

An Empirical Analysis of Relationship between Teacher's Quality and Student's Achievement

Fang Li

University of International Business and Economics
Beijing, China

Qiang Feng

University of International Business and Economics
Beijing, China

Chaofeng Zhang*

University of International Business and Economics
Beijing; Sichuan University of Arts and Science, Dazhou,

*Corresponding author

Abstract—This paper focus on the analysis of effect of teacher's characteristics on students' academic development based on the real statistical data of a university in China. Using the fixed effect panel data model, the empirical results show that the impact of full-time teachers on student's long term academic performance is better than that on the short term performance while adjunct instructors have opposite effect. Teachers' attribute (i.e., full-time teachers or adjunct teachers) statistically shows significant impact on students' academic results, but the existing correlation between teachers' evaluation and teachers' value-added effect is weak.

Keywords — university teacher; students' achievement; panel data; fixed effect

I. INTRODUCTION

Chinese colleges underwent a rapid expansion of their student intake from 1999 to 2006. The central and local departments of education are actively responding to this problem, as well as all the colleges and universities, which carry out a series of measures including ceasing enrollment expansion, increasing relevant investment and recruiting more teachers.

While increasing the workload of their existing teachers and enlarging the class size, many schools are forced to employ adjunct teachers from other institutions and encourage the postgraduate students to teach undergraduate courses. Some schools even include teaching duties in the curriculum of their postgraduate programs. As these measures become conventional, an instructor group consisting of full-time teachers, adjunct teachers, and postgraduate students has evolved and aroused public attention.

The diversification of instructor group in colleges also exists in other countries. It is reported by the NCES (National Center for Education Statistics) that more and more universities began to hire adjunct instructors since the late 1980s. Adjunct instructors made up 44% of the college staff in 2003. It has long been a concerned issue whether to employ adjunct instructors, but there is clear conclusion. It is criticized that that adjunct instructors generally do not have Ph.Ds, resulting in the deterioration of didactical quality and an increase in drop-out rate (Bettinger & Long, 2006; Ehrenberg & Zhang, 2005). However, the supporters argue that the experience in private sectors and professional teaching work enable the adjunct

instructors to enhance the students' learning experience while the full time professors can concentrate more on their academic research and improve the productivity of the school (Leslie&Gappa,1995; Autor, 2001).

Will the teachers hired from different channels have different impacts on the students' academic performance? Are there any difference between the teaching effects of teachers with superior titles and that of the young teachers represented by instructors? Answering these questions is of great value for optimizing the structure of college teachers and for improving the quality and efficiency of higher education. Nevertheless, in China, the empirical study on these problems is still blank.

This paper investigates the different effects of different teacher characteristics on students' academic performance. We consider the students taught by college teachers that are hired through different ways (i.e. full time, adjunct, and student instructor) as well as the teachers with different titles (i.e. full professor, associate professor and assistant professor). We analyze a data set of English language education from a Chinese university and examine the effects of teachers' characteristics on students' academic performance. The English teachers are categorized into four types: domestic full-time teachers, postgraduate student instructors, adjunct teachers and foreign teachers. In China, the main difference among these teachers is their personnel relationship with the school. The personnel relationship of full-times teachers belongs to the school itself, which ensures the stability of their employment. They cannot be fired unless in extreme cases. While other types of teachers vary in their initial employment. Teachers selected from postgraduate students are those with excellent grades and outstanding academic performance. The adjunct teachers are often recommended to the school by full-time teachers and are assigned tasks after interview. As for the foreign teachers, they are brought in by both the teaching department and school's department of international cooperation. In this paper, we only focus on the first three types, since foreign teachers only make up a small proportion of the staff composition, thus the effects cannot be accurately estimated. Among these three groups, full-time teachers are paid according to their titles and are provided with salaries and allowances from the government and school. The postgraduates are paid monthly based on their teaching hours. The hourly wage of adjunct teachers depends on their teaching experience, degree and rank. Notice that the last two

types of teachers will be assessed by their previous performance in class to decide whether they employed in the next academic year. Standards of teaching performance may include students' course grades and their appraisals towards the class.

