The Impact Of School Socioeconomic Status On Student-Generated Teacher Ratings

Steve Agnew, University of Canterbury, New Zealand

ABSTRACT

This paper uses ordinary least squares, logit and probit regressions, along with chi-square analysis applied to nationwide data from the New Zealand ratemyteacher website to establish if there is any correlation between student ratings of their teachers and the socioeconomic status of the school the students attend. The results show that students from mid socioeconomic status schools score their teachers significantly higher than students from other socioeconomic status schools. This has implications for the national measurement of teacher performance, as well as implications for individual teachers working in different socioeconomic status schools.

Keywords: Teacher Ratings; Socioeconomic Status; Teacher Quality; Teacher Effectiveness

INTRODUCTION

T he objective of this work is to establish if there is a correlation between the socioeconomic status (SES) of the school a student attends and student ratings of their teachers. In New Zealand, the measurement used to represent the socioeconomic status (SES) of a school is the school's decile, which is calculated by the Ministry of Education. All state schools in New Zealand are given a decile rating which represents the SES of the school's catchment area. According to the New Zealand Ministry of Education website (Ministry of Education, 2010), in order to assess a school's decile rating five factors are used to measure the socioeconomic standing of its community: 1) household income, 2) occupation, 3) household crowding, 4) educational qualifications, and 5) income support (the percentage of parents who receive a benefit). Decile 1 schools are the 10% of schools with the highest proportion of students from low socioeconomic communities, whereas decile 10 schools are the 10% of schools with the lowest proportion of these students. A low decile school would therefore draw on communities where there is a greater density of households that have low incomes, manual occupations, greater household crowding, lower educational qualifications and greater dependence on income support.

Ali, McWhirter and Chronister (2005) state that socioeconomic status influences the educational and occupational opportunities available to individuals and the attainment they achieve. They base this assertion on research carried out by Brown (2000), Fouad & Brown (2000), Gilbert & Kahl (1993) and Turner & Lapan (2003). Ali et al (2005) go on to state that socioeconomic status (SES) "influences the way individuals perceive their opportunities and affects their access to educational and vocational resources" (p. 41 & 42). They also continue that

Social Cognitive Career Theory (SCCT) outlines the ways in which personal factors, such as SES, interact with contextual factors (e.g. social support) to influence the development of career interests, the selection of career goals, and career behaviours….For example, an adolescent from a lower SES background is more likely to have poorer quality schooling, fewer career role models, and less financial support for post-secondary options than higher SES adolescents. (p. 41 & 42).

Based on the findings of Ali, McWhirter and Chronister (2005), we would expect lower socioeconomic (SES) students to have poorer schooling, and thus give their teachers lower ratings. There is certainly evidence to

suggest that lower decile schools in New Zealand are harder to staff and have a greater proportion of inexperienced teachers teaching in them. The document *Monitoring Teacher Supply* (Ministry of Education, 2006) found, based on a survey of principals at the start of 2006:

In 2006, there was again a clear pattern evident by school decile, as in previous years. The proportion of vacancies and re-advertised positions to staffing entitlement was highest in low decile schools (deciles 1 – 3) and lowest in high decile schools (8 – 10) (p. 13).

In 2006, 15% of low decile (1-3) schools had a vacancy compared to 13.5% of medium decile (4-6) schools and 11.8% of high decile (8-10) schools (Ministry of Education, 2006). When the same comparison is made for re-advertised positions (any position which has been advertised nationally more than once with no appointment being made after the first advertisement), there is a similar result with 6.1% of low decile schools re-advertising compared to 4.7 and 2.8% for medium and high decile school, respectively. This equated to 47.8 positions in low decile schools, 45.7 positions in medium decile schools, and 22.2 positions in high decile schools (Ministry of Education, 2006). For the period 2003 to 2006, schools covered 43% of their vacancies through trained relief staff and 11.9% of vacancies with staff that had limited authority to teach (Ministry of Education, 2006). This suggests that students in low decile schools have a higher chance of being taught by a relief teacher or a teacher with a limited authority to teach. Low decile schools also had a greater proportion of beginning teachers in 2006, with 5.8% being beginning teachers, compared to 4.9% in medium decile schools and 4.5% in high decile schools (Ministry of Education, 2006).

