Online Appendix

Appendix Table A1

Regression Estimates from First Stage Teacher Value-Added Estimation					
	1	2		1 Cont'd	2 Cont'd
	Math	Reading		Math	Reading
Grade 4	-0.895	-0.734	Class Size	-0.005	-0.002
	[0.017]**	[0.018]**		[0.000]**	[0.000]**
Grade 5	-0.872	-0.722	Teacher: 1-3 years experience	0.068	0.03
	[0.018]**	[0.019]**		[0.018]**	[0.019]
Student Male	0	0.009	Teacher: 4-10 years experience	0.078	0.04
	[0.003]	[0.003]**		[0.018]**	[0.018]*
Student Black	-0.072	-0.082	Teacher: 10-24 years experience	0.071	0.033
	[0.007]**	[0.008]**		[0.018]**	[0.019]+
Student Hispanic	-0.023	0.004	Teacher: 25+ years experience	0.057	0.019
	[0.009]**	[0.009]		[0.019]*	[0.020]
Student American Indian	-0.099	-0.069	School: %Black	0.063	0.229
	[0.011]**	[0.012]**		[0.061]	[0.064]**
Student Mixed Ethnicity	-0.086	-0.059	School: %White	0.105	0.28
	[0.011]**	[0.012]**		[0.060]+	[0.063]**
Student White	-0.105	-0.074	School: %Hispanic	0.416	0.47
	[0.007]**	[0.008]**		[0.085]**	[0.089]**
Parental Education: Some High School	-0.002	0.001	School: %Free-Lunch Eligible	0.067	0.025
	[0.003]	[0.003]		[0.016]**	[0.017]
Parental Education: High School Graduate	0.006	0.007	School: Log Enrollment	0.001	0.002
	[0.004]	[0.004]+		[0.007]	[0.008]
Parental Education: Some College	0.002	0.008	School: Urban Fringe (Large City)	-0.068	-0.062
	[0.003]	[0.004]*		[0.018]**	[0.019]**
Parental Education: Prof. Graduate School	0.01	0.006	School: Mid-Sized City	-0.064	-0.07
	[0.003]**	[0.003]		[0.017]**	[0.018]**
Parental Education: Junior College Graduate	0.026	0.007	School: Urban Fringe (Mid-Sized City)	-0.077	-0.06
	[0.004]**	[0.004]+		[0.018]**	[0.019]**
Parental Education: College	0.037	-0.014	School: Large Town	-0.032	-0.033
	[0.008]**	[0.008]+		[0.030]	[0.031]
Parental Education: Graduate School	-0.016	0.312	School: Small Town	-0.075	-0.072
	[0.565]	[0.593]		[0.019]**	[0.020]**
Teacher and Student Are Same Race	0.009	0.006	School: Rural (Inside CBSA)	-0.088	-0.075
	[0.003]**	[0.003]*		[0.018]**	[0.019]**
Teacher and Student Are Same Sex	0.006	-0.004	School: Rural (Outside CBSA)	-0.042	-0.049
	[0.003]*	[0.003]		[0.018]*	[0.019]**
			Observations	535332	533060

Standard errors in brackets. + significant at 10%; * significant at 5%; ** significant at 1%.

Note: These models are estimated using student data from the years 1995 through 2000. All regressions include year fixed effects. The reference student's ethnic group is Asian students. The reference parental education group is no high school. The reference city size category is large city. The omitted teacher experience group is teachers with zero years of experience.

Appendix Table A2

	Panel A				Panel B	
	Selected Summary Statistics by Teacher's Status in Previous Year				Difference Between Selected Characteristics of Movers and Peers	
	Same Grade and School	Same School, Different Grade	Different School	New to Data	Same School, Different Grade	Different School
Percentage of All Teachers	65.82	5.95	7.38	20.85		
Experience	14.62	13.05	11.82	6.70	-0.008	-0.272
	(9.68)	(9.16)	(9.18)	(9.28)	(0.208)	(0.186)
Teacher Exam Score	-0.02	0.02	-0.03	0.05	0.018	-0.012
	(0.82)	(0.79)	(0.78)	(0.70)	(0.017)	(0.015)
Advanced Degree	0.23	0.22	0.21	0.15	0.007	0.016
	(0.42)	(0.41)	(0.41)	(0.36)	(0.009)	(0.008)*
Regular Licensure	0.52	0.50	0.54	0.36	0.029	0.05
	(0.50)	(0.50)	(0.50)	(0.48)	(0.005)**	(0.004)**
Certified	0.07	0.07	0.06	0.02	0.005	0.003
	(0.25)	(0.25)	(0.23)	(0.14)	(0.006)	(0.005)

Panel A: Standard deviations in parentheses.
Panel B: Standard errors in parentheses. +, *, and ** indicate significance of a *t* test that the mean is equal to zero at 10%, 5%, and 1%, respectively.

