

FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

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

Robert G. Valletta
Federal Reserve Bank of San Francisco

K. Jody Hoff
Federal Reserve Bank of San Francisco

Jane S. Lopus
California State University, East Bay

April 2013

Working Paper 2012-03

<http://www.frbsf.org/publications/economics/papers/2012/wp12-03bk.pdf>

The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System.

LOST IN TRANSLATION? TEACHER TRAINING AND OUTCOMES IN HIGH SCHOOL ECONOMICS CLASSES

Robert G. Valletta
(corresponding author)

Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
(415) 974-3345
rob.valletta@sf.frb.org

K. Jody Hoff

Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
(415) 974-2952
jody.hoff@sf.frb.org

Jane S. Lopus

California State University, East Bay
Hayward, CA 94542
(510) 885-3140
jane.lopus@csueastbay.edu

April 15, 2013

(Previous versions March and August 2012)

Forthcoming in *Contemporary Economic Policy*

The authors thank Katherine Kuang for helpful research assistance. They also thank seminar participants at the January 2011 American Economic Association meetings and the June 2011 AEA National Conference on Teaching Economics for their comments, and especially Ron Baker for his helpful discussion at the latter meeting. Two anonymous referees provided helpful comments as well. Prior versions of the paper were presented under the title “Teacher and Student Characteristics as Determinants of Success in High School Economics Classes.” The views expressed in this paper are solely those of the authors and are not attributable to the Federal Reserve Bank of San Francisco or the Federal Reserve System.

LOST IN TRANSLATION? TEACHER TRAINING AND OUTCOMES IN HIGH SCHOOL ECONOMICS CLASSES

Abstract

Using data from a 2006 survey of California high school economics classes, we assess the effects of 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. The results show a substantial impact of specialized teacher experience and college-level coursework in economics. However, the latter is associated with higher scores on the multiple-choice test and lower scores on the essay test, suggesting that a portion of teachers' content knowledge may be "lost in translation" when conveyed to their students.

(JEL classifications: A21, I21)

LOST IN TRANSLATION? TEACHER TRAINING AND OUTCOMES IN HIGH SCHOOL ECONOMICS CLASSES

I. INTRODUCTION

Assessing the contribution of teacher quality to student achievement is a key issue in the educational research and policy fields (e.g., National Research Council 2010). A central finding from this literature is that the quantitative contribution of teacher quality to student outcomes is large, with substantial gains in student achievement possible in response to systematic improvements in teacher quality (Hanushek 2011). However, research findings regarding the relationship between teacher quality and measurable characteristics such as training and experience have been mixed, with little support found for systematic, consistent contributions of identifiable contributors to teacher preparation (e.g., Rockoff 2004; Rivkin, Hanushek, and Kain 2005; Aaronson, Barrow, and Sander 2007; Hanushek and Rivkin 2010; Clotfelter, Ladd, and Vigdor 2010; Harris and Sass 2011). This failure to identify systematic contributions of observable teacher characteristics may reflect imperfect identification of relevant teacher skills and the specific settings in which they make a difference.

In this paper, we contribute to the literature on teacher quality and student outcomes by investigating the impact of teacher characteristics on students' success in high school economics classes. Much of the voluminous literature on educational outcomes focuses on students in primary school. By contrast, 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. We distinguish between teachers' specific training and teaching experience in the economics field and their general training, including post-graduate degrees (primarily in the education field). Our breakdown between categories of teacher skills follows

the distinction in the education literature between “content knowledge,” which focuses on a teacher’s understanding of the specific subject being taught, and “pedagogical content knowledge,” which refers to a broader understanding of how learners acquire knowledge in that subject (e.g., Shulman 1986; Ball, Thames, and Phelps 2008; National Research Council 2010). This literature provides an informative backdrop for the interpretation of our results.

Our data are from a special survey of California high school economics teachers conducted in 2006. The survey produced value-added outcome data, which includes test scores before and after a limited instructional window, for nearly 1,000 students in 48 matched-pair classes taught by 24 teachers. These data were first used by Lopus and Hoff (2009), who examined the contribution of specialized instructional materials to student learning. In the present paper, we have expanded the analysis to address the broader question of how teacher characteristics affect the learning process, and how the contribution of teacher characteristics compares with the contribution of student characteristics. These questions were not addressed in the earlier Lopus and Hoff paper. In addition, relative to that paper, we more fully exploit the study’s experimental design by employing an econometric approach that enables us to explicitly account for observed and unobserved teacher effects using fixed-effects and random-effects estimation. The survey data provide a broad set of student and teacher characteristics, along with pre-test and post-test outcomes for two testing modes: multiple choice questions and an essay question. The essay mode represents a novel element of our analysis, given the economic education literature’s typical reliance on standardized multiple choice questions.