However, McKeachie (1979) makes a pertinent point when he states, "Probably the single most important student variable affecting satisfaction is student expectations. Students who expect a course or teacher to be good, generally find it to be so" (p. 390). Further studies that support the notion of student expectations having an impact on teacher evaluations include Perry, Niemi and Jones (1974), who found that prior student expectations of teaching performance based on instructor reputation, had an influence on student ratings. A similar study by Leventhal, Abrami and Perry (1976) found that students who chose a class section based on the reputation of the instructor gave higher ratings to their teachers than did their classmates. Using multilevel regression, Griffin (2001) found that students who heard positive information regarding the instructor's reputation rated the instructor higher than students who heard negative information about the instructor. Gigliotti (1987) found that negative expectancy violations resulted in unfavourable course evaluations. Gigliotti and Buchtel (1988) found that student evaluations of college courses and instructors are significantly determined by whether outcomes correspond to expectancy. They also found that students' feelings of satisfaction with a course and with themselves are influenced by the same process.

Relating this to SES, Chapman and Lawes (1984) report that "studies of low achieving or failure-prone individuals have reported the tendency for success to be seen as externally caused and failure as internally caused" (p. 177). According to New Zealand Qualifications Authority (NZQA) released data from 2005 to 2007, at levels one, two and three, students at high decile schools acquire NCEA at a much higher rate than students at mid-decile schools, who acquire it at a much higher rate than students at low decile schools (NZQA, 2009). Students from low decile schools have lower rates of success in obtaining NCEA, and if the findings of Chapman and Lawes (1984) are correct, are more likely to attribute less than expected performance to themselves as opposed to the teacher. Conversely, students from high decile schools may have higher expectations that their education will be of a higher quality, but be more likely to attribute less-than-expected academic performance to their teacher. This raises the interesting question "Is there a common correlation between the SES of a school and how the students attending that school rate their teachers?" The statistics from the Ministry of Education referred to earlier in the introduction suggest that students from high SES schools are less likely to experience a lower quality of teaching due to the lower number of relief teachers and teachers with a limited authority to teach, when compared to schools of lower SES. The existing research also suggests that if the students from high SES schools expect their education to be of a good quality, they will rate their teachers accordingly. Both of these factors suggest that students from high decile schools will rate their teachers higher than students from low decile schools. However, there is also some literature that suggests high SES students are more likely to attribute failure to the teacher, where low SES students are more likely to internalize failure. To ascertain if there is a correlation between the SES of the school a student attends, and student ratings of their teachers, the materials and methods outlined in the Material and Methods section were used.

MATERIAL AND METHODS

Data were collected from the ratemyteacher.co.nz website between December 15 and December 22, 2006. This was timed to be after the school year had finished and examinations had been completed, but before examination results were released in mid-January. The ratemyteachers.co.nz website was introduced in New Zealand in 2006 and is an expansion of the www.ratemyteachers.com site based in the USA. This study concentrates on the NZ site.

People logging onto the site are able to rate a teacher on three characteristics: 1) easiness of the subject, 2) clarity of the teacher, and 3) helpfulness of the teacher. Each characteristic is rated on a scale of 1 to 5, with one being the lowest and 5 the highest. These three numbers are averaged to give the teacher an average quality rating out of 5. As at January 7, 2007, there had been 142,404 ratings. Some 26,345 teachers have been rated (although this figure is slightly overstated as some teachers are listed twice in some schools and some teachers have been listed, but not rated) and 1,896 schools were listed on the site (although some schools had no teachers rated).

This study focuses on composite of year 7 to 13 and year 9 to 13 schools. The mean teacher rating was recorded for each school from the ratemyteacher website. The number of ratings was compared to the official school roll for each school, and schools for which the number of ratings was less than 10% of their school roll were removed from the study. The purpose of this was to remove schools that had few responses, where the average teacher rating may have been unduly affected by "rogue" ratings, whilst at the same time, not discriminating against small schools. The average teacher rating for each school also had to be adjusted where a teacher had been listed but no rating had been given. This was recorded as a zero rating on the ratemyteacher site, and thus brought the average quality rating down. These zero ratings were removed, and the averages adjusted accordingly, for the purpose of this study. In some schools, the same teacher was listed and rated more than once. This will have had a very minor effect on some of the average teacher quality ratings for certain teachers; however, there is no reason to believe they are dominant in any particular decile group, and thus should not have distorted the relative rankings.

