, and Green 1983).

teacher-specific contributions to student achievement (Rockoff 2004; Rivkin et al. 2005; Aaronson et al. 2007). Conflicting results on factors affecting teacher quality are often attributed to differences in methodology and data (Kane et al. 2010; Rothstein 2010).

As discussed by Harris and Sass (2007), findings on the impact of observable teacher characteristics on student achievement are mixed, with some studies showing positive effects for elementary and middle school math but not for reading. At the high school level, a number of studies have found positive effects for teacher training in content-specific areas but not for formal education, including the attainment of advanced degrees. For example, Aaronson et al. (2007) and Betts, Zau, and Rice (2003) investigated the effects of teachers' college major and found no systematic impact on student achievement in high school. Surprisingly, Harris and Sass (2007) found that teachers with majors in math or math education are *less* productive in the teaching of high school math than teachers who lack such background.[2]

In contrast to the ambiguous effects of formal education, additional teaching experience generally has been found to enhance teacher quality at both the elementary and high school levels (Rivkin et.al. 2005; Clotfelter et al. 2007; Harris and Sass 2007; Rice 2010). However, the impact of teaching experience begins to taper off after the first year and by the fourth or fifth year does not appreciably affect teacher quality (Clotfelter et al. 2007; Boyd et al. 2008; Rice 2010). Harris and Sass (2007) utilized panel data matching students and teachers to specific classrooms to estimate the effects of teacher education and training, distinguishing between specific types of undergraduate coursework and also between different types of professional development training. Their results indicated generally positive but mixed effects for years of

---

[2] Clotfelter, Ladd, and Vigdor (2007) also uncovered evidence that teachers' advanced degrees can be associated with lower student outcomes. In particular, using a panel of 9th and 10th grade students from North Carolina, they found a large negative effect on end-of-course test scores for teachers holding a Ph.D. Given the small number of such teachers in their sample, this may reflect unobserved characteristics of those teachers rather than a general pattern.

experience but no significant effects for obtaining an advanced degree. These findings corroborate results from a number of other studies (Rice 2010).

Some studies also find large effects of teacher credentials such as content-specific licensing or training requirements. For example, using end-of-course test scores for a large sample of high school students in North Carolina, Clotfelter et al. (2007) found that the effects of teacher credentials are larger than the effects of student characteristics. Similarly, Harris and Sass (2007) found that professional development in the form of content-specific training raised student achievement, using administrative data for the universe of third through tenth grade students in Florida. In another recent study, Subedi , Swan, and Hynes (2009) examined middle school gain scores in mathematics from a single school district and found significant positive effects for content certification in mathematics and teaching experience.

II.B. *Economic education literature*

A parallel literature within economic education has focused on how teacher characteristics and the other factors discussed in the preceding section affect achievement in high school economics classes. Watts (2005) summarized much of the existing empirical research on pre-college economic education, focusing on studies published since 1990. His major conclusions, as reiterated in Watts and Walstad (2011), identified three key factors that influence student learning of economics in high schools:

1. Students who take a separate class in economics and therefore spend more time on economics learn more than students who study economics infused into other subjects.

2. Teacher knowledge of economics and teacher training in economics enhance student learning.

3. Instructional materials that reflect accurate economics content and incorporate interesting instructional methods enhance student learning. However, no consistent evidence favors specific materials or teaching methods.

4

To determine why some teachers are more effective than others, Bosshardt and Watts (1990) investigated teacher effects on student learning in high school economics classes using fixed- and random-effects models. The teacher characteristics they examined included college credits in economics, non-credit workshops in economics, years of teaching experience, and the extent of teachers' past instruction in economics. They found that the most effective teachers were those who had completed more courses in economics. Other significant factors included a proxy for student IQ and school characteristics.

Walstad (2001) also found that teachers who take more courses in economics are more effective in teaching economics and identified five to six courses as a critical threshold level. Overall, teacher characteristics other than coursework in economics had inconsistent or insignificant effects on student learning in various studies. These characteristics include completion of non-credit workshops, years of teaching experience, years since the last economics course was taken, and percentage of teaching load that is economics (Walstad 1992). It is important to note that these findings pertain to objective test modes such as multiple choice questions.

With respect to student effects, measures of student aptitude or intelligence are consistently found to be positively correlated with student achievement in economics (Watts 2005). Related to this, prior knowledge or ability captured by a pretest score is typically the single most important variable for explaining student achievement. Students in higher level courses such as Advanced Placement (AP) courses outperform students in non-AP courses (Butters and Asarta 2011) Other student characteristics related to economics knowledge and learning in high schools include student gender and race or ethnicity, with male students often

outperforming female students and whites outperforming other races and ethnic groups.[3]

Teachers' attitudes about economics are found to affect student attitudes and student attitudes are found to affect student learning, although the direction of causality between student attitudes and student learning is not clear (Watts 2005.)