Combined with all the factors, this paper establishes the analysis of different effects on students' academic development from full-time teachers and part-time teachers, based on the students' final test scores of the English courses and their evaluation results for teachers from a proper evaluation index.

II. LITERATURE REVIEW

Along with the prevalence of the educational accountability system in the western countries, the education improvement model has been applied to policy analysis and become a vital instrument for educational accountability, providing an objective and normative standard. Some scholars believe that the accountability system based on this model had a clear positive impact on students' achievements (Hanushek and Raymond, 2005). However, opposite opinion is that this model cannot appraise the real effects of teachers and universities, and wrong appraisals bring negative stimulus to the teachers (Gorard, 2006). The reliability of the education improvement model faces challenges from three aspects. First, the model setting is questionable. For instance, in the North Carolina and South Carolina model, how much progress the students should make is calculated by how they performed before and how much the school devoted in and school's performance is measured by the real progress minus the expected value. This kind of setting is inclined to underestimate the schools that have more good students, narrowing the improvement difference. Currently, fixed effects models with school dummy or teacher dummy have become the mainstream of education improvement model settings. But the specific form remains controversial (Tekwe, 2004). The second question results from measurement error. Measurement error of the dependent variable doesn't influence the estimation consistency, but measurement error of the independent variable does when it is correlated with the observations. Comparing the results of IV estimation (scores in two terms before as the instrument variable of that in last term) with that from the OLS estimation, Ladd and Walsh (2003) concludes that the existence of measurement error enlarges the difference in education improvement and is the most important source of bias. Though, student's test score fluctuates around its real academic level, those who performed outstanding in last exam are more likely to get a lower grade next time, i.e. the mean reversion phenomenon. Therefore, the discrepancy between the IV estimation using scores in two terms before and the OLS estimation is not entirely resulted from measurement error, but also from the impact of mean reversion.

Within the education improvement analysis framework, some consensuses of the influence of teachers have been formed. Volatility of teachers' influence has been verified, no matter based on random interference experiments or observation studies, during higher education or preschool education, in developed or developing countries. The one-year-

later volatility achieves as high as 50%-80% (Currie and Thomas, 1995; Jacob et, 2008; Kane and Staiger, 2008; Rothstein, 2008; Banerjee et, 2007). Improvement effects among teachers vary greatly, but few could be interpreted by dominant traits such as working experience and educational degree of the teachers (Rivkin et. 2005;Noell and Burns, 2006;Heck, 2007). Various studies point to a common conclusion: it is the class activity itself that actually counts.

Education improvement model is like a black box. Though it provides increasingly reliable information about which school or which teacher is effective during the educational process, it tells nothing about "why". Some researchers involve themselves into the class activities, trying to open the black box of education. A research on teacher appraisal carried out in Ohio, United States shows that teachers can't induce study behavior directly, but enable students to make educational improvement through creating learning time, encouraging atmosphere and participation in valuable activities for them (Lasley et., 2006). Heck (2007) also argues that active social relationship between students and teachers, supports from teachers, and teachers attitudes and beliefs all play crucial roles in students' scores. The researches above are based on long term careful observations in classes. While many of the conclusions are fairly insightful and persuasive, better recognition of causal relationship is still needed.

III. DATA STATEMENT AND ANALYSIS

Our data set is collected from a Chinese university. It includes information of students' enrollment, entrance English test scores, all types of English classes and their teachers, students' appraisal and CET-4/6 scores (standardized English tests). Due to the availability of the data, we only select the scores of English courses in every semester during a sample period from 2005 to 2008. The information of students' college reading and oral test scores and the information of the corresponding instructors are included. There are nearly 40,000 observations in the original sample, but due to massive missing data, we only keep the individual observations with complete data. Each individual student has experienced several English courses for four semesters, and for each semester (SEM for short), there are college English class and oral English class with 1306, 1207 students in it respectively. In the end of the semester, students are required to evaluate their teachers anonymously on 10 different indicators that represent the teachers' quality. We adopt students' test scores as dependent variable, and teachers' characters like degree, title, and whether being part-time as independent variables. Time invariant control variables of students' characters include gender, nationality, and original English level (reflected by their English entrance test scores). Descriptive statistics of the data related to college English class test scores are presented in Table 1.

