Parent Characteristics, Teacher Qualification, Peer Effect, and Student Achievement in Third Grade: A Multiple-Group Mixture Structural Equation Model

Hui Jiang and Sui Huang

College of Teaching and Learning, The Ohio State University

(Received 26 March 2012; accepted 12 September 2012)

This paper uses the multiple-group mixture SEM model to examine the effects of parent characteristics, teacher qualification, and peer effect on students' academic performance. The potential differential effects for higher-achieving and lower-achieving groups are also examined. Using a national sample of 7022 third-grade students in the ECLS-K study, the project finds that parent characteristics and peer effect contribute positively to academic achievement, while teacher qualification has little effects. However, higher-achieving groups appear to be more sensitive to teacher effects than lower-achieving group. Moreover, parent characteristics remain the strongest predictor of children's academic success across samples.

I. INTRODUCTION

Research on predictors of students' academic achievement has developed along many lines, among which the study of family background (particularly the characteristics of parent(s)), teacher qualification, and peer effect are three important ones.

The first line of research focuses on the influence of parental characteristics and practice on student achievement. Many early studies find that children from families of high socioeconomic status (i.e., SES, typically measured by family income and parents' education levels) tend to go further in school and obtain more academic degrees than do children from families of lower SES (e.g., Sewell, Hauser, & Featherman, 1976). Particularly, literature on achievement has shown consistently that parent education is a very strong predictor of children's academic success (e.g., Haveman & Wolfe, 1995; Klebanov, Brooks-Gunn, & Duncan, 1994). Later studies suggest that SES measured by parents' levels of education and the family's income contribute indirectly to children's academic achievement through parents' beliefs and parenting behaviors (e.g., Davis-Kean, 2005). Studies have also suggested that the effects of families may interact with school and peer effects (Evans, Oates, & Schwab, 1992).

Research findings concerning the impact of teacher qualification have been mixed. On the one hand, some studies find teacher qualification to be a powerful predictor of student achievement. For example, Darling-Hammond (2000) emphasizes the importance of teacher education and certification, as she finds that the effects of well-prepared teachers can outweigh student background factors such as poverty and minority status. Laczko-Kerr and Berliner (2002) find that elementary school students taught by under-certified teachers did significantly worse on all subjects than those of regularly certified teachers. On the other hand, Rowan, Correnti, and Miller (2002) argue that teacher certification has no discernable effects on student achievement in elementary school; teachers' holding master's degrees has even had a negative effect on elementary students' achievement. Moreover, whereas various studies suggest that there are significant returns in the gains of student achievement as teachers have more experience, the returns tend to diminish beyond three to five years of teaching experience (e.g., Kane, et al., 2006).

The third line of research focuses on how peers influence student achievement. A common hypothesis is that other factors being equal, student outcomes are higher in the presence of peer groups that have higher ability levels. The famous Coleman report (Coleman et al., 1966) concludes that peer effects in public schools contributed to differences in the achievement of black and white students. A variety of econometric literature on the impact of school vouchers assumes that student achievement is influenced by the characteristics, achievement, or behavior of a person's classmates (e.g. Caucutt, 2002; Epple & Romano, 1998). Qualitative reports (e.g., Jencks & Mayer, 1990) suggest that there are negative effects of living or growing up in a poor neighborhood on school achievement, partly due to peer effects. A more recent study by Hanushek et al. (2003) reveals that even after removing student and school-by-grade effects in addition to the effects of observable family and school characteristics, peer achievement still has a significant positive effect on the achievement growth of individual students. In other words, students appear to benefit academically from higher achieving schoolmates.

This paper aims to examine the three major aspects - family characteristics, teacher qualification, and peer effects - simultaneously as potential predictors of student academic outcomes. There are two major questions:

1. How are parent characteristics, teacher qualification, and peer effect related with individual students' academic performance in elementary school?
2. Do the effects of parents, teachers, and peers on academic achievement differ for higher-achieving students and lower-achieving students?

II. THEORETICAL MODEL

This paper proposes a conceptual structural model to explore the research questions raised in the previous section. In this hypothesized model, parent characteristics (measured by parents' levels of education, occupational prestige, and family income), teacher qualification (measured by teachers' certification status, years of teaching experience, and levels of education), and peer effect (measured by classmates' average achievement levels) all contribute to individual students' academic performance. The outcome, i.e., the individual students' academic performance, is measured by students' cognitive assessment scores in reading, math, and science. The conceptual model is illustrated in Figure 1.

Based on this theoretical model, the paper will then use structural equation modeling (SEM) to examine the extent to which the conceptual model fit both a sample of higher-achieving students and a sample of lower-achieving students in third grade, and also to examine whether differences exist between these two groups of students. The dataset used is taken from the national database of Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (ECLS-K). See the next section for the description of the data.