III. DATA

The data used in this paper were collected as part of a project to assess the effectiveness of a video curriculum program developed by FRBSF for use in high school economics classes (*Open and Operating: the Federal Reserve Responds to September 11*).[4] The accompanying guide relates the concepts covered in the curriculum to the Voluntary National Content Standards in Economics (NCEE 2000). The project was administered in California, where a semester course in economics has been required for high school graduation since 1989. Economics teachers at all public high schools in California (approximately 1000) were invited to participate in the assessment project, conducted in fall 2006. In order to introduce strong controls into the data, the only teachers included in the final sample were those who were teaching two economics classes for similar groups of students (Advanced Placement, honors or college prep, non-college bound, or mixed). For each teacher in the final sample, one class served as the experimental class using the *Open and Operating* ("O&O") curriculum on monetary policy, and the other class served as the control class, with the material on monetary policy taught in the teachers' traditional manner. Teachers were asked to administer student questionnaires and to complete a

---

[3] The gender effect is not consistent in past research (see Johnson, Robson, and Taengnoi 2011) , and it is likely that racial gaps reflect school district and other student background characteristics.
[4] The title refers to the Federal Reserve's press release of September 11, 2001 following the terrorist attacks: "The Federal Reserve is open and operating. The discount window is available to meet liquidity needs." The 16-minute video describes the background and functions of the Federal Reserve System, monetary policy, how the central bank's responsibilities have evolved over time, and how the Fed responded to the September 11 crisis. See Lopus and Hoff (2009) for additional details.

teacher questionnaire. Evaluation of student outcomes was based on pre- and post-tests for a set of 20 multiple choice questions and an essay question, as described in more detail below.

Sixty-two teachers responded that they were scheduled to teach two similar economics classes during fall 2006 and that they were willing to participate in the study. Materials were sent in early September and teachers were instructed to randomly assign one class to be the experimental class and one class to be the control class. Forty-three teachers returned some materials and 24 returned the complete sets of materials used in this study.[5] Among the 48 classes taught by these 24 teachers, 1290 students returned some information, with 982 returning complete information used in the regression analysis of multiple choice scores in this paper and 963 for the essay scores.

The pre- and post-tests administered to students were developed for use in this study, since no valid, normed, and reliable instrument such as the *Test of Economic Literacy (TEL)* (Walstad and Rebeck 2001) exists relating to the concepts covered in the O&O curriculum. The 20 multiple-choice questions and correct answers were taken from existing instruments such as the *TEL* where appropriate. The essay question asks students to write one or two paragraphs (as if for a newspaper) about how the Federal Reserve System could respond to a situation such as high inflation, unemployment, a banking panic or other crisis. Inclusion of this question represents a novel form of testing and instructional assessment in the present study relative to most past research. To ensure grading objectivity and comparability across classes, grading for this question was performed by a panel of 6 experienced high school economics teachers recruited from a pre-existing teacher database compiled by the FRBSF economic education group. None of the teachers included in the grading panel were otherwise involved with the

---

[5] Teaching assignments for some of the original 62 teachers changed so that they were no longer teaching two similar classes and could no longer participate in the study. Some teachers did not require their students to complete all assessment activities or otherwise returned data that were not usable.

study. A grading "session" was conducted at the FRBSF offices in San Francisco where teachers were briefed on the study objectives and given examples of graded responses to review and discuss prior to scoring the study essay question. The panel utilized a scoring rubric to evaluate the essay responses, with grades ranging from 0 (lowest) to 3 (highest).[6] Each essay was read by two teachers. If their scores differed, a third teacher also graded the essay.

Table 1 reports variable definitions and descriptive statistics. Post-test multiple choice and essay scores serve as dependent variables in the regression analyses, controlling for pre-test scores, as described in the next section. The mean values indicate that scores were raised substantially by the teaching curriculum, particularly for the multiple choice questions; the post-test average of about 13 represents nearly a 60 percent gain relative to the pre-test average of about 8. The improvement in essay scores is also substantial but not especially impressive in regard to the final level, with an increase essentially from 0 to 1 on average. Some students achieved the maximum score on the multiple choice post-test, and some received the maximum on the essay pre-test as well as the post-test.

Student characteristics, class characteristics, and teacher characteristics will be used as control variables in our analyses, to capture the influences on student achievement discussed in Section II. The descriptive statistics in Table 1 reveal a diverse student body in regard to racial and ethnic composition, prior academic performance as reflected in own and peer GPA, and parents' educational attainment. Their attitude towards economics mostly ranges from indifference to open dislike. Class characteristics also reveal substantial diversity along key dimensions. About 5 hours of class time was devoted to the monetary policy curriculum on

---

[6] The specific guidance for scoring the essay exams was as follows: 0 = No response or no knowledge of what the Fed is or does; 1 = May include incorrect information but some knowledge of what the Fed is or does; 2 = Mostly correct information and basic knowledge about the role of the Fed; 3 = Correctly describes role of Fed and provides relevant details.

average, with the full range extending from a minimum of 2 hours to a maximum of 12.[7] Most

of the classes fell in the category of "mixed" in terms of students' academic aspirations and

likelihood of college attendance.[8]

Teacher characteristics are of particular interest in this study. Most teachers in the

sample have taught economics for at least 10 years and have been teaching for nearly 20 years

overall (the medians for these variables are not shown but are very close to the means listed in

Table 1). One quarter of the teachers (6) have taught economics for their entire careers, ranging

from 5 to 20 years in length. Only a few teachers have taught for 5 years or less, which

precludes separate identification of early career learning effects in our regression analyses. One-

third of the teachers possess a college undergraduate major or minor degree in economics, while

two-thirds hold an advanced degree (typically a Master's in Education); about half of those who

were undergraduate economics majors hold an advanced degree, so the overlap between these

two groups is small enough for estimation of their independent impacts on student achievement.[9]

IV. REGRESSION FRAMEWORK AND RESULTS

IV. A. *Regression specification*

Our analysis relies on the well-established value-added approach for estimating the

contributions of student, teacher, and classroom characteristics to educational outcomes (see e.g.

