Purpose

In the past few decades, leading school reform policies embraced the principle of incentives as a key component to building a stronger accountability system. For instance, the landmark policy No Child Left Behind (NCLB) engendered highly visible incentives to hold schools accountable. Under the incentive scheme, successful schools are rewarded with public recognition and financial bonuses while those missing the adequate yearly progress (AYP) target are labeled as failing schools and subject to a variety of consequences (U.S Department of Education, 2001).

At the same time, given the vital role of teacher quality in student learning, most of the accountability policies take particular aim at the training, recruiting, and rewarding of effective teachers. For instance, reform efforts to reauthorize NCLB called on states to revise the evaluation and compensation system of teachers and build "a new culture of accountability": teacher accountability (President Obama, 2009).

Given this rich policy context, it is essential to examine the effects of accountability policies on school staffing practices, teacher effectiveness and teacher assignment. Existing research has focused on the selection and de-selection of teachers as a result of accountability pressure with little evidence on school staffing practice and teacher assignment.[1] This paper focuses on the sanction and reward elements of the incentive scheme and provides causal evidence on the implication of getting an A or F performance grade on teacher assignment within schools.

The challenge, however, is that over time schools learnt their way around accountability systems, making it difficult to get unbiased estimates of the policy effect. For example, in

[1] For instance, prior research has examined teachers' decisions to enter or exit the teaching profession.

Florida, schools would have to fail all three subjects (Reading, Math, and Writing) to receive an F grade. So long as students did well enough in at least one subject, schools would escape the F stigma. This opened up opportunities to game to the test: schools put more emphasis on Reading since it is much easier to show improvement as compared to Math and Writing (Rouse et.al, 2007; Goldhabor & Hannaway, 2004; Chakrabarti, 2006). To overcome this endogeneity problem, this paper explores an exogenous grading change that happened in Florida in 2002 and relies on this accountability shock to identify policy effects.

In summer 2002, as part of NCLB requirements, Florida Department of Education added student learning gain into its grading formula, generating an exogenous shock to school grades (Florida Department of Education, 2002). As a result of the grading change, 28 (15) percent of elementary schools that would have earned a C (D) under the old grading system received an F; and 48 (23) percent of schools that would have earned a B (C) under the old system received an A (Rouse, et.al, 2007).The grading change provides a unique opportunity to examine the effects of sanctions and rewards on school staffing changes, since it generates exogenous changes in school grades which are unrelated to school characteristics.[2] This paper takes advantage of this policy shock and answers the following questions:

1. Descriptively, how are teachers assigned to students within schools?
2. What are the impacts of getting an A or F grade on teacher assignment?
3. To what extent is the assignment shaped by principal characteristics?

The contribution of this study to the current literature is two-fold. First, it provides the first causal evidence on the impacts of accountability on teacher assignment within schools.

---

[2] Though it is true that schools were aware of the possible grading change prior summer 2002, as explored in Rouse et.al (2007), they did not have full information and did not know how exactly the new grades would be calculated. Technically, so long as schools are not able to *precisely* manipulate their school grades, we can estimate the causal effects of getting an A or F performance grade on teacher assignment.

Second, it explores the role principal characteristics played in the assignment of teachers and contributes to our understanding of how different schools respond to accountability pressure.

Research Methodology

The Impact of Accountability on Teacher Assignment

Grade assignment. This paper assumes that a teacher's probability of being assigned to tested grades is a function of teacher characteristics T, grade assignment in previous school year y_{it-1}, principal attributes P, school characteristics S and a policy variable w_{it}, indicating whether schools received an A or F grade in summer 2002:

$$P(y_{it} = 1 | T, S, P) = \Phi(\beta T_{it} + \chi y_{it-1} + \tau w_{it} + \varphi P_{it} + \phi S_{it}) \quad (1)$$

where Φ is the Cumulative Distribution Function (CDF) of the standard normal.

Classroom assignment. Similarly, for classroom assignment:

$$C_{it} = \alpha + \beta T_{it} + \gamma q_{it-1} * I_{it-1} + \tau w_{it} + \varphi P_{it} + \phi S_{it} + \varepsilon_{it} \quad (2)$$

where C_{it} indicates classroom characteristics, I_{it-1} is an indicator of whether the teacher was assigned to tested grades (3-5) in previous school year, q_{it-1} individual teacher value-added scores in previous year and T_{it}, S_{it}, P_{it} are same as above.[3]

In equation (1) and (2) τ captures the effects of being rewarded or sanctioned. There are, however, issues with the identification and inference of τ. First, as discussed before, due to the possible gaming around accountability, schools that meet the AYP target will be systematically different from those missing the target. Any direct estimation of τ without correcting for

[3] Alternatively, classroom assignment can be modeled using Ordered Probit Models. For instance, each classroom is ranked on a 1-3 scale based on student characteristics. For example, 1 indicating the classroom has a low percent of minority students and 3 indicating a high percent of minority students. The thresholds that partition the real line into a series of regions corresponding to the various ordinal categories are arbitrary and typically estimated in practice. In this paper, estimates from OMP are qualitatively similar to the linear specification.

selection will be biased. Second, since the teacher effectiveness measure q is an estimated regressor, the usual standard errors in general will be incorrect.

Identification of τ

The grading shock in 2002 provides a unique opportunity to estimate τ since it generates exogenous changes in school grades that are unrelated to school characteristics. Therefore, I employ a difference-in-difference (DiD) method to estimate τ**.** Alternative identification utilizes a regression-discontinuity (RD) design that exploits the administrative rules that determine school grades.[4]

Difference-in-Difference (DiD). Let $E[y_{it}(1)]$ denotes the outcome for teachers in schools that received an A performance grade and $E[y_{it}(0)]$ for those that did not. w is still the treatment indicator, X school covariates such as principal and school characteristics, and c time-constant individual school heterogeneity. Assuming the conditional mean independence:

$$E[y_{it}(0) \mid w_i, X_i, c_j(0), c_j(1)] = E[y_{it}(0) \mid X_i, c_j(0)] \qquad (3)$$

$$E[y_{it}(1) \mid w_i, X_i, c_j(0), c_j(1)] = E[y_{it}(1) \mid X_i, c_j(1)] \qquad (4)$$

Further assume that once we control for school points and heterogeneity, the difference between control and treatment group reflects the treatment effect:

$$E[y_{it}(1) \mid X_i, c_i(1)] = E[y_{it}(0) \mid X_i, c_i(0)] + \tau_t, \quad t = 1, \ldots, T. \qquad (5)$$

Finally, we have:

[4] Since RD estimates require the use of school points information, which is not available till 2001 school year, this study uses DiD as main identification strategy and RD as robustness check.

