

# Research Note: On Teacher Effect vs. *Other Stuff* in New Jersey's Growth Percentiles

**Bruce D. Baker**, Rutgers University, Graduate School of Education  
June 2, 2014

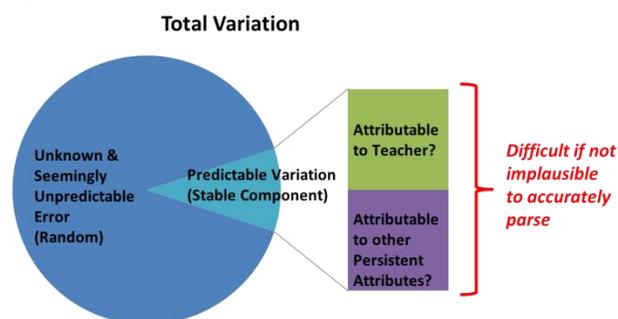
In this research note, I estimate a series of models to evaluate variation in New Jersey's school median growth percentile measures. These *measures of student growth* are intended by the New Jersey Department of Education to serve as measures of both school and teacher effectiveness. That is, the effect that teachers and schools have on marginal changes in their median student's test scores in language arts and math from one year to the next, all else equal. But all else isn't equal, and that matters greatly!

Variations in student test score growth estimates, generated either by value-added models or growth percentile methods, contain three distinct parts:

1. **"Teacher" effect:** Variations in changes in numbers of items answered correctly that may be fairly attributed to specific teaching approaches/ strategies/ pedagogy adopted or implemented by the child's teacher over the course of the school year;
2. **"Other stuff" effect:** Variations in changes in numbers of items answered correctly that may have been influenced by some non-random factor other than the teacher, including classroom peers, after school activities, health factors, available resources (class size, texts, technology, tutoring support), room temperature on testing days, other distractions, etc;
3. **Random noise:** Variations in changes in numbers of items answered correctly that are largely random, based on poorly constructed/asked items, child error in responding to questions, etc.

In theory, these first two types of variations are predictable. I often use a version of Figure 1 below when presenting on this topic. We can pick up variation in growth across classrooms, which is likely partly attributable to the teacher and partly attributable to *other stuff* unique to that classroom or school. The problem is, since the classroom (or school) is the unit of comparison, we really can't sort out what share is what?

**Figure 1**



We can try to sort out the variance by adding more background measures to our model, including student individual characteristics, student group characteristics, class sizes, etc., or by constructing more intricate analyses involving teachers who switch settings. But we can never really get to a point where we can be confident that we have correctly parsed that share of variance attributable to the teacher versus that share attributable to *other stuff*. And the most accurate, intricate analyses can rarely be applied to any significant number of teachers.

Thankfully, to make our lives easier, the New Jersey Department of Education has chosen not to try to parse the extent to which

variation in teacher or school median growth percentiles is influenced by *other stuff*. They rely instead on two completely unfounded, thoroughly refuted claims:

- 1) By accounting for prior student performance (measuring “growth” rather than level) they have fully accounted for all student background characteristics (refuted here<sup>1</sup>); and
- 2) Thus, any uneven distribution of growth percentiles, for example, lower growth percentiles in higher poverty schools, is a true reflection of the distribution of teacher quality (refuted here<sup>2</sup>).

In previous analyses I have explored predictors of New Jersey growth percentiles at the school level, including the 2012 and 2013 school reports. Among other concerns, I have found that the year over year correlation (across schools) between growth percentiles is only slightly stronger than the correlation between growth percentiles and school poverty.<sup>3</sup> That is, NJ SGPs tend to be about as correlated with *other stuff* as they are with themselves year over year. One implication of this finding is that even the year-over-year consistency is merely consistently measuring the *wrong effect* year over year. That is, the effect of poverty.

In the following models, I take advantage of a richer data set in which I have used a) school report card measures, b) school enrollment characteristics and c) detailed statewide staffing files and have combined those data sets into one, multi-year data set which includes outcome measures (SGPs and proficiency rates), enrollment characteristics

(low income shares, ELL shares) and resource measures derived from the staffing files.

Following are what I would characterize as exploratory regression models, using 3-years of measures of student populations, resources and school features, as predictors of 2012 and 2013 school median growth percentiles.