Rivkin et al. 2005; Kane et al. 2010; Rothstein 2010; Chetty, Friedman, and Rockoff 2011). Pre-

---

[7] Teachers were asked to allocate equal amounts of time to the topics in both classes. All did except for two, who reported spending slightly more time on the material in the experimental class. These differences are small compared with the range of time spent across teachers. In the regressions, we control for actual time spent in each class.

[8] Because complete data records were not obtained for all students in each class, we are unable to form and use an accurate class-size variable in our empirical analyses.

[9] Among the advanced degree holders, only one came from a traditional Master's program in Economics (actually one course short of a Master's in Economics); this is too small for accurate inference, and we therefore do not distinguish this individual from advanced degree holders more generally.

tests and post-tests were administered to participating students in our sample of California high
school economics classes before and after they completed the unit on monetary policy and
central banking, so our experimental design fits well within the value-added framework.
Moreover, our matched experimental sample and ability to model individual teacher effects helps
us to overcome some common concerns about the influence of unobservables in the estimation of
education production functions.

We begin with a value-added equation of the following form:

$$A_{ijk} = \alpha_0 + \alpha_1 A_{ijk}^0 + \beta_1 C_j + \beta_2 S_{ijk} + \mu_k + \varepsilon_{ijk} \qquad (1)$$

This equation specifies that the achievement outcome A (multiple choice or essay test score) of
student i in classroom j taught by teacher k depends on the student's pre-test score A^0 (which
represents individual ability and cumulative educational inputs prior to the experiment), plus
vectors of classroom instructional characteristics C_j, individual student characteristics S_{ijk}, and
teacher effects μ_k.[10] The α and β terms are coefficients to be estimated, and ε_{ijk} is an error term
that has zero mean conditional on the right-hand side variables. The intent of Equation (1) is not
to pin down the sources of the teacher-specific contribution to student achievement, but instead
to model these effects as unobserved intercept shifts (fixed effects) or as a teacher-specific
component of variance in the error term (random effects). After estimating this equation using
the fixed-effects and random-effects estimators, we implemented standard tests of the alternative
econometric models. Acceptance of the null hypothesis of random effects implies that the

---

[10] A common alternative approach estimates determinants of the "gain score" by moving the pre-test score
to the left-hand side and using the score change as the dependent variable. This specification is
equivalent to restricting the coefficient on the pre-test score in (1) to be equal to 1. This restriction is
strongly rejected in our empirical results reported in Tables 2 and 3.

unobserved teacher effects are uncorrelated with the other variables (classroom and student) in

the model, implying in turn that we can obtain unbiased estimates of the coefficients on A^0, C,

and S along with the coefficients on a set of explicit teacher variables, or T_k in equation (2)

below:[11]

$$A_{ijk}^s = \lambda_0 + \lambda_1 A_{ijk}^r + \gamma_1 C_j + \gamma_2 S_{ijk} + \gamma_3 T_k + \eta_{ijk} \qquad (2)$$

We estimate equations (1) and (2) and apply specification tests for the multiple choice

outcomes. The multiple choice scores range from 0 to 20, and their distribution has nearly

equivalent mean and variance. We investigated use of a Poisson regression model, which often

has attractive properties for estimating models based on count data such as our multiple choice

test score. Our specification checks indicated that the Poisson and linear models yield similar

point estimates, but the linear model generates more precise estimates in our specific setting. We

therefore use the linear model for the analysis of the multiple choice outcomes.[12] The essay

question scores range from 0 to 3, lending themselves naturally to an ordered response model;

we use an ordered logit model for the results reported in the next sub-section.

IV.B. Regression Results

Table 2 lists regression results for four different specifications of the multiple choice

model. The listing of control variables is organized into the groups defined in equations (1) and

---

[11] Rejection of the fixed-effects specification in favor of random effects implies that observable
characteristics of classrooms and students are assigned randomly among teachers (no sorting). Because
we do not have repeat observations on individual students, we are unable to account for sorting based on
unobserved student-specific effects (e.g., Rothstein 2010), but we are able to incorporate measures of
student achievement more generally (GPA) and attitudes towards economics.
[12] Burnett and La Croix (2010) also analyzed a 20-question exam for high school economics students
using an OLS model and an alternative model for count data, the negative binomial; their results are
nearly identical across the two specifications.

(2). We list the coefficient on the pre-test score first. The first group of variables reflects fixed

classroom characteristics, including measures of whether the class received the Fed O&O

instructional materials ("experimental class"), the amount of time the teacher reported spending

on the monetary policy material, and three class-level dummies.[13] The next group contains

student characteristics including gender, race and ethnicity, self-reported high school GPA, peer

GPA (for all other observed students in the class), parents' education, and the student's attitude

toward studying economics. The final group contains teacher characteristics including years of

experience teaching economics, years of general teaching experience, gender, whether the

teacher has an undergraduate major or minor in economics, and whether the teacher has an

advanced degree. The teacher characteristics are only included in column 3, as explained below.