$$
\begin{aligned}
&E[y_{it} \mid w_i, X_i, c_j(0), c_j(1)] \\
&= (1 - w_{it})E[y_{it}(0) \mid w_i, X_i, c_j(0)] + w_{it}E[y_{it}(1) \mid w_i, X_i, c_j(1)] \\
&= E[y_{it}(0) \mid X_i, c_j(0)] + w_{it}[E[y_{it}(1) \mid X_i, c_j(1)] - E[y_{it}(0) \mid X_i, c_j(0)]] \\
&= E[y_{it}(0) \mid X_i, c_j(0)] + \tau_t w_{it}
\end{aligned} \quad (6)
$$

A regression form of equation (6) is:

$$y_{it} = \alpha + \beta T_{it} + \gamma q_{it-1} * I_{it-1} + \tau w_{it} + \varphi P_{it} + \phi S_{it} + c_j + d_t + \varepsilon_{it} \quad (7)$$

where d_t is the vector of school year fixed effects and the rest follows above.[5] Estimation of equation (7) is carried out using fixed effects (FE).[6]

Regression discontinuity (RD) design. Given the administrative rules that determine school grades each year, for instance, schools receive an A performance grade if their school points are greater than 410; and an F grade if school points are fewer than 395, alternative identification comes from comparisons between schools that are just above or below the A/F threshold. Let $E[Y_i(1) \mid X]$ denotes teacher assignment in schools that earned an A, $E[Y_i(0) \mid X]$

[5] The inclusion of aggregate time effects is to take care of any stake-wide policy that affected all schools during the time span. The model is subject to endogeneity bias when the policy does not affect all schools equally. For instance, the Class Size Reduction (CSR) of 2004, which may have profound influences on student grouping, teacher mobility and assignment. Anecdotal evidence suggests that schools may respond to CSR differently as it is more difficult to implement the class size cap in hard-to-staff inner city schools or rural schools. For this reason, this study included class size measure as additional controls in the hope that it could partially address the problem.

[6] In cases where w_{it} may be correlated with unit-specific trends in the response, a correlated random trend model is used (Woolridge, 2009):

$$y_{it} = \alpha + g_i t + \beta T_{it} + \gamma q_{it-1} * I_{it-1} + \tau w_{it} + \varphi P_{it} + \phi S_{it} + c_j + d_t + \varepsilon_{it} \quad (1)$$

where g_i is the trend for individual teacher i. This set up allows arbitrary correlation between $(c_, g_i)$ and w_{it}, for example, as a teacher matures, she is more likely to be assigned into tested grade, regardless of the policy intervention. By first differencing, we have:

$$\Delta y_{it} = g_i + \beta \Delta T_{it} + \gamma \Delta(q_{it-1} * I_{it-1}) + \tau \Delta w_{it} + \varphi \Delta P_{it} + \phi \Delta S_{it} + (d_t - d_{t-1}) + \Delta \varepsilon_{it} \quad (2)$$

which can be estimated by differencing again or FE. If we want to allow the effect of the policy to change over time, we can add interaction term between time dummies and policy indicator into the equation. Throughout this paper, I use equation (7) to estimate τ and above equations for robustness checks.

for those that do not, and $E[A_i \mid X]$ the probability of getting an A grade, the local average treatment effect (LATE) is estimated by:

$$\tau^A = \frac{\lim_{\varepsilon \downarrow 0} E[Y_i \mid X_i = c+\varepsilon] - \lim_{\varepsilon \uparrow 0} E[Y_i \mid X_i = c+\varepsilon]}{\lim_{\varepsilon \downarrow 0} E[A_i \mid X_i = c+\varepsilon] - \lim_{\varepsilon \uparrow 0} E[A_i \mid X_i = c+\varepsilon]} \qquad (8)$$

provided that around cutoff value c there is a local randomization and school points are continuous at c. Since for F schools, school points perfectly predicts schools' receipts of F grade, the numerator goes to one and the treatment effect is estimated by taking the difference at the limit:

$$\begin{aligned} \tau^F &= \lim_{\varepsilon \uparrow 0} E[Y_i \mid X_i = c+\varepsilon] - \lim_{\varepsilon \downarrow 0} E[Y_i \mid X_i = c+\varepsilon] \\ &= E[Y_i(1) - Y_i(0) \mid X = c] \end{aligned} \qquad (9)$$

Empirical estimates of τ are carried out using polynomial regression.[7] More specifically, τ^A is estimated using two-stage least square (2-SLS) with A being the instrument for w; and τ^F is estimated by ordinary least squares (OLS). The regression form is:

$$y_{it} = \alpha + \beta T_{it} + \eta y_{it-1} + \gamma q_{it-1} * I_{it-1} + \tau W_{i2002} + \gamma X_{it} + \varphi P_{it} + \phi S_{it} + d_t + \varepsilon_{it} \qquad (10)$$

where the inclusion of y_{t-1} is to take advantage of the longitudinal nature of the data and reduce sampling error. [8] X includes third order polynomials of school points and the rest are same as above.[9]

[7] In theory, we can use local linear regression (LLR) that restricts the data to a close neighborhood around the cutoff. However, this is not feasible to do in practice as it would result in extreme small sample for F schools.

[8] Due to the longitudinal nature of the data, it might be beneficial to include individual school fixed effects to account for individual heterogeneity that may be correlated with treatment status. However, as discussed in Lee & Lemieux (2011) and Hahn, Todd & van der Klaauw (2001), the source of identification in RD design comes from comparisons between schools that barely passed and barely missed the threshold, thus the inclusion of individual fixed effects does not bring any gains to the identification.

[9] As we shall see later, the inclusion of third order polynomials is sufficient to control for the differences in school characteristics in most cases. Though the cross-validation criteria tend to favor higher order of polynomials (fifth and sixth), those higher order polynomials are not jointly significant and scatter plots suggest they may lead to over-fitting of the data.

Inference of τ

For teachers teaching tested grades, their effectiveness measure, q enters directly into equation (7) and (10). q is unobservable but it is related to observable data through the function $q = f(x, \delta)$ where f is a known function and x is a vector of observed variables (i.e., student characteristics). For each observation i, $\hat{q}_{it} = f(x, \hat{\delta})$ effectively estimates q_{it}. Replacing q_{it} with $\hat{q}_{it}$ causes no consistency but inference problem. Since q is estimated from previous equation, the usual inference for OLS and 2-SLS estimators in general are no longer valid (Wooldridge, 2002). Though the problem is less salient with a large sample, we still need to adjust the standard errors to account for the fact that q is a generated regressor. A formal derivation of the adjusted standard errors is given in the Appendix 10.1. It is derived in the form of 2-SLS estimators with generated regressor and covers OLS estimators.