In the next section, we discuss the relevant prior literature in more detail, including a discussion of how our paper adds to this literature. In Section III, we describe our data, followed in Section IV by the discussion of our value-added methodological framework and results. The

value-added approach, which focuses on examining the determinants of improvement in student outcomes, is commonly employed in the education and economics literatures (see e.g. Rivkin et al. 2005; McCaffrey and Hamilton 2007; Rothstein 2010; Chetty, Friedman, and Rockoff 2011). Our regression results based on this framework indicate that student characteristics including own GPA and peer GPA have the expected large effects on improvements in student outcomes. Teacher characteristics such as experience teaching economics and formal education in economics also are associated with substantial gains in student performance, by amounts that are statistically significant and nearly as large as the effects of student characteristics. However, the impact of teachers' formal education in economics varies across the two 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 knowledge alone may add less value than a deep pedagogical understanding of how to convey specialized high school subject matter such as economics to students. We detail the implications of our findings in the Conclusion section.

II. LITERATURE REVIEW

Education Literature

Teacher quality has been identified in a variety of studies as a significant determinant of student achievement (e.g., Hanushek 2011; Hanushek and Rivkin 2010; Kane et al. 2011). 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

teacher-specific contributions to student achievement (Rockoff 2004; Rivkin et al. 2005; Aaronson et al. 2007). As discussed by Harris and Sass (2011), 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, such as the attainment of advanced degrees. For example, Aaronson et al. (2007) and Betts, Zau, and Rice (2003) investigated the effects of teachers' college majors and found no systematic impact on student achievement in high school. Harris and Sass (2011) generally found no significant effects of teachers' advanced degrees on student achievement, using administrative data for the universe of third through tenth grade students in Florida.¹

In contrast to the ambiguous effects of formal education, direct teaching experience generally is associated with enhanced student outcomes at both the elementary and high school levels, although the effects are largely limited to early years of experience (Rivkin et al. 2005; Boyd et al. 2008; Clotfelter et al. 2010; Rice 2010; Harris and Sass 2011). 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. (2010) found that the effects of teacher credentials are larger than the effects of student characteristics. Similarly, Harris and Sass (2011) found that professional development in the form of content-specific training raised student achievement.

¹ Clotfelter et al. (2010) uncovered evidence that teachers' advanced degrees may be associated with lower student achievement. In particular, using a panel of ninth and tenth grade students from North Carolina, they found a large negative effect on end-of-course test scores for teachers holding a doctorate. Given the small number of such teachers in their sample, this may reflect unobserved characteristics of those teachers rather than a causal relationship.

Conflicting results on factors affecting teacher quality are often attributed to general differences in methodology and data, particularly given the likely influence of nonrandom matching of students and teachers (Kane et al. 2011; Rothstein 2010). However, these conflicting results may also reflect inadequate precision in the identification of critical teacher skills and the settings in which they enhance student performance. Education researchers have long emphasized the distinction between teachers' content knowledge (CK) of their specific field and their pedagogical content knowledge (PCK), or ability to convey the subject matter to their students (Shulman 1986; Ball, Thames, and Phelps 2008; National Research Council 2010). In recent research based on a large sample of German secondary school students, Baumert et al. (2010) reported that, while both CK and PCK affect student outcomes in mathematics and are highly correlated, CK was less predictive of instructional quality and student progress than PCK. Our focus on high school economics is well suited to provide a further test of this distinction between CK and PCK, particularly given the measures of teacher background available in our data.

Economic Education Literature

A parallel literature within the economic education field has focused on how teacher and student characteristics 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. He noted that existing research emphasized the importance of teacher training in economics for student performance in economics classes. For example, Bosshardt and Watts (1990) and Walstad (2001) found that the most effective teachers were those who had completed more courses in economics. Similarly, Swinton et al. (2010) found that additional content training, obtained through attendance at specialized in-service

training workshops in economics, enhanced student performance in Georgia's mandatory high school economics curriculum and examination. By contrast, teacher characteristics other than coursework in economics have had inconsistent or insignificant effects on student learning in various studies. These characteristics include completion of noncredit workshops, years of teaching experience, years since the last economics course was taken, and percentage of teaching load that is economics (Walstad 1992).