The regressions reported in Table 2 are based on random effects in columns 1 and 3 and

fixed effects in column 2. Each of these accounts for a teacher-specific component to the error

structure. The random effects (RE) specification incorporates a teacher-specific random

component, which is assumed to be uncorrelated with the other variables in the model. By

contrast, the fixed effects (FE) specification is equivalent to a model that includes a dummy

variable for every teacher and as such accounts for correlation between teacher-specific factors

and the other variables in the model (and hence precludes the inclusion of any observable teacher

variables).

Columns 1 and 2 are reported primarily for specification testing. Column 1 is the basic

RE specification without any teacher variables, to match the FE specification in column 2. The

---

[13] Because the class pairs for participating teachers were required to be at the same level, class level
indicators cannot be included in regressions that include teacher fixed effects (or are intended to be
comparable with such specifications), as in columns 1 and 2 of Table 2 and 3. Similarly, because only
two teachers reported spending slightly more time in the experimental class, the effects of the "time
spent" variable cannot be reliably identified in regressions that account for teacher fixed effects, hence it
is excluded from the first two columns.

coefficients are nearly identical across columns 1 and 2. The Hausman test statistic for the null hypothesis of random effects, listed near the bottom of the table for columns 1 and 2, is well below values that would imply rejection of the null hypothesis (statistical equivalence of the results across Columns 1 and 2) at conventional significance levels. This equivalence suggests that we can parameterize the teacher effects directly using observable teacher characteristics, without imparting substantial bias to the estimated coefficients on other variables.

Column 3 lists the results for the RE specification with a group of explicit teacher variables.[14] We focus here on a subset of key student and teacher characteristics (their relative magnitudes are discussed in the next sub-section). The coefficient on the pre-test score is positive and significant, indicating that students who start from a higher baseline achieve higher final scores. However, this coefficient is substantially smaller than one, indicating that the size of the typical gain declines with the level of the pre-test score (conditional on the other covariates). Students in the experimental classes that received the Fed O&O instructional materials recorded significantly higher post-test scores (conditional on pre-test scores) than did the students in the control classes; this largely replicates the key finding of Lopus and Hoff (2009), despite the expanded set of control variables used here. Additional time spent on the curriculum is associated with lower student achievement, which likely reflects extra time and effort exerted by teachers in classrooms where absorption of the material was observed to be slow.

Several student characteristics have large and statistically significant effects on post-test scores, most notably their high school GPAs, peer GPA, and self-reported attitude towards

---

[14] An alternative approach to the assessment of teacher effects would rely on a regression of the teacher fixed effects (estimated from the column 2 specification) on observable teacher characteristics (see e.g. Bosshardt and Watts 1990). In our data this approach generates point estimates for the teacher variables that are similar to those reported in column 3, but with much larger standard errors.

economics (which is likely a proxy for their expected performance or studying intensity for the course).[15] These findings are consistent with prior results in economic education (e.g. Watts 2005). In addition, female and minority students experience somewhat smaller score gains than do male and white students.

The results for the teacher characteristics in Table 2 indicate that an undergraduate major or minor in economics significantly improves student outcomes, as do advanced degrees and years of experience teaching economics.[16] These results generally are consistent with a number of studies cited earlier, regarding the favorable impact of teachers' specialized training and experience for student achievement (Harris and Sass, 2007; Clotfelter et al. 2007; Boyd et al. 2008; Rice 2010). Both undergraduate economics training and advanced degrees (typically in the education field in our sample) appear to enhance teacher effectiveness for the multiple choice testing mode in economics. Similarly, the positive coefficient on years spent teaching economics suggests either that sustained course specificity enables teachers to enhance their proficiency, or else that teachers sort over time into course specializations that make the best use of their talents and interests. By contrast, the coefficient on years of general teaching experience is negative and nearly significant at the 5% level, suggesting that time spent teaching subjects other then economics may reduce teacher effectiveness in teaching economics.

Table 3 lists results for the ordered logit model of essay test scores, with a three-column structure that parallels Table 2. RE and FE estimators and a corresponding Hausman test are not

---

[15] Our estimated peer effects are more robust (e.g., to teacher fixed effects) than those found by Clark, Scafidi, and Swinton (2011). This difference probably arises because we are able to directly identify peers who share a classroom, whereas Clark et al. can only identify peers who share a teacher (but are not necessarily in the same class).

[16] Despite the findings noted earlier (section II) that the first five years of teacher experience are especially valuable, we are unable to identify such differences in the estimated experience profiles because almost none of our teachers have fewer than five years of general or economics teaching experience. As such, the coefficients on the experience variables represent the impact of experience beyond five years.

feasible for the ordered logit model, so in columns 1 and 2 we list results for the basic model and a similar specification that includes an explicit set of teacher dummies (coefficients not reported). The estimated coefficients are relatively similar across columns 1 and 2, suggesting that observed and unobserved teacher effects are not significantly correlated with the other variables in the model. The primary exception is the coefficient on peer GPA, which declines substantially in size and statistical precision when teacher dummies are included. This suggests a tendency for high-achieving students to be sorted into classrooms taught by teachers with favorable characteristics.