To ensure there were a sufficient number of samples in each region, schools have been grouped into one of four groupings, according to their decile rating. The four groupings are:

- Deciles 1, 2 and 3
- Deciles 4, 5 and 6
- Deciles 7, 8 and 9
- Decile 10 and private schools

The smallest sample size of any decile grouping was 28 schools, which represents many hundreds of individual teacher ratings. The overall sample size was 269 schools. Decile 10 and private schools were given their own grouping as they represented the highest SES school grouping and schools which caregivers and parents have been required to pay fees to attend. The majority of private schools are likely to have been categorized as decile 10 had they been state schools. They are also likely to have similar expectations by students of the quality of education they will receive as decile 10 school students.

The mean teacher ratings, by school, that were calculated for each decile grouping allowed for comparison across several different regions. Ordinary Least Squares (OLS) regressions were run to establish if any differences in the means between decile groupings were statistically significant. The regression model had the following form:

Mean School Teacher Rating = $\alpha + \beta_1$ High Decile + β_2 Urban + β_3 North Island + ε

where:

α = intercept
High Decile = the dummy variable for the highest decile grouping in that regression
City = the dummy variable for schools identified by the Ministry of Education as having a city local territorial authority (1=City; 0=Districts)

North Island = the dummy variable for the island each school was located in (1=North Island; 0=South Island)
ε= the error term

The city and North Island variables were included to see if the geographic location of the school had a significant impact on the mean school teacher rating.

To enable the use of contingency tables and the calculation of odds ratios, the data were then recoded. Rather than using the mean teacher rating for each school, the national mean teacher rating was calculated. Schools were then given a dummy variable of 1 or 0, depending on whether their mean school teacher rating was above or below the national mean. Logit and probit regressions were then run, with the probit regressions reporting changes in probabilities and the logit regressions reporting odds ratios. The changes in probabilities reported by the probit regressions were to be the same as the OLS coefficients and represented the risk differences when calculating effect size. A chi-square analysis was also performed on the data to confirm levels of significance.

RESULTS AND DISCUSSION

The national mean and median teacher ratings by school were both 3.67(2dp). Table 1 shows the mean school teacher rating calculated across various geographic regions.

Table 1: Mean Teacher Ratings by Decile Grouping and Region

Decile Grouping	NZ	NZ City	NZ Districts	North Island	South Island	Nth Island Districts	Nth Island City	Sth Island Districts	Sth Island City
1, 2 & 3	3.618	3.522	3.681	3.630	3.584	3.723	3.442	3.481	3.662
4, 5 & 6	3.739	3.795	3.695	3.739	3.738	3.694	3.806	3.698	3.777
7, 8 & 9	3.651	3.671	3.631	3.688	3.585	3.675	3.702	3.552	3.617
10 & 99	3.628	3.653	3.446	3.642	3.574	3.482	3.659	3.376	3.624

Column one shows the mean school teacher rating for the entire sample. It is this data that are used in the statistical analysis to test for significant correlations. The difference in the mean scores for the different decile groupings is not large. There are larger differences between decile groupings in some of the sub-samples; however, smaller sample sizes in some of these groupings means that tests of significance are not appropriate. An interesting pattern to emerge is the decile 4, 5 & 6 grouping scoring the highest rating, and the decile 10 and private schools scoring the lowest or second lowest rating, which is consistent across most of the data set. This is highlighted in Table 2 where the rank order of each decile grouping means teacher rating is shown.

Table 2: Relative Rankings of Teacher Ratings

Decile Grouping	NZ	NZ City	NZ Districts	North Island	South Island	Nth Island Districts	Nth Island City	Sth Island Districts	Sth Island City
1, 2 & 3	4	4	2	4	3	1	4	3	2
4, 5 & 6	1	1	1	1	1	2	1	1	1
7, 8 & 9	2	2	3	2	2	3	2	2	4
10 & 99	3	3	4	3	4	4	3	4	3

With the exception of some payments paid to teachers working in hard-to-staff schools (traditionally low decile and/or remote rural schools), the teachers' salary scale is the same across all deciles in the state sector. Given that many private schools also pay a premium above the state scale to attract "quality" teachers, it could be expected that "better" teachers would be attracted to higher decile and private schools, where teaching conditions are perceived to be easier, and in some private schools, the pay is higher. Therefore, if ratemyteacher ratings are indicative of teacher quality, we could expect higher ratings in higher decile schools. Higher decile schools have certainly traditionally found it easier to fill staff vacancies than lower decile schools. A wider pool of applicants should theoretically increase the chances of hiring a "higher quality" teacher. The fact that the mean teacher rating for the decile 10 and private schools was consistently the lowest or third lowest is therefore somewhat surprising.