Data

Data is from the *Florida Education Data Warehouse (EDW),* a state-wide longitudinal database that tracks individual students from pre-K through postsecondary institutions. The sample includes all students, teachers and principals in public elementary schools (and k-8 combination schools that served 1-5 grades) from 1998-99 through 2004-05 school years, which gives 5,242,522 student-year observation including 1,960,205 unique students, 62,988 unique teachers, 222,147 unique classrooms, 1,670 unique schools and 2,423 unique principals. A descriptive summary of student, teacher, and school characteristics is presented in Table 1.

Results[10]

Descriptively, How Are Teachers Assigned to Students within Schools?

[10] Due to space limits, only difference-in-difference estimates are reported. Results based on regression discontinuity are qualitatively similar and available upon request.

Table 4.1 presents the summary of teacher qualifications by student characteristics. It is clear that minority, LEP and students who are eligible for free/reduced price lunch are more likely to be taught by minority, uncertified, inexperienced and less effective teachers. Interestingly, Black and Hispanic students are more likely to be taught by teachers of the same race, echoing the findings by Dee (2004) and Hanushek et.al, (2005). Panel B reports teacher qualifications by student preparation levels. Not surprisingly, high performing students are more likely to be assigned to effective teachers and less likely to be taught by minority, inexperienced and uncertified teachers.

Table 4.2 compares within teacher variance in student prior test scores to two simulated scenarios, one where students are randomly assigned to teachers (within and across schools) and the other where students are perfectly sorted based on their rankings in test scores distribution. Though teachers and students are not randomly assigned, the results suggest that the actual teacher-student sorting is closer to what we would expect from the random assignment scenario than what we would expect from a perfect sorting scheme.

What Are the Impacts of Getting an A or F Grade on Teacher Assignment?

Grade assignment. With everything else being equal, getting an A performance grade increases teacher's probability of being assigned to tested grades (3-5) by 0.0028 while getting an F increases the probability by 0.0386 (Table 5.1). However, both results are not significant at any conventional level. This is not surprising given that the number of teachers in a year is fixed and an increase in teacher i's probability of being assigned to tested grades must be offset by a decrease in teacher j's assignment.

Table 5.2 explores the heterogeneities in policy effects and examines how teacher assignment is shaped by principal characteristics. In schools that earned an A in 2002, novice teachers are on average more likely to teach tested grades; however, there are significant disparities by principal experience: experienced principals are less likely to assign novice teachers to tested grades. Similar findings hold for F grades. Further, we see more heterogeneities in F schools: experienced principals are more likely to assign certified, and teachers with an advanced degrees to tested grades.

Classroom assignment (conditional on grade assignment). Table 6.1 examines the effects of getting an A grade on teachers' classroom assignment. The results suggest that principals do respond to accountability systematically by assigning teachers, especially novice, least effective, and most effective teachers in particular ways. For example, novice teachers are more likely to be assigned to classrooms with higher percent of Black students or high-performing students. At the same time, we also see heterogeneities by principal characteristics. For instance, experienced principals are less likely to assign novice or highly effective teachers to Black students. In F schools, we find more staffing changes as a result of the grading shock (Table 6.2). There is some evidence that principals are assigning effective teachers to minority, FRPL or low-performing students, however, since teachers are more likely to leave F graded schools (Feng, Figlio & Sass, 2011), it is hard to see whether principals are doing so to fulfill NCLB expectations or responding to teacher mobility.

Non-tested grades. Table 7.1 and 7.2 presents the results for non-tested grades. In general, I find fewer staffing changes in non-tested grades, consistent with the notion that under accountability pressure, schools are concentrating their resources on tested grades.

Conclusions

This paper takes advantage of an exogenous grading change that happened in 2002 to analyze the causal effects of getting an A or F grade on teacher assignment within schools. The results suggest that getting an A or F grade affects teacher assignment, both grade assignment and classroom assignment systematically. Principals' experience and newness to schools also play significant roles in teacher assignment. Though the results are robust across various alternative specifications, there are two caveats to the findings:

First, this paper examines the effect of getting an A performance grade on teacher assignment. The reward or benefit of getting an A grade, however, goes beyond conventional and/or pecuniary measures. Schools may simply benefit from the labeling or branding effect of being an A school. Also, there is increasing evidence showing school quality has been capitalized into neighborhood housing market (Black, 1999; Figlio & Lucas, 2004). Since without appropriate data it is hard to tease out the pure signaling effect, the results from this study may serve as the lower bound of the policy effects.

Second, this paper so far has been describing the patterns of teacher assignment within schools; it does not make any normative statements, for example, it does not conclude whether the assignment practice is desirable or undesirable.[11] Future study may look at the impacts of teacher placement on student achievement, particularly those minority and low income students. It answers important policy questions as how we can improve student achievement through redistribution of teachers within schools.

[11] Embedded in this statement is not value-laden judgment but whether or not the assignment practice improves student achievement.

Key References

Cohen-Vogel, L. (2010). Staffing to the test: Are today's school personnel practices evidence based? *Educational Evaluation and Policy Analysis*, 33(3).

Feng, L., Figlio, D. & Sass, T. (2011). School accountability and teacher mobility. NBER working paper #16070.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*,69(1), 201-209.

Imbens, G., & Kalyanaraman, K. (2009). Optimal bandwidth choice for the regression discontinuity estimator. NBER working paper #14726.

Jacob, B.A. (2010). Do principals fire the worst teachers? NBER working paper # 15715.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2), 281-355.

Rockoff, J., Staiger, D.O., Kane, T.J., & Taylor, E.S. (forthcoming). Information and employee evaluation: Evidence from a randomized intervention in public schools. *American Economic Review.*

Rothstein, J. (forthcoming). Teacher quality in educational production: Tracking, decay and student achievement. *Quarterly Journal of Economics.*

Rouse, C.E., Hannaway, J., Goldhaber, D., & Figlio,D. (2007). Feeling the Florida heat? How low-performing schools respond to voucher and accountability pressure. *NBER working paper 13681*.