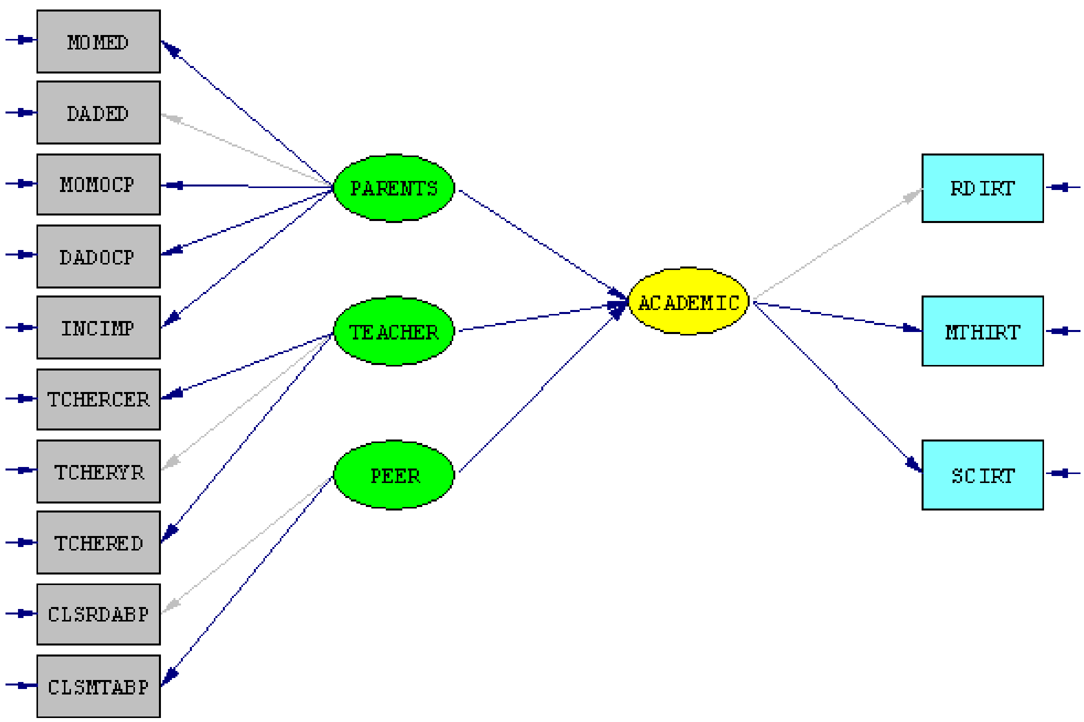

Figure 1. The conceptual model of students' academic performance

III. METHODS

A. Data

The data used in this study were obtained from the database of the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), conducted by the National Center for Education Statistics (NCES). This study followed a nationally representative cohort of children from kindergarten in the fall of 1998 through elementary school until eighth grade and collected data from the sampled children, their parents, teachers, and school principals in seven waves. Therefore, the ECLS-K database contains repeated measures of children's cognitive skills and knowledge, as well as children's physical health, social development, and emotional well-being. It also contains information on the characteristics of the schools and families of the sampled children, providing comprehensive contextual information to study children's growth (Tourangeau et al., 2009).

B. Sample

This study uses data from the 5th wave of the study, which was conducted in the spring of 2002, when the majority of participants were in third grade. The full sample includes all 13604 children who were in third grade and followed at the time of the study. After deleting listwise observations with missing data points on relevant variables, the final sample size is 7022.

Comparison between the final sample of children (n=7022) and the nationally representative sample (weighted) demonstrates that the final sample is not significantly different from the population in terms of gender ratio (p_{male}=50.5%, π_{male}=51.8%). There are significant differences between the final sample and the population in terms of achievement levels as measured by the average reading and math IRT scores, with the children included in the final achieving significantly higher on average ($\bar{X}_{\text{reading}}$=133.81, s_{reading}=25.72, μ_{reading}=122.66, p<.05; $\bar{X}_{\text{math}}$=104.75, s_{math}=23.05, μ_{math}=96.27, p<.05). In terms of ethnicity, white children were overrepresented in the final sample (p_{white}=69.8%, π_{white}=57.0%, p<.05) while black and Hispanic children were underrepresented. This may be due to the fact that teachers and parents of white children had a higher response rate compared to other ethnic groups. Therefore, caution needs to be taken if generalization is to be made to the whole population.