Turning to the model that controls for observed teacher characteristics in column 3, the results are similar to the multiple choice models in regard to the important effects of students' own and peer GPA. Among teacher characteristics, the number of years teaching economics significantly enhances student achievement, similar to the results for the multiple choice outcomes. The estimated effect of an undergraduate degree in economics is negative and significant, but this is more than offset by a positive and significant effect of advanced degrees in other fields.[17] This finding suggests that in regard to teaching how to structure a logical argument and a clear narrative in an essay question, general teaching skills are critical, whereas undergraduate training in economics may undermine this process. We note that the negative effect of undergraduate economics training is surprising but is broadly consistent with the finding of Harris and Sass (2007) described earlier, regarding the negative effects of teachers' undergraduate math training on their high school students' math achievement. These findings suggest that the undergraduate economics curriculum content may not be closely aligned with the subject matter teachers are required to impart to their students once they arrive in the

---

[17] We obtain similar results for the impact of teachers' undergraduate economics training on the multiple choice and essay equations when we replace the indicator for an undergraduate major or minor in economics with teachers' number of undergraduate course units in economics.

classroom. Put differently, the specialized knowledge acquired through undergraduate economics may be "lost in translation" at the level of high school teaching, undermining rather than enhancing the ability to teach high school students how to structure an economic argument.

IV.C. Assessing the magnitudes of student, class, and teacher contributions to learning

The regression results discussed in the previous section identified significant effects for all three categories of inputs in the educational process. In this section, we assess the relative magnitudes of some key effects. Because the variables differ in their dimensions and scale, we translated them into consistent and comparable scales for evaluating their relative magnitudes. The resulting magnitude calculation is straightforward for the multiple choice models because each coefficient represents the effect of a one-unit change in the variable on the numerical score. For the ordered logit model, the coefficients require a transformation into probability space, which relies on the estimated coefficients in conjunction with the estimated constants; the latter are different for each outcome category in the model (see Wooldridge 2002, section 15.10.1, for details).

Table 4 lists the results for selected variables that produce statistically significant coefficients in column 3 of Tables 2 and 3. For each variable listed, we indicate the unit of change assessed (a change of 1 for dummy variables, one standard deviation for other variables) and the calculated effect on the outcome. For the multiple choice models in Panel A, the effect is in terms of the numerical score. For the essay models, the effects in Panel B of Table 4 are calculated as the percentage point increase in the probability of receiving a score of 2 rather than 0. These effects on essay scores are most meaningful by comparison with the shares of students who earned those scores (20.5% scored a 2 and 34.5% scored a 0).

16

The results in Table 4 indicate that students' attitude toward economics has the largest effects on achievement: compared with those who dislike economics, students who are excited about studying it see an increase in their multiple choice scores by more than an additional question and an increase in the probability of receiving an essay score of 2 (rather than 0) of 11 percentage points; this latter effect is especially large relative to the 20.5 percent sample incidence of this score. Higher own and peer GPAs also raise post-test scores by substantial amounts, particularly for the essay test. The Fed O&O instructional materials raised multiple choice scores by nearly as much as a standard deviation increase in peer GPA.

The effects of teacher characteristics are especially notable, given their implications for teacher training. Undergraduate degrees in economics and advanced degrees in general both enhance achievement on the multiple choice tests by amounts nearly as large as student enthusiasm about learning economics. Teachers' advanced degrees also enhance achievement on the essay exam by an amount nearly as large as student GPAs. However, undergraduate training in economics diminishes student achievement on the essay exams by an amount almost as large as the increase associated with an advanced degree. A standard deviation increase in the number of years of economics taught also enhances student achievement, by an amount equal to about one-quarter to one-half of the impact of the other key variables listed.

V. CONCLUSIONS

Using results of a special survey that was administered in 2006 to about 1000 high school economics students in California, we investigated the factors that contributed to student achievement on multiple-choice and essay exams on a monetary policy and central banking curriculum, relying on a standard value-added framework. We focused primarily on the

contributions of student characteristics such as their GPAs and attitudes toward learning economics, and teacher qualifications such as undergraduate economics training and advanced degrees. The results of our specification tests suggested that teacher and student characteristics generally are uncorrelated in our multiple choice testing sample, indicating that we can obtain unbiased estimates of the effects of the full range of student and teacher characteristics. Students' attitudes towards economics and their own and peer GPA have large effects on achievement. The effects of teachers' specialized training such as college economics coursework and advanced degrees were nearly as large as the primary student characteristics, although undergraduate economics training was associated with lower student achievement on the essay test.

We largely confirmed past findings regarding the important role of student enthusiasm and prior achievement for their performance in high school economics classes, along with the substantial contributions of specialized teacher training and experience. However, the inclusion of an essay exam in our study revealed the novel finding that teachers' undergraduate training in economics enhanced student performance on the multiple choice test but reduced it on the essay test. We have dubbed the negative impact of teachers' undergraduate economics training on students' essay performance as the "lost in translation" effect. By contrast, teachers' general advanced degrees, which are primarily in education, enhanced performance on both types of exams.