Based on the information from the Ministry of Education provided earlier, which stated that high decile schools had fewer re-advertised positions and so were less likely to need to employ relief staff, beginning teachers or teachers with limited authority to teach - the premise that the decile 10 and private schools find it easier to attract quality teachers and so should receive higher ratings - would be expected. However, the ratemyteacher site ratings do not reflect this. Not only does the decile 4, 5 & 6 group consistently receive the highest rating, it does so across every subset of the sample apart from one. The decile 4, 5 & 6 group teacher ratings also tend to fluctuate less than the other groups from region to region.

Decile groups 4, 5 & 6 and 7, 8 & 9 are consistently ranked first and second, respectively, across the different samples. Rather than representing actual teacher quality, it is conceivable that these teacher ratings are a reflection of where the student expectations of their teachers and the quality of teaching are most closely matched, although it must be noted that the decile 4, 5 & 6 grouping was the only one that displayed a significantly different mean rating. Despite the decile 1, 2 & 3 group performing better in the South Island, and scoring the highest ranking in the North Island District sample, when the North Island City scores are added, their nationwide rating comes out the lowest. Even though as low decile school students, their expectations are lower, these lower expectations appear to be still not being met as well as students in mid-decile schools. This is supported by the lower rates of school achievement and higher rates of exposure to relief teachers, beginning teachers, and limited authority to teach teachers. The fact that despite high rates of student achievement, the decile 10 and private schools group consistently received the lowest or second lowest ratings suggests that the high expectations of the students may not be being met to the same extent as mid-decile school students.

To establish if there were any significant differences between the mean teacher rating for each of the decile groupings, OLS regressions were run with the results shown in Table 3.

Table 3: OLS Regression Coefficients Showing the Importance of Decile and Location on Mean Teacher Ratings

	1 to 3 Dropped	4 to 6 Dropped	7 to 9 Dropped	10 and 99 Dropped
Deciles 1 to 3		-0.12**	-0.03	0.02
		(-2.01)	(-0.54)	(0.31)
Deciles 4 to 6	0.12**		0.09**	0.14***
	(2.01)		(2.19)	(2.74)
Deciles 7 to 9	0.03	-0.09**		0.05
	(0.54)	(-2.19)		(1.06)
Deciles 10 & 99	-0.02	-0.14***	-0.05	
	(-0.31)	(-2.72)	(-1.06)	
North Island	0.06	0.06	0.06	0.06
	(1.57)	(1.57)	(1.57)	(1.57)
City	0.06	0.06	0.06	0.06
	(1.58)	(1.58)	(1.58)	(1.58)

*** $p < 0.01$, ** $p<0.5$, * $p<.0.10$

Column one shows the regression coefficients when the decile 1, 2 & 3 grouping is dropped. This means that the ratings of the other decile groupings are compared back to the decile 1, 2 & 3 grouping. The decile 4, 5 & 6 grouping is the only one to show a significantly different rating when compared to the decile 1, 2 & 3 grouping. It is significant at the 95% confidence interval. The coefficient of 0.12 represents the difference between decile 4, 5 & 6 grouping rating of 3.739 and the decile 1, 2 & 3 decile grouping rating of 3.618 (allowing for rounding). Because the coefficient is positive, the decile 4, 5 & 6 grouping had a higher mean than the decile 1, 2 & 3 grouping it is being compared against. Column three shows the results when the decile 7, 8 & 9 grouping is dropped (compared against). Once again, the decile 4, 5 & 6 grouping was the only one to have a significantly different rating; and again, it was positive, meaning the decile 4, 5 & 6 grouping had a statistically significant higher rating. Column four shows the results when the decile 10 and private schools are dropped. The same pattern emerges with the decile 4, 5 & 6 grouping having a positive, statistically significant coefficient, meaning a statistically significant higher rating. Column two effectively shows a summary of the statistically significant results from the other columns. Ratings for the decile 4, 5 & 6 group are significantly higher than all other decile groupings. The biggest difference is the comparison with the decile 10 and private schools, followed by decile 1, 2 & 3 schools. Decile 7, 8 & 9 has the