C. Instrument

Direct Cognitive Assessments (DCA)

DCA was designed to assess children's academic performance in reading, mathematics, and science. In the first stage of testing children took routing items. They were then assigned to different levels of difficulty in the second-stage of the test based on their performance on the routing items. Raw scores were converted into IRT-scale scores, which are criterion-referenced measures of children's academic performance in a particular time point. Previous research has shown that IRT scores obtained from DCA have reliabilities ranging from .88 to .96. It also demonstrates good content validity and criterion validity (Tourangeau et al., 2009).

Parent Questionnaire (PQ)

Parent interview for the spring-third grade data collection lasted on average 62 minutes and asked approximately 500 questions covering topics such as home environment, child care, parent characteristics, family structure, parental involvement in school, and child health. Most of the interviews were conducted by telephone, but a small percentage was conducted in person.

Teacher Questionnaires (TQ)

Each child's teacher received a self-administered questionnaire consisting of three distinct parts. The first section, part A (TQ-A), is the class-level questionnaire. It asked about the teacher's classroom and the characteristics of the students, instructional activities and practices, and student evaluation methods. Part B (TQ-B) is the teacher-level questionnaire. It asked questions on school activities, teacher's views on teaching, the school environment and climate, as well as the background information of the teacher. Part C (TQ-C) is the child-level questionnaire. Questions involved child's academic performance and social skills were asked in this section.

D. Variables and Measures

There are 13 observed variables included in the analysis. A brief description of these variables is given in Table 1.

Table 1. Description of observed variables [a)]

Variable name and description	Instrument	Mean	*SD*
MOMED: mother's education level.	PQ	4.79	1.78
DADED: father's education level.	PQ	4.81	2.02
MOMOCP: average prestige score for mother's occupation.	PQ	34.10	21.76
DADOCP: average prestige score for father's occupation.	PQ	42.38	14.04
INCIMP: income class.	PQ	9.21	2.63
TCHERCER: teacher's certification level.	TQ-B	3.83	.82
TCHERYR: number of years of teaching.	TQ-B	14.97	10.09
TCHERED: teacher's education level.	TQ-B	2.21	.91
CLSRDABP [b)]: proportion of class with reading scores above grade level.	TQ-A	.27	.18
CLSMTABP [b)]: proportion of class with math scores above grade level.	TQ-A	.21	.18
RDIRT: reading IRT scale score.	DCA	133.81	25.72
MTHIRT: math IRT scale score.	DCA	104.75	23.05
SCIRT: science IRT scale score.	DCA	.446	.497

There are four latent variables in the analysis, described in the following list.

PARENTS: parent characteristics and family background. This construct is measured by MOMED (mother's level of education), DADED (father's level of education), MOMOCP (prestige score for mother's occupation), DADOCP (prestige score for father's occupation), and INCIMP (income class).

TEACHER: teacher qualification. This construct is measured by TCHERCER (teacher's certification level), TCHERYR (teacher's years of teaching experience), and TCHERED (teacher's level of education).

PEER: peer/class performance. This construct is measured by CLSRDABP (percentage of classmates with reading scores above grade level) and CLSMTABP (percentage of classmates with math scores above grade level).

ACADEMIC: student academic achievement. This construct is measured by RDIRT (reading IRT scale score), MTHIRT (math IRT scale score) and SCIRT (science IRT scale score).

E. Analysis

In order to determine the extent to which the proposed theoretical models is supported by the collected sample data, structural equation modeling (SEM) is used to test the fit of the model. SEM, based on the general linear models, allows a researcher to test how sets of variables define constructs and how these constructs are related to each other. Because the conceptual model involves both continuous and ordinal variables, a Mixture Model should be used. Also, in order to test the potential differential effects between the sample of higher-achieving students and the sample of lower-achieving students, a Multiple Group Model is used. Multiple-group SEM allow for testing for group differences in the specified model or testing for differences in specific parameter estimates by imposing constraints (Schumacker & Lomax, 2004).

F. Assessment of Model fit

Chi-square value is the traditional measure for overall model fit and is the most popular index reported in studies. However, studies suggest that the chi-square test is usually very sensitive to sample size. Specifically, chi-square statistics tended to reject the model when sample sizes are large (e.g., Hooper et. al., 2008; Joreskog & Sorbom, 1993). In the model proposed in this study, the multivariate normality assumption is hardly satisfied, and the sample size is over 7000. Therefore, the standard chi-square test may lead to rejection of model fit even when the model is adequately specified. Thus, instead of emphasizing chi-square as the major index for model fit, we mainly looked at other fit indices such as Satorra-Bentler Scaled chi-square, CFI, NNFI, RMSEA, and PNFI to obtain the overall model fit (Hooper et. al., 2008; Schumacker & Lomax, 2004).