Table 1. Descriptive Statistics of Student, Teacher and Principal Characteristics in Florida Elementary Schools, 1998-2004

	Mean	SD
Student Characteristics		
Percent of students who are:		
Female	0.49	0.50
Black	0.23	0.42
Hispanic	0.22	0.42
White	0.50	0.50
Eligible for free or reduced price lunch	0.55	0.50
Designated as Limited English Proficent	0.11	0.32
Gifted	0.05	0.22
Disabled	0.13	0.34
FCAT-SSS Math Score	312	59
FCAT-SSS Reading Score	300	61
FCAT-NRT Math Score	639	41
FCAT-NRT Reading Score	642	43
DSS Math Score	1487	291
DSS Reading Score	1457	370
Principal Characteristics		
Percent of principals who are new to school	0.15	0.35
Principals' years of experience:		
As school administrator	10.86	6.91
In current job assignment	9.43	8.28
In current district	13.89	8.02
Teaching FL public schools	12.91	6.92
Teacher Characteristics		
Age	41	11
Male	0.07	0.26
Black	0.16	0.37
Hispanic	0.11	0.31
White	0.73	0.45
Having a professional certification	0.92	0.27
Ever certified in NBPTS	0.04	0.21
Passed FTCE Math on first attempt	0.55	0.50
Test scores in FTCE Math	192	41
Passed FTCE Reading on first attempt	0.76	0.43
Test scores in FTCE Reading	223	37
Years of experience teaching Florida public	10.4	9.3
Having an advanced degree	0.29	0.45
Having an advanced degree in Education	0.07	0.25
CLAST Math score	301	26
CLAST Reading score	311	25
SAT Verbal	458	90
SAT Quant	451	94
SAT Total	941	150
First major is education	0.94	0.24
Education major	0.94	0.23

Table 4.1 Summary of Teacher Characteristics by Student Characteristics and Preparation Levels

	Student Characteristics							
	Black	non-Black	Hispanic	non-Hispanic	FRPL	non-FRPL	LEP	non-LEP
Panel A: by student characterisitcs								
Black	0.28	0.13	0.20	0.15	0.20	0.11	0.20	0.16
Hispanic	0.09	0.12	0.31	0.05	0.15	0.07	0.36	0.08
Professional Certification	0.89	0.93	0.91	0.92	0.91	0.94	0.90	0.92
NBPTS Certification	0.03	0.05	0.04	0.05	0.04	0.06	0.03	0.05
Years of Experience	9.7	10.5	9.6	10.5	9.9	10.8	9.2	10.5
Lagged VAM-Math	0.052	0.091	0.091	0.079	0.071	0.094	0.091	0.081
Lagged VAM-Reading	0.043	0.082	0.068	0.075	0.055	0.094	0.065	0.074
Panel B: by student preparation	Math				English			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Black	0.22	0.17	0.14	0.11	0.22	0.18	0.15	0.11
Hispanic	0.15	0.11	0.10	0.09	0.15	0.09	0.08	0.08
Professional Certification	0.92	0.93	0.94	0.95	0.91	0.93	0.94	0.95
NBPTS Certification	0.04	0.04	0.05	0.06	0.04	0.04	0.05	0.06
Years of Experience	9.59	10.18	10.50	10.92	9.95	10.51	10.91	11.28
Lagged VAM-Math	0.0157	0.0184	0.0313	0.0552	0.0255	0.0155	0.0263	0.0525
Lagged VAM-Reading	0.0078	0.0135	0.0272	0.0570	0.0131	0.0098	0.0285	0.0564

Table 4.2 Within Teacher Variance in Student Prior Test Scores

	Lagged Math Score	Lagged Math Gain	Lagged Reading Score	Lagged Read Gain
Observed	0.6568	0.4394	0.6436	0.4538
Random assigned within schools	0.7774	0.3915	0.7875	0.4034
Perfect sorting within schools	0.0564	0.0645	0.077	0.0655
Random assigned across school	0.8975	0.4201	0.8941	0.4148
Perfect sorting across schools	0	0	0	0

1. Observed calculates the within teacher variance in student test scores based on the observed assignment of students to teachers

2.Random assignment sorts students to teachers within or across schools based on a randomly generated number from a uniform distribution. The random assignments are repeated 100 times before averaging across all teachers, all years, and all random assignments.

3. Perfect sorting assigns students to teachers based on students' rank in test score distribution.

4. Lagged scores are first available in 1998-99 and 1-year lagged gain becomes available in 2001-02.

Table 5.1 Effects of A/F Grade on Teachers' Probabilities of Being Assigned to Tested Grades, 1998-2004

	Getting an A grade				Getting an F grade			
	All	Novice teachers	Certified teachers	Had an advanced degree	All	Novice teachers	Certified teachers	Had an advanced degree
A*2002	0.0028	0.0008	0.0008	-0.0283				
	[0.0098]	[0.0220]	[0.0102]	[0.0206]				
F*2002					0.0386	0.1464	0.0222	-0.2595
					[0.1193]	[0.1173]	[0.1213]	[0.1609]
Teaching tested grade in prior yea	0.8712***	0.7939***	0.8753***	0.8753***	0.8712***	0.7940***	0.8753***	0.8753***
	[0.0023]	[0.0051]	[0.0022]	[0.0034]	[0.0023]	[0.0051]	[0.0022]	[0.0033]
Teacher Characteristics								
Black==1	-0.006	0.0045	-0.0103*	-0.0151	-0.006	0.0047	-0.0104*	-0.0152
	[0.0061]	[0.0125]	[0.0063]	[0.0109]	[0.0061]	[0.0124]	[0.0063]	[0.0109]
Hispanic==1	0.0171*	0.0353***	0.0131	0.0078	0.0171*	0.0350***	0.0131	0.0081
	[0.0092]	[0.0136]	[0.0095]	[0.0157]	[0.0092]	[0.0136]	[0.0095]	[0.0158]
Novice	-0.0136		-0.0150*	-0.0024	-0.0137		-0.0150*	-0.0022
	[0.0086]		[0.0089]	[0.0175]	[0.0086]		[0.0089]	[0.0174]
Had a professional certification	0.0300**	0.0199		0.0389	0.0300**	0.0199		0.0389
	[0.0149]	[0.0123]		[0.0383]	[0.0149]	[0.0123]		[0.0383]
Average years experience	-0.0026***		-0.0028***	-0.0042**	-0.0026***		-0.0028***	-0.0042**
	[0.0010]		[0.0010]	[0.0017]	[0.0010]		[0.0010]	[0.0017]
experience square	0		0	0.0001	0		0	0.0001
	[0.0000]		[0.0000]	[0.0000]	[0.0000]		[0.0000]	[0.0000]
Had an advanced degree	0.0122***	0.0291***	0.0121***		0.0122***	0.0291***	0.0121***	
	[0.0045]	[0.0108]	[0.0046]		[0.0045]	[0.0108]	[0.0046]	
Principal Characteristics								
New to school	-0.0005	0.0083	0.0014	-0.0038	-0.0006	0.0082	0.0013	-0.0036
	[0.0053]	[0.0126]	[0.0054]	[0.0103]	[0.0053]	[0.0126]	[0.0054]	[0.0103]
Experience as administrator	0.0017*	0.0029	0.0015*	0.0007	0.0017*	0.0029	0.0016*	0.0007
	[0.0009]	[0.0023]	[0.0009]	[0.0018]	[0.0009]	[0.0023]	[0.0009]	[0.0018]
Experience square	0	-0.0001	0	0	0	-0.0001	0	0
	[0.0000]	[0.0001]	[0.0000]	[0.0001]	[0.0000]	[0.0001]	[0.0000]	[0.0001]
# of teacher-year obs	123,693	23,184	119,486	37,915	123,693	23,184	119,486	37,915
Pseudo R-Sq	0.66	0.528	0.668	0.669	0.66	0.528	0.668	0.669