With respect to student characteristics, measures of student intellect or prior classroom performance in multiple subjects are consistently found to be positively correlated with student achievement in economics (Watts 2005). Students in higher level 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.² 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).

We rely on this prior research to guide our empirical specification. However, it is important to note that past findings regarding student performance in high school economics courses generally pertain to objective test modes such as multiple choice questions.³ For the present study, we supplemented standard multiple choice questions with an essay question, as described in more detail in the next section. The inclusion of an essay question enables us to examine whether the contribution of specific forms of teacher training and experience are

² The gender effect is not consistent in past research (see Johnson, Robson, and Taengnoi 2012), and it is likely that racial gaps reflect school district and other student background characteristics.

³ Exceptions to this focus on multiple choice questions include the paper that first used our data set (Lopus and Hoff 2009), and a much earlier paper focusing on gender differences in achievement in university-level economics classes (Ferber, Birnbaum, and Green 1983).

consistent with the distinction between content knowledge and pedagogical skills discussed in the previous sub-section. In addition, relative to the educational literature's focus on students in primary school, our focus on student achievement in specialized high school subject matter provides a different perspective on the contribution of specific forms of teacher training to student achievement.

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 the Federal Reserve Bank of San Francisco (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 (National Council on Economic Education 2000). The project was administered in California, where a semester course in economics has been required for high school graduation since 1989.⁵ Economics teachers at all public high schools in California (approximately 1,000) 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, noncollege bound, or mixed). For each teacher in the final sample, one class served as the experimental

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.

⁵ See also Gratton-Lavoie and Gill (2009) for a related study that uses value-added performance data for California high school economics students. Our paper is distinguished from theirs by our focus on observable teacher characteristics, which are not available in their data.

class, with the *Open and Operating* (O&O) video and curriculum used in addition to the materials normally used to teach about the Federal Reserve and monetary policy. The other class served as the control class, with the material on monetary policy taught in the teachers' traditional manner. The O&O curriculum requires about 2½ hours of class time; in the control class, other instructional materials were used in place of O&O, so that approximately equal amounts of time were spent on monetary policy in both classes. Participating teachers administered student questionnaires and completed a teacher questionnaire. Evaluation of student outcomes was based on pre- and post-tests for a set of 20 multiple choice questions and a separate essay question, described in more detail below. Participating teachers were instructed to incorporate the post-test scores into students' course grades and to announce this to students in advance, so it is likely that students were motivated to absorb the curriculum and do well on the post-test exam (which is also reflected in the typical score gains between the pre-test and post-test, as discussed in conjunction with Table 1 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 the other to be the control class. Forty-three teachers returned some materials and 24 returned the complete sets of materials used in this study.⁶ 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 returning complete information for the analysis of essay scores.

⁶ 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.

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 specifically 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 questions used were identical in the pre-test and post-test phases, except that the order was changed. The essay question asked 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; like the multiple choice questions, it was worded identically in the pre-test and post-test phases. This question was open-ended in the sense that students chose for themselves what type of monetary policy challenge to address.

Inclusion of the essay question represents a novel form of testing and instructional assessment in the present study relative to most past research. Student performance on this question reflects not only specific knowledge of the economics content taught in the monetary policy module but also their ability to write clearly and structure a complete, coherent argument based on their understanding of the relevant economics concepts. As such, students' essay performance reflects teachers' pedagogical skills as well as their content knowledge.

To ensure grading objectivity and comparability across classes on the essay question, scoring for this question was performed by a panel of six 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 study, hence eliminating potential conflicts of interest in grading. 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 essay questions. The panel utilized a scoring rubric to evaluate the essay responses, with grades ranging from 0 (lowest) to 3 (highest).⁷ Each essay was read by two teachers. If their scores differed, a third teacher also graded the essay.

There are advantages and disadvantages to using specialized survey data such as those used for this paper. The disadvantages include small sample sizes and the possibility of greater measurement error than is present in administrative data sets. However, our survey data possess the advantage of focusing on student achievement in a specific subject under quasi-experimental conditions, unlike studies using administrative data in which the subject matter and learning environments are more varied. In addition, a key advantage of our survey data is the availability of a detailed set of teacher and student characteristics that can be used in the estimation of our student achievement (value-added) equations, as discussed further in Section IV.