TABLE 1. SUMMARY STATISTICS

Year	SEM	Mean	St.d	Min	Max	Part-time (%)	Average Evaluation	
							Part-time	Full-time
2005	1	78.93	6.91	60	93	52.39	89.97	90.82
	2	78.82	6.47	53	96	34.64	89.68	90.51
	3	77.57	6.15	60	92	21.16	91.52	92.15
	4	78.15	6.43	60	94	21.50	87.94	87.82
2006	1	78.92	6.52	47	94	33.09	88.65	88.74
	2	77.65	6.47	51	94	21.94	85.41	86.5
	3	76.16	6.07	52	94	14.75	90.75	89.28
	4	75.74	5.91	53	92	13.49	90.37	91.52
2007	1	76.72	8.21	57	91	42.01	90.18	91.04
	2	76.06	6.97	58	93	32.88	87.83	87.96
	3	75.09	6.42	53	91	18.72	91.51	93.31
	4	75.21	7.13	42	91	25.11	90.59	89.36
2008	1	79.19	7.55	60	94	64.58	91.44	91.58
	2	78.87	6.87	57	94	40	88.4	89.62
	3	78.47	6.63	60	92	27.5	90.34	91.46
	4	78.4	6.79	59	95	6.67	91.67	91.88

Table 1 shows that the percentage of part-time teachers has a positive relationship with students' test scores, yet students' evaluations towards them are comparatively lower. There four semesters in total, including two classes in each term, namely, college English and oral English.

We intend to estimate our model with panel data so as to control the unobservable individual heterogeneity and to describe and analyze the dynamic adjustment process, dealing with the error terms. Two common models of panel data are fixed effect model and random effect model. In fixed effect model individual effect correlates with other explanatory variables, which is not necessary in random effect model and omitting variables in the model specification will lead to inconsistency of parameter estimation. Therefore, we can select a proper model through Hausman test to see whether the fixed effect is correlated with other explanatory variables.

IV. EMPIRICAL RESULTS AND ANALYSIS

A. Econometric Model

In order to distinguish the impacts of teacher's category on students' test scores, we need to control some additional influential factors of students, such as gender and original

TABLE 2. ESTIMATION RESULTS OF COLLEGE ENGLISH CLASS MODEL

	1	2	3	4	5	6	7	8	9	10
adjunct	-0.688*** (0.182)			-0.366* (0.176)			-0.826*** (0.202)	-0.564** (0.185)		
degree		-0.285** (0.0916)			-0.0622 (0.0852)				-0.198* (0.0932)	
rank			-0.114 (0.0724)			0.0788 (0.0723)				-0.0634 (0.0804)
term	-2.529*** (0.390)	-2.596*** (0.388)	-2.688*** (0.386)	-2.679*** (0.300)	-2.749*** (0.297)	-2.786*** (0.296)	-2.522*** (0.395)	-2.654*** (0.305)	-2.710*** (0.304)	-2.756*** (0.302)
termsq	0.410*** (0.0773)	0.420*** (0.0770)	0.433*** (0.0770)	0.437*** (0.0603)	0.446*** (0.0601)	0.448*** (0.0602)	0.401*** (0.0778)	0.427*** (0.0608)	0.434*** (0.0607)	0.439*** (0.0606)
C	40.25*** (0.994)	40.23*** (0.996)	40.18*** (1.002)	81.31*** (0.316)	81.24*** (0.320)	81.11*** (0.321)	44.12*** (4.407)	85.08*** (4.349)	86.07*** (4.411)	85.51*** (4.344)
N	5224	5224	5224	5224	5224	5224	5224	5224	5224	5224
R2	0.392	0.392	0.391				0.399			

(standard error * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

In parameter estimation, time invariant variables like gender, nationality, and original English level are controlled. Model 1 specifies whether full-time teachers are influential to students' college English test scores. Model 2 specifies how teacher's educational degree impacts students' college English test scores. And model 3 specifies how teacher's title impacts