Notes:
1. All models include additional school control such as school demographics, student educational needs, and prior achievement and school year fixed effects
2. Standard errors account for school-year clutering
3. *** p<0.01, ** p<0.05, * p<0.1

Table 5.2 Impacts of Principal Characteristics on Teachers' Grade Assignment, 1998-2004

	All teachers			Novice teachers			Certified teachers			Had an advanced degree		
A*2002	0.0028	0.0136	0.0147	0.0008	0.0603*	0.0567*	0.0008	0.0077	0.0089	-0.0283	0.0014	0.007
	[0.0098]	[0.0136]	[0.0139]	[0.0220]	[0.0309]	[0.0318]	[0.0102]	[0.0145]	[0.0148]	[0.0206]	[0.0289]	[0.0290]
Princ Expr*A*2002		-0.0007	-0.0008		-0.0052**	-0.0050**		-0.0006	-0.0006		-0.0024	-0.0026
		[0.0008]	[0.0008]		[0.0021]	[0.0021]		[0.0009]	[0.0009]		[0.0017]	[0.0016]
New*A*2002			-0.0057			0.0189			-0.0065			-0.0304
			[0.0184]			[0.0490]			[0.0200]			[0.0394]
F*2002	0.0386	-0.2007	-0.1943	0.1464	0.3266***	0.3246***	0.0222	-0.2247	-0.2373	-0.2595	-0.4820*	-0.6763***
	[0.1193]	[0.1834]	[0.2991]	[0.1173]	[0.0487]	[0.0534]	[0.1213]	[0.1818]	[0.3045]	[0.1609]	[0.2497]	[0.0339]
Princ Expr*F*2002		0.019	0.0186		-0.0307**	-0.0301**		0.0209*	0.0216		0.0209	0.0530**
		[0.0120]	[0.0158]		[0.0134]	[0.0140]		[0.0121]	[0.0159]		[0.0191]	[0.0232]
New*F*2002			-0.0103			no sample			0.0192			no sample
			[0.2774]						[0.2819]			

Notes:
1. All models include additional school control such as school demographics, student educational needs, and prior achievement and school year fixed effects
2. Standard errors account for school-year clutering
3. *** p<0.01, ** p<0.05, * p<0.1

Table 6.1 Dif-in-dif estimates of the effects of an A grade on teacher classroom assignment, grade 3-5, 1998-2004

	Math				English			
	Black	Hispanic	FRPL	Avg Score	Black	Hispanic	FRPL	Avg Score
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All Teachers								
A*2002	-0.0005	0.0003	-0.001	0.0383	-0.0005	0.0003	-0.001	0.0037
	[0.0014]	[0.0012]	[0.0018]	[0.0217]*	[0.0014]	[0.0012]	[0.0018]	[0.0202]
New to School*A*2002	0.0013	0.0006	-0.001	0.0398	0.0013	0.0006	-0.001	0.061
	[0.0016]	[0.0017]	[0.0021]	[0.0288]	[0.0016]	[0.0017]	[0.0021]	[0.0271]**
Experience*A*2002	-0.0001	0	0.0001	0.0013	-0.0001	0	0.0001	0.0009
	[0.0001]	[0.0001]	[0.0001]	[0.0013]	[0.0001]	[0.0001]	[0.0001]	[0.0012]
Novice Teachers								
A*2002	0.0066	-0.0066	-0.0072	0.0718	0.0066	-0.0066	-0.0072	0.0614
	[0.0035]*	[0.0043]	[0.0053]	[0.0429]*	[0.0035]*	[0.0043]	[0.0053]	[0.0429]
New to School*A*2002	-0.001	0.0026	0.0054	0.0608	-0.001	0.0026	0.0054	0.0933
	[0.0041]	[0.0048]	[0.0073]	[0.0514]	[0.0041]	[0.0048]	[0.0073]	[0.0523]*
Experience*A*2002	-0.0004	0.0003	0.0002	0.0011	-0.0004	0.0003	0.0002	-0.0019
	[0.0002]**	[0.0002]	[0.0003]	[0.0027]	[0.0002]**	[0.0002]	[0.0003]	[0.0028]
Teachers w/ professional certification								
A*2002	-0.0006	0.0015	-0.0004	0.0298	-0.0006	0.0015	-0.0004	-0.0054
	[0.0015]	[0.0013]	[0.0020]	[0.0219]	[0.0015]	[0.0013]	[0.0020]	[0.0204]
New to School*A*2002	0.0008	0.0009	-0.0018	0.0393	0.0008	0.0009	-0.0018	0.066
	[0.0017]	[0.0019]	[0.0023]	[0.0293]	[0.0017]	[0.0019]	[0.0023]	[0.0283]**
Experience*A*2002	-0.0001	0	0.0001	0.0012	-0.0001	0	0.0001	0.0013
	[0.0001]	[0.0001]	[0.0001]	[0.0013]	[0.0001]	[0.0001]	[0.0001]	[0.0011]
Teachers w/ advanced degree								
A*2002	-0.0005	-0.0002	0.0032	0.051	-0.0005	-0.0002	0.0032	0.0279
	[0.0029]	[0.0031]	[0.0043]	[0.0329]	[0.0029]	[0.0031]	[0.0043]	[0.0303]
New to School*A*2002	0.0004	0.0092	0.0023	0.0567	0.0004	0.0092	0.0023	0.0973
	[0.0031]	[0.0042]**	[0.0056]	[0.0458]	[0.0031]	[0.0042]**	[0.0056]	[0.0415]**
Experience*A*2002	-0.0002	0.0001	-0.0002	0.0001	-0.0002	0.0001	-0.0002	0.0006
	[0.0002]	[0.0002]	[0.0002]	[0.0019]	[0.0002]	[0.0002]	[0.0002]	[0.0017]
Q1								
(Teachers at 1st Quartile of VAM)								
A*2002	-0.0124	-0.0069	-0.0051	-0.0067	-0.0083	-0.0037	-0.0222	0.0886
	[0.0082]	[0.0082]	[0.0116]	[0.0703]	[0.0084]	[0.0082]	[0.0121]*	[0.0812]
New to School*A*2002	-0.0007	-0.0069	-0.0114	0.2462	-0.0038	0.0012	-0.0088	0.0412
	[0.0089]	[0.0107]	[0.0161]	[0.0872]***	[0.0096]	[0.0129]	[0.0187]	[0.0927]
Experience*A*2002	0.0004	0.0003	0.0003	0.0009	-0.0001	0.0003	0.0014	-0.0038
	[0.0005]	[0.0005]	[0.0007]	[0.0043]	[0.0005]	[0.0005]	[0.0007]*	[0.0051]
Q2								
A*2002	0.0089	-0.0154	-0.0114	0.0036	0.0135	-0.008	-0.0036	-0.0171
	[0.0089]	[0.0098]	[0.0118]	[0.0639]	[0.0079]*	[0.0084]	[0.0113]	[0.0614]
New to School*A*2002	0.004	0.0185	-0.0007	-0.0325	0.002	0.0127	0.0071	0.0866
	[0.0088]	[0.0130]	[0.0181]	[0.0932]	[0.0084]	[0.0092]	[0.0157]	[0.0530]
Experience*A*2002	-0.0004	0.0003	0.0005	0.0016	-0.0003	0.0001	0.0004	-0.0046
	[0.0004]	[0.0004]	[0.0007]	[0.0037]	[0.0004]	[0.0004]	[0.0006]	[0.0033]
Q3								
A*2002	0.0093	0.0022	0.0206	0.0528	0.0107	0.0078	0.0133	-0.0772
	[0.0084]	[0.0086]	[0.0123]*	[0.0646]	[0.0088]	[0.0094]	[0.0118]	[0.0647]
New to School*A*2002	-0.0078	-0.0092	0.0161	0.1045	-0.0032	0.0012	-0.0023	0.011
	[0.0075]	[0.0088]	[0.0139]	[0.0616]*	[0.0084]	[0.0107]	[0.0154]	[0.0691]
Experience*A*2002	-0.0006	-0.0003	-0.0011	-0.0005	-0.0005	-0.0002	0.0002	-0.0005
	[0.0004]	[0.0004]	[0.0007]	[0.0034]	[0.0004]	[0.0004]	[0.0006]	[0.0031]
Q4								
A*2002	0.021	-0.0135	-0.0031	0.0221	0.0132	-0.0056	0.0272	-0.0184
	[0.0086]**	[0.0099]	[0.0120]	[0.0941]	[0.0085]	[0.0087]	[0.0115]**	[0.0858]
New to School*A*2002	-0.0059	0.0079	0.0147	-0.001	-0.0099	0.0215	-0.002	0.0187
	[0.0101]	[0.0099]	[0.0154]	[0.1137]	[0.0082]	[0.0102]**	[0.0151]	[0.0920]
Experience*A*2002	-0.0007	0.0008	-0.0004	-0.005	-0.0003	0.0004	-0.0012	0.0048
	[0.0004]*	[0.0004]*	[0.0006]	[0.0048]	[0.0004]	[0.0004]	[0.0007]*	[0.0044]