Table 1 reports variable definitions and descriptive statistics. Post-test multiple choice and essay scores serve as dependent variables in the regression analysis, controlling for pre-test scores, as described in the next section. The mean values indicate that scores were raised substantially by introducing the 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 score on the essay pre-test as well as the post-test.

⁷ 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.

We use student characteristics, class characteristics, and teacher characteristics as control variables in our analysis 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 grade point average (GPA) as reported near the beginning of the current course, and parents' educational attainment. Their attitudes toward economics mostly consist of indifference or 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 average, with the full range extending from a minimum of 2 hours to a maximum of 12.⁸ Most of the classes were mixed in terms of students' academic aspirations and likelihood of college attendance.⁹

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 analysis. One-third of the teachers possess a college undergraduate major or minor degree in economics, while two-thirds hold some type of advanced degree (typically a master's in education); half of those who were undergraduate economics majors hold an advanced degree, so the overlap between

⁸ 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.

⁹ Because complete data records were not obtained for all students in each class, we are unable to form an accurate class-size variable to use in our empirical analysis.

these two groups is small enough for estimation of their independent impacts on student achievement.¹⁰

IV. REGRESSION FRAMEWORK AND RESULTS

Regression specification and caveats

Our analysis relies on the well-established value-added approach for estimating the contributions of student, teacher, and classroom characteristics to educational outcomes (e.g., Rivkin et al. 2005; Rothstein 2010; Chetty, Friedman, and Rockoff 2011; McCaffrey and Hamilton 2007). This approach begins with an acknowledgment that students' achievement in a particular grade level or course of study is likely to depend on their background characteristics and prior achievement. These characteristics may in turn be systematically related to other modeled factors, such as teacher background characteristics, and as such may bias the estimated effects of these factors in statistical studies. The value-added approach therefore focuses on the incremental gains observed for students during a particular grade or course of study, with an initial assessment or pre-test serving as a measure of students' prior and potential achievement.

Our experimental design fits well within the value-added framework. Tests were administered to participating students in our sample of California high school economics classes before and after they completed a short unit on monetary policy. Although the value-added approach incorporates the pre-test measure of baseline achievement, it is typically bolstered through incorporation of controls for observable student characteristics such as prior grades, which we have in our data (described below). Moreover, our matched experimental sample and

¹⁰ 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 statistical inference, and we therefore do not distinguish this individual from advanced degree holders more generally.

ability to model individual teacher effects helps us 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.

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

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 observation period), plus vectors of classroom instructional characteristics C_j , individual student characteristics S_{ijk} , and teacher effects μ_k .¹¹ 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, see equation (2)). 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 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 in an equation that also includes a set of explicit teacher characteristic variables, or T_k in equation (2):

¹¹ 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 equation (1) to be equal to 1. This restriction is strongly rejected in our empirical results reported in Tables 2 and 3.

$$A_{ijk} = \lambda_0 + \lambda_1 A_{ijk}^0 + \gamma_1 C_j + \gamma_2 S_{ijk} + \gamma_3 T_k + \eta_{ijk} \quad (2)$$

where

$$\eta_{ijk} = \xi_k + v_{ij}$$

(an error term that has zero mean conditional on the right-hand side variables, divided into separate components for students sharing the same teacher k and an individual student in a particular class ij).

There is no guarantee that we can obtain unbiased estimates of the coefficients on the teacher characteristic variables (γ_3) in equation (2). Rejection of the fixed-effects specification in favor of random effects simply implies that the effects of observable student and classroom characteristics are unrelated to the teachers to whom students are assigned, hence the estimated effects of student and classroom characteristics are unbiased in the random effects specification with explicit teacher variables. To account for sorting based on unobserved student-specific effects that may bias the coefficients on teacher characteristics, we would need repeat observations on individual students across multiple teachers in our data set (Rothstein 2010). This is the approach taken in studies that use data on early elementary school students from Tennessee's Student-Teacher Achievement ratio (STAR) project, although such near-randomized data may have their own drawbacks (see Hanushek 1999 and Krueger 1999 for additional discussion). A similar approach is taken in a limited set of studies that examines student outcomes at the middle school and high school levels (e.g., Harris and Sass 2011, Clotfelter et al. 2010).