Our results are based on a relatively small sample and are restricted to instruction in economics. As such, they may not generalize to larger samples and other subjects. However, our finding of a "lost in translation" effect, in which content expertise may undermine the ability of teachers to impart knowledge at the appropriate level for the high school curriculum, merits

further investigation. By contrast, our findings for advanced degrees in general suggest that they enable teachers to achieve more consistent success in the classroom, perhaps by adapting content knowledge to the instructional needs of students. Given the upcoming wave of retirements by baby-boom generation educators (Aaronson and Meckel 2009), the number of newly minted college graduates who embark on high school teaching careers is likely to increase. Our findings suggest that educational policymakers should carefully consider how their specific skills can best be adapted to teaching the high school curriculum

References:

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools," *Journal of Labor Economics* 25: 95–135.

Aaronson, Daniel, and Katherine Meckel. 2009. "How will baby boomer retirements affect teacher labor markets?" *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol.33(4): 2-15.

Betts, Julian R., Andrew C. Zau and Lorien A. Rice. 2003. *Determinants of Student Achievement: New Evidence from San Diego*. San Francisco: Public Policy Institute of California.

Bosshardt, William and Michael Watts. 1990. "Instructor Effects and Their Determinants in Precollege Economic Education." *Journal of Economic Education* 21 (3, Summer): 265-276.

Boyd, Donald, Hamilton Lankford, Susanna Loeb, Jonah E. Rockoff, and James Wyckoff. 2008. "The Narrowing Gap in New York City Teacher Qualifications and Its Implications for Student Achievement in High-Poverty Schools." NBER Working Paper No. 14021. Cambridge, MA.

Burnett, Kimberly, and Sumner La Croix. 2010. "The Dog ATE My Economics Homework! Estimates of the Average Effect of Treating Hawaii's Public High School Students with Economics." University of Hawaii: UHERO Working Paper 2010-1.

Butters, Roger B. and Carlos J. Asarta. 2011. "A Survey of Economic Understanding in U.S. High Schools." *Journal of Economic Education* 42(2): 200-205.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2011. "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood." NBER Working Paper 17699. Cambridge, MA.

Clark, Christopher, Benjamin Scafidi, and John R. Swinton. 2011. "Do Peers Influence Achievement in High School Economics? Evidence from Georgia's Economics End of Course Test." *Journal of Economic Education* 42 (1): 3-18.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2007. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." CALDER Working Paper 11. Washington, DC: The Urban Institute.

Ferber, Marianne A., Bonnie G. Birnbaum and Carole A. Green. 1983. "Gender Differences in Economic Knowledge: a Reevaluation." *The Journal of Economic Education* 14(2): 24-37.

Hanushek, Eric A. and Steven G. Rivkin. 2010. "Using Value Added Measures of Teacher Quality." CALDER Policy Brief 9. Washington, DC: The Urban Institute.

Harris, Douglas N., and Tim R. Sass. 2007. "Teacher Training, Teacher Quality, and Student Achievement." CALDER Working Paper 3. Washington, DC: The Urban Institute.

Johnson, Marianne, Robson, Denise and Taengnoi, Sarinda, The Gender Gap in Economics: A Meta Analysis (August 22, 2011). Available at SSRN: http://ssrn.com/abstract=1914553 or http://dx.doi.org/10.2139/ssrn.1914553

Kane, Thomas J. and Eric S. Taylor, John H. Tyler, and Amy L. Wooten. 2010. "Identifying Effective Classroom Practices Using Student Achievement Data." NBER Working Paper No. 15803. Cambridge, MA.

Lopus, Jane S. and Jody Hoff. 2009. "An Empirical Analysis of Alternative Assessment Strategies in the High School Economics Class." *The American Economist* 53 (2): 38-51.

National Council on Economic Education. 2000. *Voluntary National Content Standards in Economics*. New York: National Council on Economic Education.

Rice, Jennifer. 2010. "The Impact of Teacher Experience: Examining the Evidence and Policy Implications." CALDER Policy Brief 11. Washington, DC: The Urban Institute.

Rivkin, Steven G., Eric A. Hanushek and John F. Kain. 2005. "Teachers, Schools and Academic Achievement." *Econometrica* 73(2): 417-58.

Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* 94(2): 247-52.

Rothstein, Jesse. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement." *The Quarterly Journal of Economics* 125(1): 175-214.

Subedi, B. R., B. Swan, and M.C. Hynes. 2009. "Predicting student gains with hierarchical models: A value-added approach to measure teacher effectiveness." Paper presented at the 2009 annual meeting of the American Educational Research Association, April 13-17, San Diego, CA.

Walstad, William B. 2001. "Economic Education in U.S. High Schools." *Journal of Economic Perspectives* 15 (3): 195-210.

Walstad, William B. and Ken Rebeck. 2001. *Test of Economic Literacy,* Third Edition. New York: National Council on Economic Education.

Walstad, William B. 1992. "Economics Instruction in High Schools." *Journal of Economic Literature*, vol. 30, No. 4 (Dec.) pp. 2019-2051.

Watts, Michael and William B. Walstad. 2011. "Research and Teaching High School Economics." In Mark C. Schug and William C. Wood, editors, *Teaching Economics in Troubled Times: Theory and practice for Secondary Social Studies.* New York: Routledge.