Note:
1. All standard errors account for the estimation of value-added scores at the first stage and school-year clustering
2. All models include teacher qualifications, principal characteristics, school covariates and school year fixed effects
3. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 6.2 Dif-in-dif estimates of the effects of an F grade on teacher classroom assignment, grade 3-5, 1998-2004

	Math				English			
	Black	**Hispanic**	**FRPL**	**Avg Score**	**Black**	**Hispanic**	**FRPL**	**Avg Score**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
All Teachers								
F*2002	0.009	-0.0098	0.014	0.2987	0.009	-0.0098	0.014	0.0872
	[0.0106]	[0.0068]	[0.0130]	[0.0950]***	[0.0106]	[0.0068]	[0.0130]	[0.1304]
New to School*F*2002	-0.0229	0.0134	-0.0182	-0.0475	-0.0229	0.0134	-0.0182	0.2339
	[0.0099]**	[0.0054]**	[0.0118]	[0.1421]	[0.0099]**	[0.0054]**	[0.0118]	[0.1258]*
Experience*F*2002	-0.0002	-0.0001	-0.0007	-0.0187	-0.0002	-0.0001	-0.0007	-0.0058
	[0.0007]	[0.0003]	[0.0006]	[0.0039]***	[0.0007]	[0.0003]	[0.0006]	[0.0083]
Novice Teachers								
F*2002	-0.0081	-0.0103	0.0249	0.1344	-0.0081	-0.0103	0.0249	-0.056
	[0.0102]	[0.0131]	[0.0069]***	[0.0740]*	[0.0102]	[0.0131]	[0.0069]***	[0.0604]
New to School*F*2002	0.0031	0.0181	0.0228	-0.2602	0.0031	0.0181	0.0228	0.1104
	[0.0075]	[0.0096]*	[0.0105]**	[0.2327]	[0.0075]	[0.0096]*	[0.0105]**	[0.0518]**
Experience*F*2002	0.0006	0.0008	-0.0013	-0.0108	0.0006	0.0008	-0.0013	-0.0094
	[0.0005]	[0.0008]	[0.0004]***	[0.0052]**	[0.0005]	[0.0008]	[0.0004]***	[0.0037]**
Teachers w/ professional certification								
F*2002	0.0168	-0.0071	0.0202	0.3862	0.0168	-0.0071	0.0202	0.1905
	[0.0094]*	[0.0061]	[0.0173]	[0.1806]**	[0.0094]*	[0.0061]	[0.0173]	[0.2411]
New to School*F*2002	-0.0306	0.01	-0.026	-0.1377	-0.0306	0.01	-0.026	0.1527
	[0.0101]***	[0.0055]*	[0.0154]*	[0.1905]	[0.0101]***	[0.0055]*	[0.0154]*	[0.2177]
Experience*F*2002	-0.0007	-0.0001	-0.0009	-0.0224	-0.0007	-0.0001	-0.0009	-0.0084
	[0.0005]	[0.0002]	[0.0007]	[0.0070]***	[0.0005]	[0.0002]	[0.0007]	[0.0108]
Teachers w/ advanced degree								
F*2002	0.0165	-0.0213	-0.021	1.3718	0.0165	-0.0213	-0.021	1.8388
	[0.0151]	[0.0080]***	[0.0084]**	[0.0862]***	[0.0151]	[0.0080]***	[0.0084]**	[0.2711]***
New to School*F*2002	-0.0129	0.0217	-0.01	-0.7099	-0.0129	0.0217	-0.01	-1.0499
	[0.0117]	[0.0041]***	[0.0041]**	[0.0459]***	[0.0117]	[0.0041]***	[0.0041]**	[0.2050]***
Experience*F*2002	0.0002	0.0013	0.001	-0.0748	0.0002	0.0013	0.001	-0.1208
	[0.0008]	[0.0003]***	[0.0003]***	[0.0040]***	[0.0008]	[0.0003]***	[0.0003]***	[0.0162]***
Q1								
(Teachers at 1st Quartile of VAM)								
F*2002	0.0054	-0.0113	-0.0205	1.0128	0.0294	-0.0294	0.1166	1.263
	[0.0159]	[0.0136]	[0.0203]	[0.1116]***	[0.0563]	[0.0346]	[0.0788]	[0.1607]***
New to School*F*2002	-0.0159	0.018	0.0053	-0.2867	-0.075	0.0545	-0.1257	-0.3904
	[0.0106]	[0.0107]*	[0.0137]	[0.0774]***	[0.0565]	[0.0342]	[0.0751]*	[0.0836]***
Experience*F*2002	0.0023	-0.0017	0.0033	-0.0516	-0.0035	0.0022	-0.0069	-0.0865
	[0.0007]***	[0.0008]**	[0.0009]***	[0.0053]***	[0.0034]	[0.0021]	[0.0046]	[0.0058]***
Q2								
F*2002	0	0	0	0.3151	0	0	0	0
	[0.0000]	[0.0000]	[0.0000]	[0.3371]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
New to School*F*2002	-0.0325	0.0282	-0.1255	0	0.0092	0.0334	-0.0054	0.5575
	[0.0176]*	[0.0133]**	[0.0247]***	[0.0000]	[0.0199]	[0.0145]**	[0.0148]	[0.2467]**
Experience*F*2002	0.0036	-0.0024	0.0248	-0.3502	0.0005	0.0017	-0.0022	0.092
	[0.0018]*	[0.0017]	[0.0022]***	[0.0119]***	[0.0025]	[0.0020]	[0.0019]	[0.0310]***
Q3								
F*2002	0.0173	-0.0108	0.0861	-0.2087	-0.0465	-0.0074	-0.0023	-0.1936
	[0.0463]	[0.0226]	[0.0422]**	[0.1535]	[0.0095]***	[0.0126]	[0.0263]	[0.1345]
New to School*F*2002	-0.0227	0.0352	-0.2047	0.5788	0	0	0	0
	[0.0155]	[0.0175]**	[0.0194]***	[0.1246]***	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Experience*F*2002	-0.0015	0.0019	-0.0091	0.0016	0	0	0	0
	[0.0013]	[0.0014]	[0.0016]***	[0.0101]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Q4								
F*2002	-0.0919	-0.0168	0.0454	-1.0329	-0.1131	0.0413	0.0286	-1.4327
	[0.0212]***	[0.0274]	[0.0226]**	[0.1825]***	[0.0344]***	[0.0212]*	[0.0452]	[0.1676]***
New to School*F*2002	0.0098	0.0302	0.0018	-0.1663	-0.0112	0.0358	-0.0245	1.1184
	[0.0133]	[0.0251]	[0.0131]	[0.1187]	[0.0075]	[0.0080]***	[0.0105]**	[0.0923]***
Experience*F*2002	0.0149	-0.0073	-0.0064	0.0733	0.0186	-0.012	-0.0058	0.0667
	[0.0020]***	[0.0036]**	[0.0019]***	[0.0160]***	[0.0016]***	[0.0020]***	[0.0021]***	[0.0182]***