A longitudinal approach is precluded in our study not only by our specific data but also by our focus on a single specialized high school subject area, economics, which students in

general only take once and for which they only encounter one teacher.¹² As such, our results for teacher characteristics are not robust to the influence of unobserved student characteristics and sorting and hence cannot be regarded as definitive. This is particularly true given our small sample of only 24 teachers (and 48 classrooms), such that the effects of observable teacher characteristics are identified for small numbers of teachers who possess those characteristics. On the other hand, we are able to incorporate detailed student background measures into the analysis, including own and peer GPAs measured at the start of the course, plus family characteristics, including parents' educational attainment. In conjunction with measured pre-test performance, these detailed controls may adequately account for unobserved student effects in our econometric estimates.

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 to analyze the multiple choice outcomes.¹³ 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 subsection.

¹² See also National Research Council (2010, Chapter 2) for discussion of the challenges posed by randomized sample designs for educational research.

¹³ Burnett and La Croix (2010) also analyzed a 20-question exam for high school economics students using an ordinary least squares model and an alternative model for count data, the negative binomial; their results are nearly identical across the two specifications.

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 (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.¹⁴ The next group contains student characteristics including gender, race or ethnicity, self-reported high school GPA, average 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.

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

¹⁴ 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 Tables 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.

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 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.¹⁵ We focus here on a subset of key student and teacher characteristics (their relative magnitudes are discussed in the next subsection). 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

¹⁵ 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 (e.g., Bosshardt and Watts 1990, Aaronson et al. 2007). 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.

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 GPAs, and self-reported attitudes toward economics (which is likely a proxy for their expected performance or studying intensity for the course).¹⁶ These findings are consistent with prior results in economic education (e.g., Watts 2005; Gratton-Lavoie and Gill 2009). 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 a teacher's undergraduate major or minor in economics significantly improves student outcomes, as do advanced degrees and more years of experience teaching economics.¹⁷ These results suggest an important role of targeted training and content knowledge (CK) for teachers' ability to successfully convey specialized subject matter (Harris and Sass 2007; Clotfelter et al. 2007; Boyd et al. 2008; Rice 2010; Swinton et al. 2010). However, advanced degrees beyond college (typically a master's degree in education in our sample) also enhance teacher effectiveness for the multiple-choice testing mode, suggesting that a deeper understanding of pedagogy (PCK) also is useful for teaching high school economics. The significant positive coefficient on years

¹⁶ 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).

¹⁷ Despite the earlier findings noted in 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 has 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.

spent teaching economics also is indicative of a positive contribution from CK.¹⁸ By contrast, the coefficient on years of general teaching experience is negative and nearly significant at the 5 percent level, suggesting that time spent teaching subjects other than economics may reduce teacher effectiveness in teaching economics. The estimated effects of teaching experience should be viewed with caution, however. As noted by Murnane and Phillips (1981), the relationship between teaching experience and student achievement may reflect the influence of unobserved factors that influence teaching performance, such as innate ability, sorting, or time devoted to teaching, rather than learning by doing.

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 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 GPAs. Among teacher characteristics, the number of years teaching economics significantly enhances student achievement, similar to the results for the multiple choice

¹⁸ In regressions not displayed, we included the percentage of each teacher's load allotted to economics instruction. That variable has a small positive but highly insignificant coefficient in the multiple choice and essay equations, with essentially no change in the other coefficients.

outcomes. However, the estimated effect of a teacher's undergraduate degree in economics is negative and significant, in contrast to the positive effect estimated for the multiple choice test.¹⁹ This finding suggests that the CK obtained through undergraduate economics is not adequate or may detract from teachers' ability to teach students how to construct the logical arguments and clear narratives for a successful essay response. Put differently, the specialized knowledge teachers acquire through undergraduate economics may be lost in translation at the level of high school teaching, undermining rather than enhancing their ability to teach high school students how to structure an economic argument. By contrast, the impact of advanced degrees in other fields (primarily education) are positive and significant for students' essay scores, suggesting that teachers with such degrees possess the pedagogical skills needed to convey to students how to structure a successful essay response.

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 translate 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).

¹⁹ We obtain results for teachers' undergraduate economics training similar to those reported in Tables 2 and 3 when we replace the indicator for an undergraduate major or minor in economics with teachers' number of undergraduate course units in economics.

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 percent scored a 2 and 34.5 percent scored a 0).