Resource measures include:

- a) **Competitiveness of wages:** a measure of how much teachers’ actual wages differ from predicted wages for all teachers in the same labor market (metro area) in the same job code, with the same total experience and degree level (estimated via regression model). This measure indicates the wage premium (>1.0) or deficit (<1.0) associated with working in a given school or district. This measure is constant across all same job code teachers across schools within a district. This measure is created using teacher level data from the fall staffing reports from 2010 through 2012.
- b) **Total certified teaching staff per pupil (staffing intensity):** This measure is created by summing the full time certified classroom teaching staff for each school and dividing by the total enrolled pupils. This measure is created using teacher level data from the fall staffing reports from 2010 through 2012.
- c) **% Novice teachers with only a bachelors’ degree:** This measure also focuses on classroom teachers, taking the number with fewer than 3 years of experience and only a bachelors’ degree and dividing by the total number of classroom teachers. This measure is created using teacher level data from the fall staffing reports from 2010 through 2012.

I have pointed out previously that it would be inappropriate to consider a teacher or school to be failing, or successful for that matter, simply because of the children they happen to serve. Estimate bias with respect to student population characteristics is a huge

<sup>1</sup>

<http://njedpolicy.wordpress.com/2013/05/02/deconstructing-disinformation-on-student-growth-percentiles-teacher-evaluation-in-new-jersey/>

<sup>2</sup>

<http://schoolfinance101.wordpress.com/2014/04/18/the-endogeneity-of-the-equitable-distribution-of-teachers-or-why-do-the-girls-get-all-the-good-teachers/>

<sup>3</sup>

<http://schoolfinance101.wordpress.com/2014/01/31/an-update-on-new-jerseys-sgps-year-2-still-not-valid/>

validity concern regarding the intended uses of New Jersey's growth percentile measures.

The potential influence of resource variations presents a comparable validity concern, though the implications vary by resource measure. If we find, for example that teachers receiving a more competitive wage are showing greater gains, we might assert that the wage differential offered by a given district is leading to a more effective teacher workforce. A logical policy implication would then be to provide resources to achieve wage premiums in schools and districts serving the neediest children, and otherwise lagging most on measures of student growth.

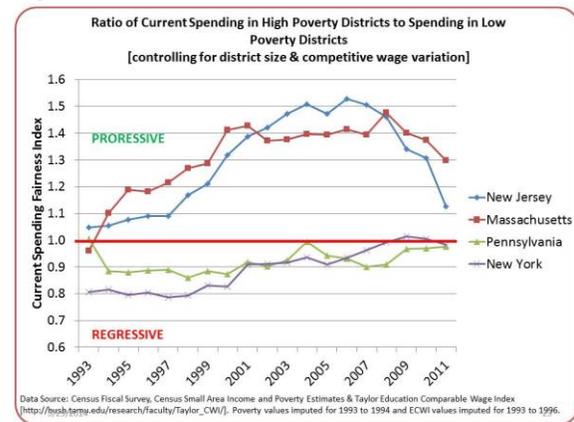
Of course, schools having more resources for use in one way – wages – also may have other advantages. If we find that overall staffing intensity is a significant predictor of student growth, it would be unfair to assert that the growth percentiles reflect teacher quality. That is, if growth in some schools is greater than in others because of more advantageous staffing ratios. Rather than *firing the teachers in the schools producing low growth*, the more logical policy response would be to provide those schools the additional resources to achieve similarly advantageous staffing ratios.

With these models, I also test assumptions about variations across schools within larger and smaller geographic areas – counties and cities. This geography question is important for a variety of reasons.

New Jersey is an intensely racially and socioeconomically segregated state. Most of that segregation occurs between municipalities far more so than within municipalities. That is, it is far more likely to encounter rich and poor neighboring school districts than rich and poor schools within districts. Yet education policy in New Jersey, like elsewhere, has taken a sharp turn toward reforms which merely reshuffle students and resources among schools (charter and district) within cities, pulling back significantly from efforts to target additional resources to high need settings.

Figure 2 shows that from the early 1990s through about 2005, New Jersey placed significant emphasis on targeting additional resources to higher poverty school districts. Since that time, New Jersey's school funding progressiveness has backslid dramatically. And these are the very resources needed for districts – especially high need districts – to provide wage differentials to recruit and retain a high quality workforce, coupled with sufficient staffing ratios to meet their students' needs.