Note:
1. All standard errors account for the estimation of value-added scores at the first stage and school-year clustering
2. All models include teacher qualifications, principal characteristics, school covariates and school year fixed effects
3. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 7.1 Difference-in-difference estimates of the effects of an A grade on teacher classroom assignment, grade 1-2, 1998-2004

	Black (1)	Hispanic (2)	FRPL (3)	Avg Score (4)
All Teachers				
A*2002	0.0049	-0.0009	0.0071	-0.0009
	[0.0022]**	[0.0024]	[0.0029]**	[0.0049]
New to School*A*2002	-0.0006	0.0022	0.0008	0.0063
	[0.0025]	[0.0039]	[0.0044]	[0.0074]
Experience*A*2002	-0.0001	-0.0001	-0.0004	-0.0001
	[0.0001]	[0.0001]	[0.0002]**	[0.0002]
Novice Teachers				
A*2002	0.009	-0.0133	0.001	-0.0188
	[0.0055]*	[0.0067]**	[0.0075]	[0.0143]
New to School*A*2002	-0.0022	0.0128	0.0042	0.0103
	[0.0065]	[0.0093]	[0.0116]	[0.0217]
Experience*A*2002	-0.0002	0.0001	-0.0001	0.0003
	[0.0003]	[0.0003]	[0.0005]	[0.0008]
Teachers w/ professional certification				
A*2002	0.0049	0.0005	0.0076	0.001
	[0.0024]**	[0.0026]	[0.0030]***	[0.0050]
New to School*A*2002	-0.0012	0.003	0.0015	0.0093
	[0.0027]	[0.0041]	[0.0044]	[0.0076]
Experience*A*2002	-0.0001	-0.0002	-0.0004	-0.0002
	[0.0001]	[0.0001]*	[0.0002]**	[0.0002]
Teachers w/ advanced degree				
A*2002	0.0066	-0.0008	0.0099	0.0026
	[0.0046]	[0.0058]	[0.0061]	[0.0091]
New to School*A*2002	-0.0019	0	-0.0004	0.0113
	[0.0043]	[0.0079]	[0.0085]	[0.0147]
Experience*A*2002	-0.0002	-0.0002	-0.0005	-0.0006
	[0.0002]	[0.0003]	[0.0003]	[0.0004]

Note:
1. All standard errors account for school-year clustering
2. All models include teacher qualifications, principal characteristics, school covariates and school year fixed effects
3. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 7.2 Difference-in-difference estimates of the effects of an F grade on teacher classroom assignment, grade 1-2, 1998-2004

	Black (1)	Hispanic (2)	FRPL (3)	Avg Score (4)
All Teachers				
F*2002	0.002	-0.0062	-0.02	-0.0387
	[0.0482]	[0.0432]	[0.0062]***	[0.0421]
New to School*F*2002	0.0308	-0.0064	0.054	0.0277
	[0.0427]	[0.0379]	[0.0053]***	[0.0368]
Experience*F*2002	-0.0025	0.0031	0.0009	0.002
	[0.0030]	[0.0024]	[0.0003]***	[0.0019]
Novice Teachers				
F*2002	0.0719	-0.053	-0.0158	-0.0098
	[0.0424]*	[0.0509]	[0.0077]**	[0.0481]
New to School*F*2002	0	0	0	0
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Experience*F*2002	-0.0088	0.008	0.0004	0.0001
	[0.0025]***	[0.0030]***	[0.0004]	[0.0028]
Teachers w/ professional certification				
F*2002	-0.0129	0.0077	-0.0215	-0.0266
	[0.0728]	[0.0683]	[0.0062]***	[0.0650]
New to School*F*2002	0.0493	-0.0245	0.0532	0.0118
	[0.0636]	[0.0597]	[0.0053]***	[0.0567]
Experience*F*2002	-0.0028	0.0035	0.0009	0.003
	[0.0043]	[0.0038]	[0.0003]***	[0.0033]
Teachers w/ advanced degree				
F*2002	-0.0225	-0.0048	-0.0108	-0.0804
	[0.0774]	[0.0650]	[0.0071]	[0.0728]
New to School*F*2002	0	0	0	0
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Experience*F*2002	-0.0003	0.0019	0.0006	0.0034
	[0.0044]	[0.0036]	[0.0003]*	[0.0035]