IV. RESULTS

A. Validating the fit of the hypothesized models – Whole sample mixture model

With the polyserial correlation matrix produced by PRELIS, we are able to test the hypothesized model (shown in Figure 1) first with the whole sample (without grouping) using LISREL. Although the chi-square statistics (χ^2=943.425, df=59, p<.05) obtained from the initial analysis didn't indicate a good model fit, we still considered the model fit acceptable looking at other fit indices (RMSEA=0.0462, p=0.992; NNFI=0.971; CFI=0.978; SRME=0.0248; PNFI=0.739). The model explained 32.1% of the variances in ACADEMIC. All paths are significant except for one: the path from TEACHER to ACADEMIC (t=1.919).

The modification indices suggested we add error covariances between MOMOCP and MOMED, DADOCP and DADED, DADED and MOMED, and MOMOCP and INCIMP. Considering these variables were measured by the same instrument (parent questionnaire), reported by the same respondents, and measuring similar aspects, they are very likely to be related to each other. Therefore we added the suggested four error covariances into the theoretical model.

Fit indices for the new model indicated the model fit had improved significantly after modification (χ^2=187.188, df=55, p<0.05; RMSEA=0.019, p=1.00; NNFI=0.995; CFI=0.997; SRME=0.014; PNFI=0.702). The model explained 35.1% of the variances in ACADEMIC. There was one non-significant path: path from TEACHER to ACADEMIC (t=1.864). All other paths were significant.

In order to obtain robust Satorra-Bentler chi-square statistics, the LISREL-SIMPLIS program was modified to include the *Covariance matrix from file*, *Asymptotic covariance matrix from file*, and *Method of estimation: maximum likelihood* commands. The final theoretical model is shown in Figure 2. The fit statistics are: Satorra-Bentler Scaled chi-Square = 60.864 (df=55, p < 0.05); RMSEA=0.004, p=1.00; NNFI=1.00; CFI=1.00; SRME=0.014; PNFI=0.704. These indices indicated a better model fit compared to the Normal Theory model results. PARENT, TEACHER, and PEER accounted for 35.1% of the variance in ACADEMIC. TEACHER remained an insignificant predictor for ACADEMIC, while all the other parameters were significant.

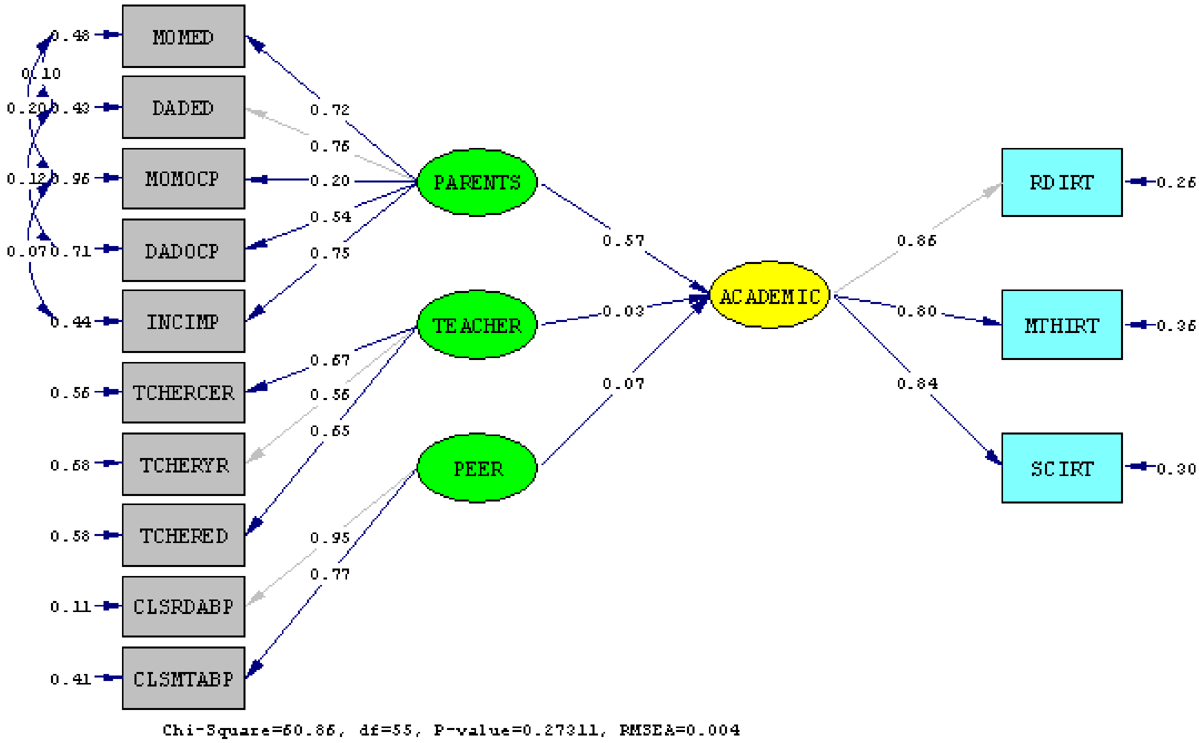

Figure 2. Path diagram of the whole sample mixture model (standardized solution)