English level. The latter can be represented by their English entrance test scores. In our econometric model, the explained variable is $Score_{kit}$, ($k=1,2$), where $k=1$ denotes the test score of college English class of student i in semester t , and $k=2$ denotes the test score of oral English class of student i in semester t . The explanatory variable $Adjunct_{kit}$ is a dummy variable which takes the value of 0 if the English teacher of student i in semester t is part-time, or takes 1 if it is full-time. Explained variable $Degree_{kit}$ takes value from 0, 1, 2, 3, which takes the value of 0 if the English teacher of student i in semester t is part-time, 1 if it is full-time with bachelor degree, 2 if it is full-time with master degree, or 3 if it is full-time with doctor degree. Explained variable $Rank_{kit}$ also takes value from 0, 1, 2, 3, which takes the value of 0 if the English teacher of student i in semester t is part-time, 1 if it is full-time with instructor title, 2 if it is full-time with associate professor title, or 3 if it is full-time with professor title. We also consider how students' evaluation to the teachers impacts their scores. Because of the anonymous principle, we are unable to match each student to the evaluation they make. Therefore, we take the average evaluation in the same semester for a teacher of all the students enrolled in the same year. Our results show that current students' evaluation results possess weak relationship with teachers' improvement. Due to limitations on space, we don't show the results of students' evaluation.

In this paper we construct the econometric model below:

$$Score_{kit} = \alpha + \delta_i + \gamma_t + \beta Adjunct_{kit} + \phi term_{kit} + \phi termsq_{kit} + \sum_{j=1}^{10} \theta_j Ev_j + \varepsilon_{kit} \quad (1)$$

Where δ_i is fixed-effect, γ_t is time-effect of specific term, and term is which term student's class is in, reflecting the difference of classes among these semesters. Meanwhile, the square of term is also included in the model, denoted as $termsq$, for every student has observations in four semesters, and thus we also give cluster standard error in parameter estimation in order to control inner group correlation.

B. Regression Results of College English Class

students' college English test scores. Since the value of part-time teacher of these variables is 0, we only need to consider one variable in each regression model to find the influence of part-time teachers. These three models, all using mixed-regression, control for heteroskedasticity. Model 4-6 reflect the results of fixed effect model regressions. Because of the

existence of heteroskedasticity and auto-correlation, we estimate the model clustered by student ID to obtain a more robust standard deviation. Model 7-10 discuss the relationship between students' evaluation of the teachers and their test scores, where model 7 is pooled regression, and model 8-10 consider the influence from full-time teachers, educational degree and title of the teachers respectively.

Results show that full-time teachers have negative impacts on students' test scores, which is even worse as their degree increases. But the impact of title is not significant. The regression result is relatively robust as seen from clustered standard deviation. Hausman test shows that the chi-square

statistic is far larger than the critical value at 5% significant level. So it is suitable to choose fixed effect models, and results of random effect models are omitted due to limited space. Furthermore, among the evaluations presented in table 1, the average evaluation to part-time teachers is comparatively lower. From the estimation results, however, there are not any obvious relationships between evaluations to teachers and students' test scores. So we neglect these effects in the table. Nevertheless, in average, students hold a higher degree of recognition towards full-time teachers.

C. Regression Results of oral English Class

TABLE 3. ESTIMATION RESULTS OF ORAL ENGLISH CLASS MODEL

	1	2	3	4	5	6	7	8	9	10
Adjunct	-0.487*** (0.121)			-0.474*** (0.116)			-0.361* (0.158)	-0.330* (0.147)		
Degree		-0.192** (0.0575)			-0.197*** (0.0557)				-0.146* (0.0696)	
Rank			-0.169** (0.0548)			-0.153** (0.0537)				-0.0503 (0.0663)
term	-0.878*** (0.303)	-0.873*** (0.303)	-0.890*** (0.303)	-0.887*** (0.256)	-0.882*** (0.256)	-0.897*** (0.256)	-0.919*** (0.304)	-0.910*** (0.265)	-0.905*** (0.265)	-0.906*** (0.265)
termsq	0.269*** (0.0606)	0.267*** (0.0606)	0.267*** (0.0606)	0.271*** (0.0508)	0.270*** (0.0509)	0.268*** (0.0508)	0.258*** (0.298)	0.262*** (0.129)	0.260*** (0.139)	0.257*** (0.130)
C	67.53*** (0.846)	67.46*** (0.848)	67.47*** (0.845)	67.67*** (1.024)	67.61*** (1.026)	67.60*** (1.024)	73.21*** (1.744)	90.89*** (1.795)	90.75*** (1.791)	90.58*** (1.790)
N	4728	4728	4728	4728	4728	4728	4728	4728	4728	4728
R2	0.169	0.168	0.168				0.192	0.071	0.070	0.069