Appendix Table A3

Predictors of Receiving a New Peer

	1	2	3	4	5	6	7	8	9
	New Teacher from Same School, Different Grade	New Teacher, Different School	New Teacher Not from State Data	New Teacher from Same School, Different Grade	New Teacher, Different School	New Teacher Not from State Data	New Teacher from Same School, Different Grade	New Teacher, Different School	New Teacher Not from State Data
Lag Mean Math Test Score Growth	-0.001	-0.01	-0.01	-0.001	-0.018	-0.013	-0.013	-0.012	-0.002
	[0.011]	[0.010]	[0.014]	[0.013]	[0.011]	[0.016]	[0.016]	[0.015]	[0.018]
Lag Mean Reading Test Score Growth	0.01	0.006	0.003	0.009	0.012	0.009	0.02	0.01	-0.001
	[0.010]	[0.010]	[0.013]	[0.012]	[0.011]	[0.016]	[0.014]	[0.014]	[0.017]
Lag %teachers: 0 years experience	-0.003	0.029	-0.173	-0.003	0.03	-0.171	0.006	0.116	-0.165
	[0.095]	[0.105]	[0.132]	[0.095]	[0.105]	[0.132]	[0.107]	[0.116]	[0.147]
Lag %teachers: 1 to 3 years experience	0.051	-0.01	-0.23	0.051	-0.009	-0.228	0.072	0.074	-0.21
	[0.086]	[0.097]	[0.120]+	[0.086]	[0.097]	[0.120]+	[0.095]	[0.103]	[0.128]
Lag %teachers: 4 to 9 years experience	0.052	-0.021	-0.201	0.052	-0.02	-0.199	0.066	0.065	-0.203
	[0.086]	[0.095]	[0.119]+	[0.086]	[0.095]	[0.119]+	[0.094]	[0.100]	[0.127]
Lag %teachers: 10 to 24 years experience	0.053	-0.046	-0.262	0.053	-0.046	-0.26	0.05	0.023	-0.269
	[0.085]	[0.095]	[0.121]*	[0.085]	[0.095]	[0.120]*	[0.094]	[0.101]	[0.128]*
Lag %teachers: 24+ years experience	0.105	-0.028	-0.245	0.105	-0.027	-0.243	0.118	0.045	-0.235
	[0.088]	[0.098]	[0.121]*	[0.088]	[0.098]	[0.121]*	[0.096]	[0.105]	[0.128]+
Lag Average Math Test Score	—	—	—	0.001	-0.012	-0.012	-0.003	-0.01	-0.004
	—	—	—	[0.012]	[0.011]	[0.014]	[0.014]	[0.013]	[0.016]
Lag Average Reading Test Score	—	—	—	-0.001	0.014	0.005	0.004	0.012	0.002
	—	—	—	[0.012]	[0.011]	[0.014]	[0.014]	[0.014]	[0.016]
Lag Mean Teacher Value-Added Math	—	—	—	—	—	—	-0.013	-0.016	-0.041
	—	—	—	—	—	—	[0.022]	[0.021]	[0.027]
Lag Mean Teacher Value-Added Reading	—	—	—	—	—	—	0.003	-0.027	-0.001
	—	—	—	—	—	—	[0.033]	[0.025]	[0.035]
Observations	19550	19550	19550	19550	19550	19550	12466	12466	12466
R-squared	0.6	0.63	0.68	0.6	0.63	0.68	0.6	0.62	0.68

Robust standard errors in brackets.

+ Significant at 10%; * significant at 5%; ** significant at 1%.

Includes grade fixed effects and school by year fixed effects.