Note:
1. All standard errors account for school-year clustering
2. All models include teacher qualifications, principal characteristics, school covariates and school year fixed effects
3. * significant at 10%; ** significant at 5%; *** significant at 1%

Appendix

10.1 Asymptotic distribution of the 2SLS estimator with generated regressors and generated instruments

The basic model is:

$$y_{it} = x_{it}\beta + u_{it}$$

where x_{it} is $1 \times K$, $\beta\, K \times 1$. Let β be the 2-SLS estimator from:

$$y_{it} = \hat{x}_{it}\beta + e_{it}$$

where $\hat{x}_{it} = f(w,\delta)$, δ is a $Q \times 1$vector. Additionally, let z_{it} be the $1 \times L$ instrument vector with the instruments for each i are $\hat{z}_{it} = g\left(v_{it,}\hat{\lambda}\right)$ and λ an $S \times 1$vector of parameters. The pooled 2SLS estimator $\hat{\beta}$, after plugging the first-stage estimates is:

$$\hat{\beta} = \left[\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{x}'_{it}\hat{z}_{it}\right)\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}\hat{z}_{it}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}\hat{x}_{it}\right)\right]$$
$$\times\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{x}'_{it}\hat{z}_{it}\right)\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}\hat{z}_{it}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}y_{it}\right)$$

write $y_{it} = \hat{x}_{it}\beta + (x_{it} - \hat{x}_{it})\beta + u_{it}$, and after substitution:

$$\sqrt{N}\left(\hat{\beta} - \beta\right) = \left[\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{x}'_{it}\hat{z}_{it}\right)\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}\hat{z}_{it}\right)^{-1}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}\hat{x}_{it}\right)\right]^{-1}$$
$$\times\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{x}'_{it}\hat{z}_{it}\right)\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}\hat{z}_{it}\right)^{-1}$$
$$\times\left(N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}\left[\left(x_{it} - \hat{x}_{it}\right)\beta + u_{it}\right]\right)$$
$$= \left(\hat{C}'\hat{D}^{-1}\hat{C}\right)^{-1}\hat{C}'\hat{D}^{-1}\left\{N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}\left[\left(x_{it} - \hat{x}_{it}\right)\beta + u_{it}\right]\right\}$$

where $\hat{C} \equiv N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}_{it}'x_{it}$ and $\hat{D} \equiv N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}_{it}'z_{it}$. By uniform weak law of large numbers (UWLLN), $\hat{C}\xrightarrow{P}E(z'x)$ and $\hat{D}\xrightarrow{P}E(z'z)$. If $g(\cdot)$ is twice-differentiable, the mean value expansion of $N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}_{it}'u_{it}$ gives:

$$N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}_{it}'u_{it} = N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}z_{it}'u_{it} + \underbrace{\left[N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\nabla_{\lambda}g(v_{it},\lambda)u_{it}\right]}_{=o_p(1)}\underbrace{\sqrt{N}\left(\hat{\lambda}-\lambda\right)}_{=O_p(1)} + o_p(1)$$

where $\nabla_{\lambda}g(v_{it},\lambda)$ is the $L\times S$ Jacobian of $g(v_{it},\lambda)'$ and we have used the $E(u_{it}\mid v_{it})=0,\ E\left[\nabla_{\lambda}g(v_{it},\lambda)'u_{it}\right]=0$. Finally, since $o_p(1)\cdot O_p(1)=o_p(1)$, we have:

$$N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}_{it}'u_{it} = N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}z_{it}'u_{it} + o_p(1)$$

Similarly, we have:

$$N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}_{it}'(x_{it}-\hat{x}_{it})\beta = -\left[N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(\beta\otimes z_{it})'\nabla_{\delta}f(w_{it},\delta)\sqrt{N}\left(\hat{\delta}-\delta\right)+o_p(1)\right] = -G\sqrt{N}\left(\hat{\delta}-\delta\right)+o_p(1)$$

where $G=E\left[(\beta\otimes z_{it})'\nabla_{\delta}f(w_{it},\delta)\right]$ and $\nabla_{\delta}f(w_{it},\delta)$ is the $K\times Q$ Jacobian of $f(w_{it},\delta)'$.

Assume that

$\sqrt{N}\left(\hat{\delta}-\delta\right)=N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}r_{it}(\delta)+o_p(1)$ with $E[r_{it}(\delta)]=0$, which generally holds for most estimators (Wooldridge, 2002). Finally we have:

$$\sqrt{N}\left(\hat{\beta}-\beta\right)=\left(C'D^{-1}C\right)^{-1}C'D^{-1}\left\{N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[z_{it}'u_{it}-Gr_{it}(\delta)\right]\right\}+o_p(1)$$

$$\sqrt{N}\left(\hat{\beta}-\beta\right)\overset{a}{\sim}Normal\left[0,\left(C'D^{-1}C\right)^{-1}C'D^{-1}MD^{-1}C\left(C'D^{-1}C\right)^{-1}\right]$$

where $M=Var\left[z_{it}'u_{it}-Gr_{it}(\delta)\right]$

Consistent estimates of $A\text{var}\left[\sqrt{N}\left(\hat{\beta}-\beta\right)\right]$ is obtained by replacing unknown parameters with sample estimators:

$$\hat{C} \equiv N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}x_{it}$$

$$\hat{D} \equiv N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{z}'_{it}z_{it}$$

$$\hat{M} = N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{z}'_{it}\hat{u}_{it}-\hat{G}\hat{r}_{it}\right)\left(\hat{z}'_{it}\hat{u}_{it}-\hat{G}\hat{r}_{it}\right)'$$

$$\hat{G} = N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\hat{\beta}\otimes\hat{z}_{it}\right)'\nabla_{\delta} f\left(w_{it},\hat{\delta}\right)$$

$$\hat{r}_{it} = r_{it}\left(\hat{\delta}\right)$$

$$\hat{u}_{it} = y_{it} - x_{it}\hat{\beta}$$