B. Testing the Model Fit for Two Achievement Groups – Multiple Group Model

After validating the fit of the whole sample model, we divided the whole sample into two groups: higher-achieving group, including 3896 students who achieved at or above 50 percentile in three subjects on average in the full sample [c)]; and lower-achieving group, including 3126 students who achieved below 50 percentile. In order to test the differential effects of PARENTS, TEACHER, and PEER on ACADEMIC for the two groups, we first established the acceptance of the measurement models and measurement invariance for the groups. Therefore, five different multiple-sample measurement models were submitted to LISREL using asymptotic covariance matrices. The results are summarized in Table 2.

Table 2. Fit indices of five comparing measurement models

Model Details	χ^2	df	p	RMSEA	NNFI	CFI	PNFI	ECVI
A. Nothing invariant	90.71	110	0.91	0.00	1.00	1.00	0.70	0.04
B. Factor loadings invariant*	138.15	119	0.11	0.01	1.00	1.00	0.76	0.04
C. Loadings & measurement variance invariant	279.95	136	0.00	0.02	0.99	0.99	0.86	0.05
D. Loadings & variance/covariance of latent variables invariant	160.37	129	0.03	0.01	1.00	1.00	0.82	0.04
E. All parameters invariant	316.40	146	0.00	0.02	0.99	0.99	0.92	0.06

* Best fitting model identified

All five models had adequate RMSEA, NNFI, CFI, PNFI, and ECVI. However, only model A and model B provided satisfactory fit in terms of Satorra-Bentler Scaled Chi-Square. Even though model B, by keeping factor loadings invariant, led to a significant increase in chi-square as compared to model A, all its fit indices unanimously suggested that the model fit was acceptable. Therefore, assuming invariant factor loadings with the variances of measurement error and the variances/covariance of latent variables free to change between two groups is reasonable.

Next we examined the difference in the coefficients of latent variables between the two groups. First a Separate Group Model was run to provide separate path analysis estimates for lower-achieving and higher-achieving groups (factor loadings fixed, path coefficients free). Then a Similar Group Model was run to test whether the two groups share a common path model (factor loadings and path coefficients both invariant). Fit indices in Table 3 indicate that no significant differences existed between the two models (difference in χ^2=2.524, *df*=3, *p*>0.05). This implied that the lower-achieving and higher-achieving groups separately fitted the path model, while both datasets also fitted a common path model.

Table 3. Fit indices of two comparing path models

Model	χ^2	*df*	*p*	RMSEA	P (close fit)	NNFI	CFI	PNFI
Separate group	137.472	122	0.160	0.006	1.000	0.999	0.999	0.778
Similar group*	139.996	125	0.170	0.006	1.000	0.999	0.999	0.797
difference	2.524	3	ns*					

*ns = not significant

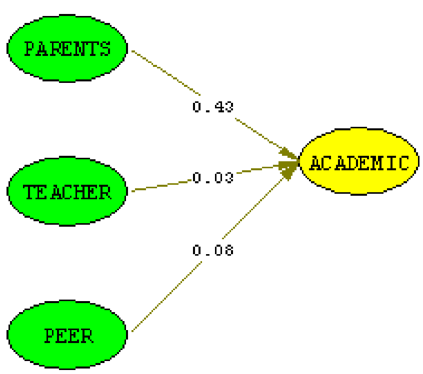

Figure 3. Path diagram of the Similar Group Model (standardized solution)

C. Interpreting the Model Results

Research Question One: How are parental characteristics, teacher qualification, and peer effect related to individual students' academic performance?

The final mixture model (as shown in Figure 2) suggested that parent statistics and peer effect both had significant positive effects on students' academic achievement (β_{parent}=0.57, t_{parent}=21.222; β_{peer}=0.07, t_{peer}=4.222)[d)], but the path from teacher qualification to academic achievement was found not to be significant ($\beta_{teacher}$=0.03, $t_{teacher}$=0.901). In other word, results of the mixture model suggested that third grade students whose parents had higher level of education, more prestigious occupation, and higher income were more likely to have higher achievement in reading, math, and science. Moreover, students with higher percentage of high-achieving classmates tended to perform better academically. On the other hand, the effects of teacher qualification (e.g., education level, certification type, years of teaching experience), though positive, were both small and non-significant, implying that higher teacher qualification did not lead to higher student achievement in this model when parental and class characteristics were controlled for. Overall speaking, the strongest predictor for students' academic performance were their parents' education level and family income. The equation for the whole sample mixture model is shown in Table 4.