(standard error * p < 0.05, ** p < 0.01, *** p < 0.001)

Model 1 specifies whether full-time teachers are influential to students' oral English test scores. Model 2 specifies how teacher's educational degree impacts students' oral English test scores. And model 3 specifies how teacher's title impacts students' oral English test scores. Since the value of part-time teacher of these variables is 0, we only need to consider one variable in each regression model to find the influence of part-time teachers. These three models, all using pooled regression, control for heteroskedasticity. Model 4-6 reflect the results of fixed effect model regressions. Because of the existence of heteroskedasticity and autocorrelation, we estimate the model clustered by student ID to obtain a more robust standard deviation. Model 7-10 discuss the relationship between students' evaluation of the teachers and their test scores, where model 7 is pooled regression, and model 8-10 consider the influence from full-time teachers, educational degree and title of the teachers respectively. Still, results show that full-time teachers have negative impacts on students' oral English test scores, which is even worse as their degree increases. Variable Rank is not significant, which is similar to the results of college English class.

V. INFLUENCE OF FULL-TIME TEACHERS ON STUDENTS' MEDIUM AND LONG TERM ENGLISH IMPROVEMENT

We take the test score of college English (I) in the first semester of a freshman as his original English level, CET-4 score as his medium term progress and CET-6 score as his long term progress. Calculate the score percentiles of college English (I), CET-4 and CET-6 for every student, and then take the difference of score percentile of college English (I) and CET-4 as student's medium term progress, and of CET-4 and CET-6 as long term progress. Meanwhile, we calculate the evaluation of full-time teachers for every student in all four

terms. If this indicator is greater than 5, then the full-time dummy variable takes 1 or 0 otherwise. The estimation results show that being full-time leads to 8 percentiles improvement in student's medium term progress, and 4.3 percentiles in long term progress. In addition, we figure out the relationship between evaluation of full-time teachers and the average level of students' medium-long term progress, plotted in the graphs below:

TABLE 4. NUMBER OF FULL-TIME TEACHERS VS. MEDIUM & LONG TERM PROGRESS

Evaluation to full-time teachers	Medium term progress		Evaluation to full-time teachers	Long term progress	
	Average level	Standard deviation		Average level	Standard deviation
1	-0.105	0.417	1	-0.084	0.325
2	-0.069	0.246	2	-0.029	0.355
3	-0.078	0.293	3	-0.059	0.241
4	-0.042	0.279	4	-0.006	0.283
5	-0.019	0.271	5	-0.003	0.242
6	0.018	0.251	6	0.002	0.259
7	0.033	0.259	7	0.025	0.275
8	0.047	0.283	8	0.015	0.273

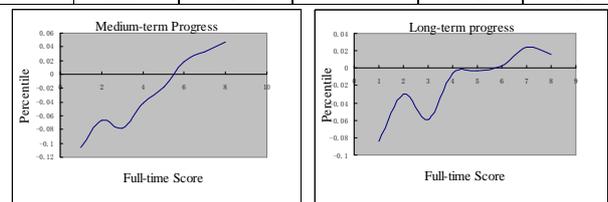


FIGURE 1. RELATIONSHIP BETWEEN NUMBER OF FULL-TIME TEACHERS AND STUDENT'S MEDIUM/LONG TERM PROGRESS IN ENGLISH

VI. CONCLUSION

The influence of full-time teachers plays a better role in students' long term progress than in short term. In the contrary, part-time teachers impact short term performance more. To further understand this result, one should be equipped with the knowledge about our current external environment and internal incentive system of higher education in China. Besides age, title and degree, the main difference between full-time teachers and part-time teachers lies in the personnel relationship with universities. In the practical process of teaching management, the personnel relationship of full-times teachers belongs to the school itself, which ensures the stability of the teachers so they won't be fired in common cases. External teachers vary in the initial employment. Generally, they are introduced to the school by full-time teachers and are assigned tasks after interview. Full-time teachers are paid according to their titles and are provided with salaries and allowances from the government and school. And external teachers, taken series of factors like experience, degree, and title into consideration, obtain hourly wage every month. Notice that external teachers will be assessed by their previous performance in class to decide whether they can be continually employed. Standards of teaching performance may include students' course grades and their appraisals towards the class. In view of this, part-time teachers must be more inclined to put efforts in improving students' short term progress and class evaluation.