**Figure 2**



## Findings

Table 1 shows the estimates from the first set of regression models which identify predictors of cross school and district, within county variation in growth percentiles. The four separate models are of language arts and math growth percentiles (school level) from the 2012 and 2013 school report cards. These models show that:

### Student Population *Other Stuff*

- % free lunch is significantly, negatively associated with growth percentiles for both subjects and both years. That is, schools with higher shares of low income children have significantly lower growth percentiles;
- When controlling for low income concentrations, schools with higher shares of English language learners

- have higher growth percentiles on both tests in both years;
- c) Schools with larger shares of children already at or above proficiency tend to show greater gains on both tests in both years;

- e) Schools with more full time classroom teachers per pupil have higher growth percentiles on both tests in both years.

**School Resource *Other Stuff***

- d) Schools with more competitive teacher salaries (at constant degree and experience) have higher growth percentiles on both tests in both years.

***Other Other Stuff***

- f) Charter schools have neither higher nor lower growth percentiles than otherwise similar schools in the same county.

**TABLE 1. Predicting within County, Cross School (cross district) Variation in New Jersey SGPs**

	LA SGP 2013		Math SGP 2013		LA SGP 2012		Math SGP 2012	
	Coef.	R.S.E.	Coef.	R.S.E.	Coef.	R.S.E.	Coef.	R.S.E.
<i>Student Background</i>								
Total % Proficient	9.660	2.166 *	8.978	3.219 *	11.217	2.685 *	4.751	3.243
% Free Lunch	-20.411	1.544 *	-15.496	2.205 *	-14.755	1.825 *	-19.062	2.288 *
% ELL	18.390	3.026 *	26.165	5.418 *	13.265	3.695 *	26.615	5.232 *
<i>Resource Levels</i>								
Wage Competitiveness	1.505	0.432 *	2.286	0.617 *	0.917	0.466 *	1.758	0.617 *
% Novice Teachers w/BA Only	1.558	2.450	-1.105	3.558	2.083	2.521	-1.960	3.689
Full Time Cert Teachers per Pupil	22.330	9.740 *	28.000	13.881 *	23.815	10.245 *	30.573	13.847 *
<i>Grade Range</i>								
Elem/Middle	1.602	0.533 *	0.797	0.829	0.752	0.577	1.030	0.769
Middle/High	-6.398	1.839 *	-7.614	2.066 *	-8.220	2.108 *	-9.923	2.480 *
Middle	-1.906	0.447 *	-1.324	0.606 *	-1.371	0.446 *	-1.075	0.588 **
K12 School	3.821	2.898	3.349	3.944	5.711	3.770	9.946	4.985 *
<i>Charter School</i>	3.014	1.953	3.029	3.145	2.063	1.939	4.168	3.106
<i>Constant</i>	46.574	2.415 *	46.857	3.626 *	41.884	2.813 *	45.941	3.629 *
<i>R-Squared</i>	0.409		0.210		0.324		0.218	

[1] Includes fixed effect for year of independent variables and for county of location (comparison is across schools within county)

[2] standard errors corrected for presence of multiple years of independent measures on each school (Robust Standard Errors, clustering on school ID)