Appendix Table A4

Interaction of Peer Quality and Own Characteristics

	Dependent Variable: Math Test Score				Dependent Variable: Reading Test Score			
	1	2	3	4	5	6	7	8
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Peer Effect	0.0599	0.0205	0.0376	0.0354	0.063	0.0198	0.0376	0.0163
	[0.0283]*	[0.0069]**	[0.0059]**	[0.0067]**	[0.0249]*	[0.0078]*	[0.0059]**	[0.0063]**
Peer Effect*Experience 1 to 3	-0.003	—	—	—	-0.0412	—	—	—
	[0.0273]	—	—	—	[0.0254]	—	—	—
Peer Effect*Experience 4 to 9	-0.0077	—	—	—	-0.039	—	—	—
	[0.0294]	—	—	—	[0.0264]	—	—	—
Peer Effect*Experience 10 to 24	-0.0361	—	—	—	-0.0384	—	—	—
	[0.0293]	—	—	—	[0.0257]	—	—	—
Peer Effect*Experience 25+	-0.016	—	—	—	-0.0278	—	—	—
	[0.0310]	—	—	—	[0.0272]	—	—	—
Peer Effect*Regular Licensure	—	0.0323	—	—	—	0.0167	—	—
	—	[0.0074]**	—	—	—	[0.0086]+	—	—
Peer Effect*Certified	—	—	0.0129	—	—	—	0.0189	—
	—	—	[0.0143]	—	—	—	[0.0183]	—
Best Teacher [a]	—	—	—	-0.0067	—	—	—	-0.0066
	—	—	—	[0.0049]	—	—	—	[0.0046]
Worst Teacher [a]	—	—	—	0.0002	—	—	—	0.0136
	—	—	—	[0.0049]	—	—	—	[0.0046]**
Prob(Worst = Best)	—	—	—	.27	—	—	—	.03
Prob(TFX(exp >10) = (TFX(exp >10))	0.08	—	—	—	.21	—	—	—
Teacher-School Effects	YES	YES	YES	YES	YES	YES	YES	YES
School-Year Effects	YES	YES	YES	YES	YES	YES	YES	YES
Observations	684752	684752	684752	684752	679230	679230	679230	679230

Robust standard errors clustered by school-teacher in brackets.

\+ Significant at 10%; * significant at 5%; ** significant at 1%.

Note: Estimated using data from 2001 to 2006. The variable “peer effect” is the mean estimated value-added of a teacher’s peers (all other teachers at the same school in the same grade during the same year). All models include indicator variables for the gender and racial matches between the teacher and the students, class size, student demographic control variables, teacher experience, indicators for missing estimated value-added, the proportion of peers with no estimated value-added, and year-by-grade fixed effects. The omitted teacher experience group is teachers with zero years of experience.

a. Best Teacher and Worst Teacher are indicator variables that take the value of one if the teacher has the highest or lowest estimated value-added among her peers, respectively, and zero otherwise.

Appendix Note 1: Estimating Teacher Fixed Effects

There are several specifications used in the literature to estimate teacher value-added [e.g. Aaronson et. al. 2007, Rockoff 2004, Hanushek, Kain, and Rivkin 2005, Jacob and Lefgren 2008]; however, the predictive power of estimated teacher fixed effects are generally robust to the chosen specification [Kane and Staiger (2008)]. We estimate teacher fixed effects using the adjusted test score growth model described in Section II. Specifically, teacher effectiveness comes from estimation of equation [3] using data from 1995 through 2000.

[3] $$A_{it} = \hat{\delta} A_{it-1} + \phi X_{it} + \varsigma Z_{st} + \eta W_{jt} + \theta_j + \xi_{gt} + \varepsilon_{ijgst} .$$

All variables are defined as before, with the addition of θ_j,which is the effect of teacher *j*. $\hat{\delta}$ is the coefficient on lagged test scores in a test score growth model obtained from a 2SLS regression using the second lag of test scores as an instrument for lagged test scores. Because we use the first year of data to compute test score growth for 1996, the actual estimation sample used spans the years 1996 through 2000. Because we need estimates of teacher value-added that can be comparable across schools, grades, and classes, we do *not* include school or student fixed effects but rather include a set of demographic controls for the students and schools.[1]