Watts, Michael. 2005. *What Works: A Review of Research on Outcomes and Effective program Delivery in Precollege Economic Education*. New York: National Council on Economic Education.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* Cambridge: MIT Press.

Table 1: Variable Definitions and Descriptive Statistics
(multiple choice sample; 982 students, 24 teachers)

Variable	Definition	Mean or sample share	Standard deviation
Test scores:			
Pre-test multiple choice	Number correct on multiple choice pretest (0-20, max=15)	8.05	2.58
Post-test multiple choice	Number correct on multiple choice post-test (0-20, max=20)	12.92	3.38
Pre-test essay	Score on essay pretest (0-3, max=3)	0.21	0.49
Post-test essay	Score on essay post-test (0-3, max=3)	1.05	0.96
Student characteristics:			
Female	Indicator that student is female	0.49	0.50
Race/ethnicity: White	Non-hispanic white	0.43	0.49
Asian	Asian or Pacific Islander	0.16	0.36
Black	Black/African American	0.04	0.20
Hispanic	Latin American/Hispanic/Chicano	0.20	0.40
Mixed race/ethnicity	Native American/other non-white	0.18	0.38
High school GPA (own)	High school GPA (self-reported)	3.11	0.59
Average GPA (peers)	Average GPA of peers (same class, in regression sample)	3.11	0.31
Parents' education:	Education of parent with highest attainment		
<High school		0.11	0.31
High school		0.12	0.32
Some college		0.30	0.46
College degree		0.30	0.46
Graduate degree		0.17	0.37
Attitude toward econ:			
Don't like	Expect to be among least favorite subjects	0.27	0.44
Indifferent	OK but not likely to be favorite subject	0.67	0.47
Excited	Econ likely to be a favorite subject	0.06	0.24

Table 1 (continued)

Variable	Definition	Mean or sample share	Standard deviation
Class characteristics (n=48)			
Experimental class	Indicator for experimental class (used O&O curriculum)	0.50	0.51
Time spent	Hours spent on curriculum material	5.21	2.60
Class level indicators: Mixed	Class level is mixed	0.63	0.49
Non-college	Class level is for non-college bound students	0.04	0.20
College prep	Class level is college preparatory	0.25	0.44
Advanced placement	Class level is Advanced Placement	0.08	0.28
Teacher characteristics (n=24)			
Years teaching econ	Number of years experience teaching economics	12.67	7.85
Years teaching	Number of years teaching experience	19.21	10.93
Female	Indicator that teacher is female	0.29	0.46
Undergrad econ major or minor	Indicator that teacher had an undergraduate major or minor in economics	0.33	0.48
Advanced degree	Indicator that the teacher has an advanced degree (beyond BA/BS)	0.67	0.48

Note: Means calculated for the sample used for the regression analysis of multiple choice outcomes (see Table 2); essay test scores based on a slightly smaller sample (963 students). The number of teachers is 24 and the number of classes is 48 (2 classes per teacher, experimental/control paired at same class level).

Table 2: Multiple Choice Regression Results (scores = 0 to 20)
(linear regression model)

VARIABLES (by category)	(1) Teacher Random Effects (RE)	(2) Teacher Fixed Effects (FE)	(3) RE (with teacher variables)
Pre-test score	0.284**	0.291**	0.227**
	(0.0346)	(0.0346)	(0.0367)
Class characteristics			
Experimental class	0.415*	0.415*	0.403*
	(0.163)	(0.162)	(0.177)
Time spent on monetary policy curriculum	--	--	-0.178**
			(0.0354)
Class level: non-college	--	--	-0.831
			(0.652)
Class level: college prep	--	--	-0.0727
			(0.236)
Class level: AP	--	--	-0.750
			(0.511)
Student characteristics			
Female	-0.335*	-0.319	-0.386*
	(0.165)	(0.164)	(0.180)
Asian	-0.460	-0.478	-0.697*
	(0.279)	(0.282)	(0.282)
Black	-0.686	-0.632	-1.365**
	(0.434)	(0.434)	(0.462)
Hispanic	-0.921**	-0.857**	-1.622**
	(0.273)	(0.276)	(0.284)
Mixed race/ethnicity	-0.395	-0.399	-0.582*
	(0.238)	(0.238)	(0.255)
High school GPA (own)	1.437**	1.444**	1.370**
	(0.164)	(0.166)	(0.179)
Average GPA (peers)	2.009**	2.070**	1.461**
	(0.547)	(0.792)	(0.455)
Parents' education: HS	0.114	0.0866	0.321
	(0.350)	(0.349)	(0.381)
Some college	0.268	0.217	0.628
	(0.310)	(0.309)	(0.334)
College degree	0.346	0.305	0.630
	(0.322)	(0.321)	(0.347)
Graduate degree	0.338	0.301	0.501
	(0.353)	(0.352)	(0.381)
			(continued)

Table 2 (continued)