Research Question Two: Do the effects of parents, teachers, and peers on academic achievement differ for higher-achieving students and lower-achieving students?

The data of two achieving groups fitted the Separate Group Model and Similar Group Model equally well, leading to two possible patterns of the differential effects.

The Similar Group Model (as shown in Figure 3) produced similar results to the whole-sample model. For both groups, parent characteristics and peer effect contributed significantly to students' academic achievement, whereas the effects of teacher qualification remained non-significant ($\beta_{parents}$=.43, $t_{parents}$=12.252; β_{peer}=.08, t_{peer}=3.958; $\beta_{teacher}$=.03, $t_{teacher}$=.966). Compared with the whole-sample mixture model, the multiple-group approach led to a decreased effect size of parent characteristics, but parental characteristics stayed the strongest predictor among the three independent latent variables. The equations for the similar group model are shown in Tables 5 & 6.

Table 4. Equation for Similar Group Model, whole sample mixture

ACADEMIC = 8.854*PARENTS + 0.0994*TEACHER + 9.118*PEER, Errorvar=316.866, R^2= 0.351				
	ACADEMIC	PARENTS	TEACHER	PEER
Standard error	0.417	0.110	2.160	12.489
t-score	21.222	0.901	4.222	25.371

Table 5. Equation for Similar Group Model, higher-achieving students

ACADEMIC = 3.778*PARENTS + 0.0595*TEACHER + 5.017*PEER, Errorvar=97.058, R^2= 0.183				
	ACADEMIC	PARENTS	TEACHER	PEER
Standard error	0.308	0.0616	1.267	7.455
t-score	12.252	0.966	3.958	13.018

Table 6. Equation for Similar Group Model, lower-achieving students

ACADEMIC = 3.778*PARENTS + 0.0595*TEACHER + 5.017*PEER, Errorvar=99.707, R^2= 0.217				
	ACADEMIC	PARENTS	TEACHER	PEER
Standard error	0.308	0.0616	1.267	6.698
t-score	12.252	0.966	3.958	14.886

Table 7. Equation for Separate Group Model, higher-achieving students

ACADEMIC = 3.580*PARENTS + 0.0883*TEACHER + 4.023*PEER, Errorvar=97.951, R^2= 0.166				
	ACADEMIC	PARENTS	TEACHER	PEER
Standard error	0.446	0.0811	1.573	7.838
t-score	8.024	1.089	2.557	12.496

Table 8. Equation for Separate Group Model, lower-achieving students

ACADEMIC = 3.984*PARENTS + 0.0200*TEACHER + 6.770*PEER, Errorvar=99.282, R^2= 0.239				
	ACADEMIC	PARENTS	TEACHER	PEER
Standard error	0.380	0.0889	1.849	6.814
t-score	10.479	0.225	3.662	14.571

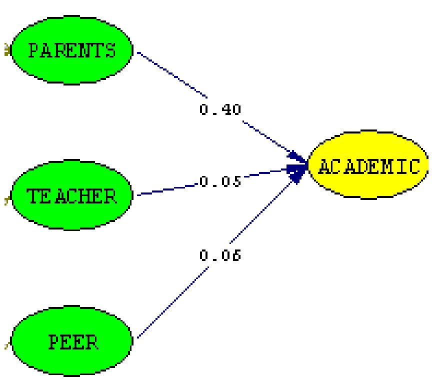

Figure 4. Path diagrams of Separate Group Model (standardized solution), higher-achieving group

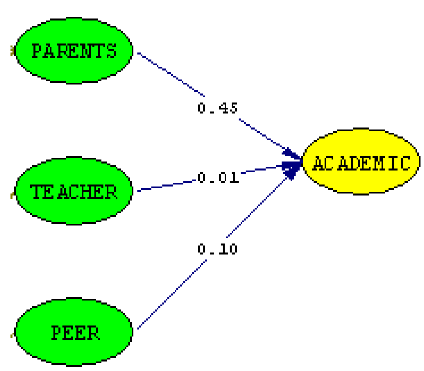

Figure 5. Path diagrams of Separate Group Model (standardized solution), lower-achieving group