The results in Table 4 indicate that students' attitudes toward economics have the largest effects on achievement. Compared with those who dislike economics, students who are excited about studying it raise their multiple choice scores by more than an additional question and are 11 percentage points more likely to receive an essay score of 2 (rather than 0); 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.

Our primary focus is on the effects of teacher characteristics. Undergraduate degrees in economics and advanced degrees beyond college 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 teaching economics

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 administered in 2006 to about 1,000 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 curriculum. Our empirical analysis relies on a standard value-added framework based on pre-test and post-test scores for both testing modes. We focused primarily on teacher qualifications such as undergraduate economics training and advanced degrees, which we compared to the contributions of student characteristics such as their GPAs and attitudes toward learning economics. 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 characteristics. We cannot guarantee that our estimates for the impacts of specific teacher characteristics are unbiased, but our extensive controls for general student background and achievement bolster the reliability of these findings.

Our empirical findings indicate that the effects of teachers' specialized training such as college economics coursework and advanced degrees were nearly as large as the effects of key student background characteristics, such as own and peer GPAs. However, teachers' undergraduate economics training was associated with lower student achievement on the essay test, whereas advanced degrees in fields other than economics (primarily in the education field) contributed to improved student outcomes for both types of testing modes. We have dubbed the

negative impact of teachers' undergraduate economics training on students' essay performance as the lost in translation effect, because it implies that high school teachers with specialized training in economics may not always be able to successfully convey their understanding of the field to their students.

Our results are based on a relatively small sample and are restricted to instruction in economics. Further, they are not robust to unobserved student characteristics that may be correlated with observed teacher characteristics. As such, they may not generalize to larger samples and other subjects. However, our finding that 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 outside of economics suggest that they enable teachers to achieve more consistent success in the classroom, perhaps by adapting content knowledge to students' instructional needs. Our overall results suggest that researchers who are investigating the factors that contribute to successful teachers should heed the longstanding distinction between content knowledge and pedagogical content knowledge emphasized in the education literature (Shulman 1986; Ball et al. 2008; National Research Council 2010). 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.
- Ball, Deborah Loewenberg, Mark Hoover Thames, and Geoffrey Phelps. 2008. "Content Knowledge for Teaching: What Makes It Special?" *Journal of Teacher Education* 59(5, Nov./Dec.): 389-407.
- Baumert, Jürgen, Mareike Kunter, Werner Blum, Martine Brunner, Thamar Voss, Alexander Jordan, Uta Klusman, Stefan Krauss, Michael Neubarnd, and Yi-Miau Tsai. 2010. "Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal* 47 (1, March): 133-180.
- 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." *Journal of Policy Analysis and Management* 27(4): 793-818.
- 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. 2010. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." *Journal of Human Resources* 45(3): 655-681.
- 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.
- Gratton-Lavoie, Chiara, and Andrew Gill. 2009. "A Study of High School Economic Literacy in Orange County, California." *Eastern Economic Journal* 35: 433-51.
- Hanushek, Eric A. 1999. "Some Findings From an Independent Investigation of the Tennessee STAR Experiment and From Other Investigations of Class Size Effects." *Educational Evaluation and Policy Analysis* 21 (2, Summer): 143-163.
- Hanushek, Eric A. 2011. "The Economic Value of Higher Teacher Quality." *Economics of Education Review* 30(2, June): 466-479.
- Hanushek, Eric A. and Steven G. Rivkin. 2010. "Generalizations about Using Value-Added Measures of Teacher Quality." *American Economic Review* 100(2, May): 267-271.
- Harris, Douglas N., and Tim R. Sass. 2011. "Teacher Training, Teacher Quality and Student Achievement." *Journal of Public Economics* 95 (Aug.): 798-812
- Johnson, Marianne, Denise Robson, and Sarinda Taengnoi. 2012. "The Gender Gap in Economics: A Meta Analysis." SSRN working paper. Available at: <http://ssrn.com/abstract=1914553>
- Kane, Thomas J., Eric S. Taylor, John H. Tyler, and Amy L. Wooten. 2011. "Identifying Effective Classroom Practices Using Student Achievement Data." *Journal of Human Resources* 46(3): 587-613.
- Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics* 114: 497-532.
- 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.
- McCaffrey, Daniel F. and Laura S. Hamilton. 2007. *Value-Added Assessment in Practice: Lessons from the Pennsylvania Value-Added Assessment System Pilot Project*. Santa Monica, CA: RAND.
- Murnane, Richard J., and Barbara R. Phillips. 1981. "What Do Effective Teachers of Inner-City Children Have in Common?" *Social Science Research* 10: 83-100.
- National Council on Economic Education. 2000. *Voluntary National Content Standards in Economics*. New York: National Council on Economic Education.