smallest difference when compared with the decile 4, 5 & 6 grouping. There are no significant differences in ratings when any decile groupings other than the decile 4, 5 & 6 grouping are compared against each other. It should be noted that, on several occasions, the coefficients in Table 2 do not exactly match the differences in ratings shown in Table 1. This is because the coefficients in Table 2 are calculated given the effect of island and territorial local authority. Although the effect of which island and territorial local authority the school was in was not found to have a statistically significant effect, those two factors were still responsible for some effect on ratings.

After the dependant variable was recoded into a binary 1 or 0, depending on whether the mean school rating was above or below the national mean; ordinary least squares regressions, probit regressions reporting changes in probability and logit regressions reporting odds ratios were run. A Chi-square statistic was also calculated to test for levels of significance. The results are shown in Table 4.

Table 4: Regression and Chi-square Analysis Outputs

Ethnicity	OLS	χ2	Logit	Probit
1-3 v 4-6	0.07	(0.47)	1.35	0.07
	(0.68)	(1,N=120)	(0.69)	(0.69)
1-3 v 7-9	-0.07	(0.44)	0.75	-0.07
	(-0.66)	(1,N=127)	(-0.66)	(-0.66)
1-3 v 10-99	-0.14	(1.34)	0.58	-0.14
	(-1.15)	(1,N=78)	(-1.15)	(-1.15)
4-6 v 7-9	-0.14**	(3.98)**	0.56**	-0.14**
	(-2.00)	(1,N=191)	(-1.99)	(-1.99)
4-6 v 10-99	-0.21**	(5.67)**	0.43**	-0.21**
	(-2.41)	(1,N=142)	(-2.36)	(-2.37)
7-9 v 10-99	-0.06	(0.56)	0.77	-0.06
	(-0.75)	(1,N=149)	(-0.75)	(-0.75)

There is now no significant difference between the decile 1-3 and decile 4-7 groupings. However, the difference between how students from the decile 4-6 school grouping rate their teachers is still significantly different from both the decile 7-9 grouping and the 10-99 grouping. For the 4-6 and 7-9 decile groupings' comparison, the probit and OLS regressions show that 14% fewer decile 7-9 schools were ranked above the national mean for teacher ratings than schools from the decile 4-6 grouping. The odds ratio calculated using a logit regression shows that the odds of a school in the decile 7-9 grouping, having a mean teacher rating above the national mean, are 56% of the odds of a decile 4-6 school having a mean teacher rating above the national mean. For the comparison between the 4-6 and 10-99 groupings, there were 21% fewer decile 10-99 schools ranked above the national mean compared to schools from the decile 4-6 grouping. The odds of a decile 10-99 school being ranked above the national mean for teacher ratings are 43% of the odds of a decile 4-6 school having a mean teacher rating above the national mean. All of these results are significant at the 05% confidence level, which is confirmed by the calculation of the chi-square statistic.

There appear to be two possible reasons for this. The first is that the best teachers do teach in decile 4, 5 and 6 schools. Measuring this is beyond the scope of this study. However, intuitively there is no obvious reason as to why this would be the case in almost every region in NZ, especially when some private schools offer higher rates of pay and high decile schools have lower rates of re-advertised positions, suggesting they are easier to staff. The second possible reason is that rather than measuring teacher quality, the rate my teacher ratings are measuring teacher performance relative to student expectations of their teachers. Although students from high decile schools have the highest rates of NCEA achievement according to the NZQA data mentioned earlier, they are also more likely to have higher expectations of their teachers, as hypothesised by Ali et al (2005). This hypothesis is further substantiated by Chapman and Lawes (1984) as mentioned in the introduction, when they suggest that low achieving or failure-prone individuals have reported the tendency for success to be seen as externally caused and failure as internally caused (1984). High decile students may therefore be more likely to see success as being internally caused and failure to be externally caused.