Results of the Separate Group Model (shown in Figures 4 & 5) indicated that the paths from PARENTS and PEER to ACADEMIC had larger path coefficients and *t*-values for lower-achieving group (β_{parents}=0.45, t_{parents}=10.479; β_{peer}=0.10, t_{peer}=3.662) than for higher-achieving group (β_{parents}=0.40, t_{parents}=8.024; β_{peer}=0.05, t_{peer}=2.557). The effect of TEACHER, while small and non-significant for both groups, was more visible in higher-achieving group (β_{teacher}=0.05, t_{teacher}=1.089) than in lower-achieving group (β_{teacher}=0.01, $\text{t}_{\text{teacher}}$=0.225). This seemed to suggest that the higher-achieving students were more sensitive to the effects of teacher qualification than the lower-achieving students. On the other hand, the effects of parents and peers appeared to be stronger for lower-achieving group than for higher-achieving group. The equations for the similar group model are shown in Tables 7 & 8.

The multi-group structural equation model confirmed our findings from the whole-sample model that parent characteristics and peer effect had significant positive influence on students' academic achievement. Parent characteristics/family background was the stronger predictor of academic success among the two factors. Teacher qualification had little if any effect on student achievement in general; however higher-achieving students appeared to be slightly more sensitive to teacher effect than lower-achieving students did.

V. DISCUSSION

SEM analysis reaffirmed previous findings on the effects of parents and peers on student achievement. Parent characteristics, in particular, were shown to have significant positive effects on children's academic performance. Peer effects were relatively weak in magnitude, but still significant. The model also suggested that the effects of parents and peers were stronger among lower-achieving students. One possible reason for this phenomenon may be that the higher-achieving students were usually more self-reliant.

However, the mechanism behind the strong positive relationship was not explored in this project. Previous studies have suggested that having parents with higher levels of education is related to a warm, social climate, strong cognitive stimulation, and enhanced learning experiences in the home (e.g., Smith et al., 1997; Corwyn & Bradley, 2002). Other studies find that parent beliefs and expectation also play an important role in children's cognitive development (e.g., Davis-Kean, 2005). Therefore, future work can be done to test these theories.

This analysis also found a generally indiscernible teacher effect. Contrary to findings of some previous studies, teacher qualification measured by teachers' level of education, certification status, and years of teaching experience did not contribute significantly to student achievement, particularly for lower-achieving group. For higher-achieving group, however, teacher effect appeared to be slightly larger and more discernible. One possible reason for this finding is that third-grade teachers did not teach their students long enough when the assessment was conducted to fully reveal any differences they might have made in student achievement, or that teacher effects were simply overshadowed by students' family background. The group difference might be due to the fact that teachers tended to have more quality interaction with higher-achieving students than with lower-achieving students.

However, another possible explanation for the nonsignificant teacher effect may be that we did not measure "teacher qualification" with the appropriate indicators. Some previous studies have pointed out that conventional broad measures of teacher qualification such as levels of education and certification status did not do a good job in predicting teacher effectiveness, as defined by the potential to create academic gains (e.g., Nye, Konstantopoulos, & Heges, 2004; Rice, 2003). Some researchers suggest that one should use more refined measures of teachers' qualification, such as the specific type of degree or training/experience in specific subject matters, to predict teachers' contribution to student achievement (e.g., Croninger et al., 2007). Therefore, future studies can consider modifying measures for teacher qualification to retest the teacher effect.

ENDNOTES AND REFERENCES:

a) The scaling of observed variables:

MOMED=1 for "8th grade or below"; 2 for "9th -12th grade"; 3 for "high school diploma/equivalent"; 4 for "voc/tech program"; 5 for "some college"; 6 for "bachelor's degree"; 7 for "graduate/professional school – no degree"; 8 for "master's degree (MA, MS)"; 9 for "doctorate or professional degree". DADED uses the same code.

INCIMP=1 for "$5000 or less"; 2 for "$5001 to $10000"; 3 for "$10,001 to $15,000"; 4 for "$15,001 to $20,000"; 5 for "$20,001 to $25,000"; 6 for "$25,001 to $30,000"; 7 for "$30,001 to $35,000"; 8 for "$35,001 to $40,000"; 9 for "$40,001 to $50,000"; 10 for "$50,001 to $75,000"; 11 for "$75,001 to $100,000"; 12 for "$100,001 to $200,000"; 13 for "$200,001 or more". Strictly speaking this is an ordinal variable, but since this variable has 13 categories and is normally distributed, we treat it as a continuous variable in the analysis.

TCHERCER=1 for "uncertified"; 2 for "temporary/probational certified"; 3 for "alternative program certified"; 4 for "regular or standard state certified"; 5 for "advanced professional certified".