VARIABLES (by category)	(1) Teacher Random Effects (RE)	(2) Teacher Fixed Effects (FE)	(3) RE (with teacher variables)
Student characteristics (con.)			
Attitude toward econ: Indiff	0.635	0.618	0.509
	(0.344)	(0.343)	(0.373)
Excited	1.134**	1.087**	1.288**
	(0.368)	(0.367)	(0.398)
Teacher characteristics			
Years teaching econ	--	--	0.0373*
			(0.0158)
Years teaching	--	--	-0.0216
			(0.0124)
Female	--	--	-0.169
			(0.238)
Undergrad econ	--	--	1.231**
(major or minor)			(0.248)
Advance degree (other)	--	--	1.100**
			(0.265)
Constant	-0.831	-0.991	1.552
	(1.793)	(2.598)	(1.398)
	Hausman test statistic: 13.94 Prob>chi2 = 0.603		
Observations	982	982	982
Number of teachers	24	24	24

Note: Teacher effects treated as random effects (columns 1 and 3) or fixed effects (column 2). Standard errors in parentheses. Omitted categories for multiple group dummy variables are mixed class level, white race for students, parents education is less than high school, and attitude towards econ is "don't like."
** $p<0.01$, * $p<0.05$

Table 3: Essay Test Regression Results (scores = 0 to 3)
(ordered logit model)

VARIABLES (by category)	(1) No teacher effects	(2) Teacher dummies	(3) With teacher variables
Pre-test score	0.700**	0.579**	0.683**
	(0.136)	(0.145)	(0.136)
Class characteristics			
Experimental class	0.117	0.115	0.137
	(0.123)	(0.127)	(0.124)
Time spent on monetary policy curriculum	--	--	-0.0451
			(0.0242)
Class level: non-college	--	--	-0.345
			(0.478)
Class level: college prep	--	--	0.352*
			(0.161)
Class level: AP	--	--	-0.0538
			(0.353)
Student characteristics			
Female	-0.163	-0.269*	-0.161
	(0.125)	(0.130)	(0.126)
Asian	-0.100	0.00853	-0.209
	(0.186)	(0.216)	(0.198)
Black	-0.526	-0.832*	-0.616
	(0.337)	(0.369)	(0.346)
Hispanic	-0.0535	-0.0906	-0.0508
	(0.189)	(0.219)	(0.201)
Mixed race/ethnicity	0.131	0.0574	0.0666
	(0.173)	(0.184)	(0.176)
High school GPA (own)	0.859**	0.924**	0.860**
	(0.125)	(0.133)	(0.126)
Average GPA (peers)	1.713**	1.331*	1.393**
	(0.234)	(0.606)	(0.321)
Parents' education: HS	0.0943	0.132	0.146
	(0.269)	(0.284)	(0.275)
Some college	0.0669	0.129	0.138
	(0.233)	(0.250)	(0.240)
College degree	0.358	0.402	0.390
	(0.238)	(0.254)	(0.244)
Graduate degree	0.404	0.435	0.420
	(0.262)	(0.276)	(0.267)

Table 3 (continued)

VARIABLES (by category)	(1) No teacher effects	(2) Teacher dummies	(3) With teacher variables
Student characteristics (con.)			
Attitude toward econ: Indiff	0.238	0.231	0.278
	(0.268)	(0.281)	(0.273)
Excited	0.547	0.556	0.631*
	(0.284)	(0.298)	(0.289)
Teacher characteristics			
Teacher dummies	No	Yes	No
Years teaching econ	--	--	0.0223*
			(0.0111)
Years teaching	--	--	0.00264
			(0.00895)
Female	--	--	0.0571
			(0.161)
Undergrad econ (major or minor)	--	--	-0.367*
			(0.175)
Advance degree (other)	--	--	0.439*
			(0.184)
Constants			
Cut 1	7.801**	6.412**	7.236**
	(0.776)	(2.056)	(1.045)
Cut 2	9.628**	8.456**	9.135**
	(0.796)	(2.063)	(1.059)
Cut 3	11.32**	10.24**	10.83**
	(0.821)	(2.076)	(1.076)
Observations	963	963	963
Number of teachers	24	24	24

Note: Teacher effects in column 2 estimated using a complete set of teacher dummies (23); coefficients not reported. Standard errors in parentheses. Omitted categories for multiple group dummy variables are mixed class level, white race for students, parents education is less than high school, and attitude towards econ is "don't like."
** p<0.01, * p<0.05

Table 4: Magnitude Assessment, Selected Coefficients

Panel A: Multiple choice scores
(Table 2, column 3)

Variable	Unit of change (SD, or 1 for dummy variables)	Effect on outcome
Student characteristics		
Experimental class	1	0.403
High school GPA (own)	0.591	0.809
Average GPA (peers)	0.312	0.455
Student attitude: Excited	1	1.288
Teacher characteristics		
Years teaching econ	7.89	0.294
Undergrad econ (major or minor)	1	1.231
Advanced degree	1	1.100

Panel B: Essay scores
(Table 3, column 3)

Variable	Unit of change (SD, or 1 for dummy variables)	Effect on outcome
Student characteristics		
High school GPA (own)	0.597	0.091
Average GPA (peers)	0.321	0.079
Student attitude: Excited	1	0.110
Teacher characteristics		
Years teaching econ	7.92	0.030
Undergrad econ (major or minor)	1	-0.061
Advanced degree	1	0.072

Panel B effect indicates percentage increase in probability of receiving post-test essay score=2 rather than 0 (base sample shares=0.205, 0.345). See text for discussion. SD (standard deviations) of class and teacher variables were calculated across students rather than classes/teachers (and therefore may differ from those listed in Table 1).