CONCLUSIONS

There is a significant difference in how students from mid decile schools rate their teachers compared to students from low and high decile schools, when comparing mean student ratings of their teachers aggregated at the school level. Although the actual difference in mean teacher ratings between schools in different decile groupings is small, 2.4% less for the decile 7-9 grouping and 2.9% less for the 10-99 decile grouping when compared to decile 4-6 schools, they are significant at the 95% confidence level. When phrased in terms of the odds of a school gaining a mean teacher rating greater than the national mean, there are large and significant differences between decile 4-6 schools and schools from both the decile 7-9 grouping and the 10-99 grouping. The purpose of this paper is to establish if a correlation exists between the SES of a school and the teacher ratings of that school. This has been proven. An opportunity also exists for further research which attempts to establish the reasons why the established correlations exist. This paper is unique in New Zealand in that it is based on nationwide data, not data within an institution. This has allowed for comparison between deciles, which previous research on teachers within an institution has not been able to do.

The implications of the findings of this paper for the teaching profession are two-fold. Firstly, when teacher effectiveness is being measured, SES differences in student ratings, which are not due to teaching ability, need to be considered. Teacher evaluations need to be recognized as being a test of expected teaching, rather than a reflection of teacher quality. It may be, for instance, that a teacher currently working in a high SES school may in fact score a higher rating in a mid SES school, where student expectations may be lower. Secondly, when trainee teachers are placed in various schools on teaching practicum, they need to be made aware that different SES schools will have differing expectations and that student evaluations may well reflect this.

AUTHOR INFORMATION

Steve Agnew is currently employed as a Teaching Fellow at the University of Canterbury. Prior to his current position, he worked in New Zealand secondary schools for eleven years, predominantly as an economics teacher. His current research interests include the teaching of economics and financial literacy in secondary school, and teacher education.

REFERENCES

1.   Ali, S. R., McWhirter, E. H., & Chronister, K. M. (2005). Self-efficacy and vocational outcome expectations for adolescents of lower socioeconomic status: A pilot study. *Journal of Career Assessment,* 13(1), 40-58.
2.   Brown, M. T. (2000). Blueprint for the assessment of socio-structural influences in career choice and decision making. *Journal of Career Assessment,* 8, 371-378.
3.   Chapman, J. W., & Lawes, M. M. (1984). Consistency of casual attributions for expected and actual examination outcome: A study of the expectancy confirmation and egotism models. *British Journal of Educational Psychology,* 54(2), 177-188.
4.   Fouad, N. A. & Brown, M. T. (2000). Role of race and social class in development: Implications for counseling psychology. *Handbook of Counseling Psychology,* 379-408.
5.   Gigliotti, R. J. (1987). Are they getting what they expect? *Teaching Sociology,* 15 (Oct), 365-375.
6.   Gigliotti, R. J., & Buchtel, F. S. (1988). Attributional bias and course evaluations. *Journal of Educational Pyschology,* 82 (2), 341-351.
7.   Gilbert, D., & Kahl, J. A. (1993). *The American class structure: A new synthesis* (4th ed.). Belmont, CA: Wadsworth.
8.   Griffin, B. W. (2001). Instructor Reputation and Student Ratings of Instruction. *Contemporary Educational Psychology, 26, 534-552.*
9.   Leventhal, L., Abrami, P. C., & Perry, R.P. (1976). Do teacher rating forms reveal as much about students as about teachers? *Journal of Educational Psychology,* 68, 441-445.
10.   McKeachie, W. J. (1979). Student ratings of faculty: A reprise. *Academe,* 65, 384-397.
11.   Ministry of Education. (2006). Monitoring teacher supply: Survey of staffing in New Zealand schools at the beginning of the 2006 school year. Retrieved from http://www.educationcounts.govt.nz.

12.   New Zealand Qualifications Authority (2009). An analysis of student performance in NCEA, 2005-2007. Wellington, N. Z.

13.   Perry, R. P., Niemi, R.R., & Jones, K. (1974). Effect of prior teaching evaluations and lecture presentation on ratings of teaching performance. *Journal of Educational Psychology,* 66, 851-856.

14.   Turner, S. L., & Lapan, R. T. (2003). The measurement of career interests among at-risk inner-city and middle-class suburban adolescents. *Journal of Career Assessment,* 11, 405-420.

15.   Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education,* 23(2), 191-212.