TCHERED=1 for "high school/associate's degree/bachelor's degree"; 2 for "at least one year beyond bachelor's"; 3 for "master's degree"; 4 for "education specialist/professional diploma/doctorate".

b) CLSRDABP and CLSMTABP were imputed from the database by dividing the number of children reading above grade level/whose math skills are above grade level by the number of children in the class.

c) The "full sample" refers to the sample including all third grade students before listwise deletion of missing data (N=13604). We did not use the 50 percentile cut in the final sample (N=7022) to sort students because students in this sample tended to achieve higher in all subjects than those of the larger, more representative sample.

d) β indicates the standardized coefficients.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence (pp. 8(1)). Journal of Education Policy Analysis.

Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment (pp.19 (2), 294-304). Journal of Family Psychology.

Caucutt E. M. (2002). Educational policy when there are peer group effects—size matters (pp.43(1), 195–222). International Economic Review.

Coleman, E.Q., Campbell, C.J., Hobson, J., McPartland, F., Mood, A.M., Weinfeld, F.D. & York, R. (1966), Equality of educational opportunity, US Government Printing Office, Washington, DC.

Corwyn, R. F., & Bradley, R. F. (2002). Family process mediators of the relation between SES and child outcomes. Unpublished manuscript, University of Arkansas at Little Rock.

Croninger, R., Ricea, J., Rathbuna, A., & Nishioa, M. (2007). Teacher qualifications and early learning: Effects of certification, degree, and experience on first-grade student achievement (pp. 26(3), 312-324). Economics of Education Review.

Epple, D., & Romano, R. (1998). Competition between private and public schools, vouchers, and peer-group effects (pp. 88(1), 33–62.). American Economic Review.

Evans, W, Oates, W., & Schwab R. (1992). Measuring peer group effects: A study of teenage behavior (pp. 100(5), 966–991). Journal of Political Economy.

Hanushek, E., Kain, J., Markman, J., & Rivkin, S. (2003). Does peer ability affect student achievement? (pp. 18(5), 527-544) Journal of Applied Econometrics.

Haveman, R., & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. Journal of Economic Literature, 33, 1829–1878.

Hooper, D. (2008). Structural equation modelling: guidelines for determining model fit (pp. 6(1), 53). Electronic Journal of Business Research Methods.

Jencks, C. & Mayer, S. (1990). The social consequences of growing up in a poor neighborhood. In: Lynn, L.E. & McGeary, M. (1990) (Ed.), Inner-city poverty in the United States, National Academy Press, Washington, DC.

Joreskog, K. G., & Sorbom, D. (1993). LISREL8: Structural equation modeling with the SIMPLIS command language. Hillsdale, NJ: Erlbaum.

Kane, T.J., Rockoff, J.E., & Staiger, D.O. (April, 2006). What does certification tell us about teacher effectiveness? Evidence from New York City. Cambridge, MA: National Bureau of Economic Research.

Klebanov, P. K., Brooks-Gunn, J., & Duncan, G. J. (1994). Does neighborhood and family poverty affect mothers' parenting, mental health, and social support? (pp. 56, 441–455) Journal of Marriage and the Family.

Laczko-Kerr, I., & Berliner, D. (2002). The effectiveness of Teach For America and other under-certified teachers on student academic achievement: A case of harmful public policy (pp. 10 (37)). Educational Policy Analysis Archives.

Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How Large Are Teacher Effects? (pp. 26 (3), 237-257) Educational Evaluation and Policy Analysis.

Rice, J.K. (2003). Teacher quality: Understanding the effects of teacher attributes. Washington, D.C.: Economic Policy Institute.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale, survey research tells us about teacher effects on student achievement: Insights from the prospects student of elementary schools (pp. 104, 1525-1567). Teachers College Record.

Schumacker, R. E. & Lomax, R.G. (2004). A beginner's guide to structural equation modeling, Lawrence Erlbaum, NJ: Mahwah.

Sewell, W., Hauser, D., & Featherman, D. (Ed.). (1976). Schooling and achievement in American society. New York: Academic Press.

Smith, J. R., Brooks-Gunn, J., & Klebanov, P. K. (1997). Consequences of living in poverty for young children's cognitive and verbal ability and early school achievement. In G. J. Duncan & J. Brooks-Gunn (Eds.), Consequences of growing up poor (pp. 132–189). New York: Russell Sage Foundation.

Tourangeau, K., Nord, C., Lê, T., Sorongon, A. G., and Najarian, M. (2009). Early Childhood Longitudinal Study, Kindergarten Class of 1998–99 (ECLS-K), Combined User's Manual for the ECLS-K Eighth-Grade and K–8 Full Sample Data Files and Electronic Codebooks (NCES 2009–004). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