The current students' evaluation results are weakly related to teachers' improvement evaluation. Both objective and subjective aspects are necessary to measure students' development comprehensively. In the objective aspect, students obtaining knowledge and skills in class can be measured by test scores. The performance of students can reflect the teaching effects. In the subjective aspect, learning experience can be measured by students' evaluation. In general, higher evaluation implies better learning experience the teacher has brought to the student. Therefore, test scores and student's evaluation form a relatively complete system to measure students' academic performance in tow aspects. Meanwhile, since factors of teachers affect students' academic performance significantly, these two aspects are also crucial dimensions to judge teacher's teaching ability, and effective indices to measure its influence to students' academic performance.

Owing to the anonymity of students' evaluation results, which are conducted via internet platforms, external interference is effectively avoided. As a result, it is popular to use students' evaluation as the indicator for class quality of teachers in current higher education. Moreover, in practical teaching management and supervisory affairs, many universities regard it as a vital reference of teacher appraisal, reward and job classification. However, it remains for argument whether students' evaluation reflects teaching abilities of teachers in a faithful and comprehensive way. Our research finds that current students' evaluation results possess

weak relationship with teachers' improvement. Indications like requirement and class management are more useful to measure the long term impacts of teachers on students' academic achievements; while indicators like teaching materials work better with the short term impacts.

Student's academic achievement evaluation and student's teaching evaluation complement each other, reference each other, and are not alternative. Through the above analysis we find education improvement evaluation based on students' academic achievement and teaching evaluation based on subjective judgments of students appear an increasing pattern as a whole, also mapping the indicators and time periods. So these two evaluation methods are complementary, having their own characteristics and emphasis. So, a comprehensive and effective teaching evaluation system should incorporate these two measures at the same time, with both quantitative and qualitative evaluations included.

REFERENCES

- [1] Altonji, J., et al.. Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, vol. 113, 2005.
- [2] Bettinger, E. P. and B. T. Long. Does Cheaper Mean Better? The Impact of Using Adjunct Instructors on Student Outcomes. *The Review of Economics and Statistics*, vol. 92, 2010, pp. 598-613.
- [3] Ehrenberg, R. G. and L. Zhang. Do tenured and tenure-track faculty matter?" *Journal of Human Resources*, vol. 40, 2005, pp.647-659.
- [4] Gorard, S. . Value-added is of little value. *Journal of Education Policy*, vol. 21, 2006, pp. 235-243.
- [5] Hanushek, E. A. and M. E. Raymond . Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, vol.24, 2005, pp. 297-327.
- [6] Heck, R. H. . Examining the relationship between teacher quality as an organizational property of schools and students' achievement and growth rates. *Educational Administration Quarterly*, vol.43, 2007, pp.399-432.
- [7] Jacob, B. A., et al.. The persistence of teacher-induced learning gains, *National Bureau of Economic Research*, 2008.
- [8] Kane, T. J. and D. O. Staiger. Estimating teacher impacts on student achievement: An experimental evaluation, *National Bureau of Economic Research*, 2008.
- [9] Ladd, H. F. and R. P. Walsh. Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, vol. 21, 2002, pp.1-17.
- [10] Noell, G. H. and J. L. Burns. Value-Added Assessment of Teacher Preparation An Illustration of Emerging Technology. *Journal of Teacher Education*, vol. 57, 2006, pp.37-50.
- [11] Rivkin, S. G., et al.. Teachers, schools, and academic achievement. *Econometrica*, vol.73, 2005, pp. 417-458.
- [12] Rothstein, J.. Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, vol.125, 2010, pp.175-214.
- [13] Tekwe, C. D., et al.. An empirical comparison of statistical models for value-added assessment of school performance. *Journal of educational and behavioral statistics*, vol. 29,2004, pp.11-36.