Researchers have pointed out that there is substantial measurement error in test scores such that the coefficient on lagged test score would be downward biased.[2] Under the assumption that measurement errors in test scores are not correlated over time, many researchers have used the second lag of test scores as an instrument for the lagged test score [as proposed in Anderson and Hsiao (1981) and Todd and Wolpin (2003)]. One downside of this approach is that it requires several years of data for each student and is impractical to implement in models that include large vectors of three fixed factors. We propose a method that builds on this solution but allows one to use more of the available data. The basic idea is that if using the second lag of test scores as an instrument for the lagged test score results in a consistent estimate of δ, then one can use this estimate to adjust the test score growth outcome variable for the full sample and obtain consistent estimates on the coefficient for other characteristics that may be correlated with lagged test scores. We present a proof of this below.

This is implemented by first estimating the instrumental variables regressions on the full sample, where the second lag of test scores is used as an instrument for the first lag of test scores. The consistent estimates of the coefficient on lagged test scores is therefore estimated *in sample*. Because this can be estimated only for students with two lags of test scores, this 2SLS model uses only grade 5 outcomes. We then estimate the sample analog of equation [4]

[1] Specifications that include student or school fixed effects identify teacher value-added based on within-school or within-student variation. If teachers are very different across schools, then much of the variation in teacher quality (i.e., the cross-school variation) will be absorbed by the school fixed effect, making estimated effects across schools impossible to compare. Including student fixed effects further exacerbates this problem by allowing only comparisons of teachers who teach the same groups of students. If those teachers who teach the gifted and talented students are of different average quality from those who teach the regular students, the estimated teacher value-added can be used only to compare teachers who share the same students, so that comparing teachers who teach different students (even within the same school) may be misguided.

[2] This is the same as saying that there is attenuation bias on the coefficient of lagged test scores in [3] due to measurement error in test scores. If lagged test scores are correlated with other covariates (very likely), this will bias the coefficients for all covariates.

(replacing δ with $\hat{\delta}$) using all the observations for which lagged test score are available. As a practical matter, although the 2SLS coefficient on lagged test scores (between 0.97 and 0.95) is *much* smaller than the OLS estimates (between 0.70 and 0.76), the peer effects results are similar across models, so that our results are not driven by any modeling assumptions from this procedure. However, the teacher value-added estimates perform better in the falsification test of Section VI (as would be expected if the 2SLS adjustment removes measurement error bias from the teacher estimates).

Proof : Consider the following. We can rewrite [3] as $A_{it} - \delta A_{it-1} = \phi H_{it} + \varepsilon_{it}$ where H_{it} denotes all observable covariates and teacher, grade, and school subscripts are suppressed. Suppose we have a consistent unbiased estimate $\hat{\delta}$ of δ such that $\lim_{n\to\infty} \hat{\delta} = \delta$ and $E(\hat{\delta}) = \delta$. Where test scores are measured with error such that $\hat{A}_{it} = A_{it} + u_{it}$ and one uses $\hat{\delta}$ in the place of δ, this can be written as [4] below.

[4] $$A_{it} - \hat{\delta} A_{it-1} = \phi H_{it} - (\delta - \hat{\delta}) A_{it-1} + \delta u_{it-1} - u_{it} + \varepsilon_{it} .$$

Equation [4] can be directly estimated using OLS where the unobserved error term is $\delta u_{it-1} - u_{it} + \varepsilon_{it}$. Because H_{it} is uncorrelated with u_{it-1}, u_{it}, and ε_{it} by assumption, the OLS estimate $\hat{\phi}$ of ϕ from [4] will be unbiased and consistent *iff* $E[\widehat{Cov}(H_{it}, (\delta - \hat{\delta}) A_{it-1})] = 0$ where $\widehat{Cov}$ is the sample covariance. Using Slutsky's theorems, because $\lim_{n\to\infty} \hat{\delta} = \delta$ and $E(\hat{\delta}) = \delta$, it follows that $\lim_{n\to\infty}[\widehat{Cov}(H_{it}, (\delta - \hat{\delta}) A_{it-1})] = 0$ and $E[\widehat{Cov}(H_{it}, (\delta - \hat{\delta}) A_{it-1})] = 0$ so that $\lim_{n\to\infty}(\hat{\phi}) = \phi$ and $E(\hat{\phi}) = \phi$ from [4].