- National Research Council. 2010. *Preparing teachers: Building evidence for sound policy*. Committee on the Study of Teacher Preparation Programs in the United States, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- 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.
- Shulman, Lee S. 1986. "Those Who Understand: Knowledge Growth in Teaching." *Educational Researcher* 15(2, Feb.): 4-14.
- Swinton, John R., Thomas De Berry, Benjamin Scafidi, and Howard C. Woodard. 2010. "Does in-service professional learning for high school economics teachers improve student achievement?" *Education Economics* 18(4): 395-406.
- Walstad, William B. 1992. "Economics Instruction in High Schools." *Journal of Economic Literature*, vol. 30, No. 4 (Dec.): 2019-2051.
- 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.
- 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

(continued)

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)

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

(continued)

Table 2 (continued)

VARIABLES (by category)	(1) Teacher Random Effects (RE)	(2) Teacher Fixed Effects (FE)	(3) RE (with teacher variables)
<u>Student characteristics (con.)</u>			
Attitude toward econ: Indiff	0.635 (0.344)	0.618 (0.343)	0.509 (0.373)
Excited	1.134** (0.368)	1.087** (0.367)	1.288** (0.398)
<u>Teacher characteristics</u>			
Years teaching econ	--	--	0.0373* (0.0158)
Years teaching	--	--	-0.0216 (0.0124)
Female	--	--	-0.169 (0.238)
Undergrad econ (major or minor)	--	--	1.231** (0.248)
Advanced degree (other than econ)	--	--	1.100** (0.265)
Constant	-0.831 (1.793)	-0.991 (2.598)	1.552 (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** (0.136)	0.579** (0.145)	0.683** (0.136)
<u>Class characteristics</u>			
Experimental class	0.117 (0.123)	0.115 (0.127)	0.137 (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)
<u>Student characteristics</u>			
Female	-0.163 (0.125)	-0.269* (0.130)	-0.161 (0.126)
Asian	-0.100 (0.186)	0.00853 (0.216)	-0.209 (0.198)
Black	-0.526 (0.337)	-0.832* (0.369)	-0.616 (0.346)
Hispanic	-0.0535 (0.189)	-0.0906 (0.219)	-0.0508 (0.201)
Mixed race/ethnicity	0.131 (0.173)	0.0574 (0.184)	0.0666 (0.176)
High school GPA (own)	0.859** (0.125)	0.924** (0.133)	0.860** (0.126)
Average GPA (peers)	1.713** (0.234)	1.331* (0.606)	1.393** (0.321)
Parents' education: HS	0.0943 (0.269)	0.132 (0.284)	0.146 (0.275)
Some college	0.0669 (0.233)	0.129 (0.250)	0.138 (0.240)
College degree	0.358 (0.238)	0.402 (0.254)	0.390 (0.244)
Graduate degree	0.404 (0.262)	0.435 (0.276)	0.420 (0.267)

(continued)

Table 3 (continued)

VARIABLES (by category)	(1) No teacher effects	(2) Teacher dummies	(3) With teacher variables
<u>Student characteristics (con.)</u>			
Attitude toward econ: Indiff	0.238 (0.268)	0.231 (0.281)	0.278 (0.273)
Excited	0.547 (0.284)	0.556 (0.298)	0.631* (0.289)
<u>Teacher characteristics</u>			
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)
Advanced degree (other than econ)	--	--	0.439* (0.184)
Constants			
Cut 1	7.801** (0.776)	6.412** (2.056)	7.236** (1.045)
Cut 2	9.628** (0.796)	8.456** (2.063)	9.135** (1.059)
Cut 3	11.32** (0.821)	10.24** (2.076)	10.83** (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)

<u>Variable</u>	<u>Unit of change (SD, or 1 for dummy variables)</u>	<u>Effect on outcome</u>
<u>Student characteristics</u>		
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
<u>Teacher characteristics</u>		
Years teaching econ	7.89	0.294
Undergrad econ (major or minor)	1	1.231
Advanced degree (other than econ)	1	1.100

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

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