\*p<.05, \*\*p<.10

**TABLE 2. Predicting within City Cross School (primarily within district) Variation in New Jersey SGPs**

	LA SGP 2013		Math SGP 2013		LA SGP 2012		Math SGP 2012	
	Coef.	R.S.E.	Coef.	R.S.E.	Coef.	R.S.E.	Coef.	R.S.E.
<i>Student Background</i>								
Total % Proficient	7.531	2.259 *	6.356	3.526 **	10.803	2.994 *	6.962	3.714 **
% Free Lunch	-14.864	2.103 *	-7.987	3.028 *	-9.232	2.557 *	-12.257	3.244 *
% ELL	6.934	3.331 *	11.559	5.726 *	7.419	4.397 **	20.487	5.275 *
<i>Resource Levels</i>								
Avg. Teacher Experience	-0.037	0.098	-0.126	0.136	-0.271	0.111 *	-0.196	0.138
% Novice Teachers w/BA Only	2.915	2.678	3.227	3.741	1.017	3.064	4.648	3.944
Full Time Cert Teachers per Pupil	-12.536	9.425	-30.022	15.031 *	-10.264	11.069	-13.685	13.991
<i>Grade Range</i>								
Elem/Middle	1.407	0.965	0.321	1.303	0.352	1.101	1.509	1.358
Middle/High	-6.029	1.726 *	-6.380	2.035 *	-7.153	2.057 *	-9.922	1.887 *
Middle	-2.215	0.405 *	-1.398	0.561 *	-1.660	0.416 *	-1.243	0.574 *
K12 School	4.289	4.967	6.940	5.652	6.809	5.777	12.097	6.051 *
<i>Charter</i>	2.844	2.003	2.346	3.352	-1.199	1.932	-1.298	3.286
<i>Constant</i>	37.175	2.473 *	52.959	3.757 *	39.203	3.027 *	48.503	3.851 *
<i>R-squared</i>	0.664		0.546		0.576		0.530	

[1] Includes fixed effect for year of independent variables and for CITY of location (comparison is across schools within CITY)

[2] standard errors corrected for presence of multiple years of independent measures on each school (Robust Standard Errors, clustering on school ID)

\*p<.05, \*\*p<.10

Table 2 includes a fixed effect for city location. That is, Table 2 runs the same regressions as in Table 1, but compares schools only against others in the same city. In most cases, because of municipal/school district alignment in New Jersey, comparing within the same city means comparing within the same school district. But, using city as the unit of analysis permits comparisons of district schools with charter schools in the same city.

In Table 2 we see that student population characteristics remain the dominant predictor of growth percentile variation. That is, across schools within cities, student population characteristics significantly influence growth percentiles. But the influence of demography on *destiny*, shall we say (as measured by SGPs), is greater across cities than within them, an entirely unsurprising finding. Resource variations within cities show few significant effects. Notably, our wage index measure does not vary within districts but rather across them and was replaced in these models by a measure of average teacher experience. Again, there was no significant difference in average growth achieved by charters than by other similar schools in the same city.

### Preliminary Policy Implications

The following preliminary policy implications may be drawn from the preceding regressions.

**Implication 1:** Because student population characteristics are significantly associated with SGPs, the SGPs are measuring differences in students served rather than, or at the very least in addition to differences in collective (school) teacher effectiveness. As such, it would simply be wrong to use these measures in any consequential way to characterize either teacher or school performance.

**Implication 2:** SGPs reveal positive effects of substantive differences in key resources, including staffing intensity and competitive wages. That is, resource availability matters and

teachers in settings with access to more resources are collectively achieving greater student growth. SGPs cannot be fairly used to compare school or teacher effectiveness across schools and districts where resources vary.

These findings provide support for a renewed emphasis on progressive distribution of school funding. That is, providing the opportunity for schools and districts serving higher concentrations of low income children and lower current growth, to provide the wage premiums and staffing intensity required to offset these deficits.<sup>4</sup>

**Implication 3:** The presence of stronger relationships between student characteristics and SGPs across schools and districts within counties, versus across schools within individual cities highlights the reality that between district (between city) segregation of students remains a more substantive equity concern than within city segregation of students across schools.

As such, policies which seek merely to reshuffle students across charter and district schools within cities and without attention to resources are unlikely to yield any substantive positive effect in the long run. In fact, given the influence of student sorting on the SGPs, sorting students within cities into poorer and less poor clusters will likely exacerbate within city achievement gaps.

**Implication 4:** The presence of significant resource effects across schools and districts within counties, but lack of resource effects across schools within cities, reveals that between district disparities in resources, coupled with sorting of students and families, remains a significant concern, and more substantive concern than within district inequities. Again, this finding supports a

---

<sup>4</sup> This finding also directly refutes the dubious assertion by NJDOE officials in their 2012 school funding report that the additional targeted funding was not only doing no good, but potentially causing harm and inducing inefficiency. <http://schoolfinance101.wordpress.com/2012/12/18/twisted-truths-dubious-policies-comments-on-the-njdoecerf-school-funding-report/>

renewed emphasis on targeting additional resources to districts serving the neediest children.

**Implication 5:** Charter schools do not vary substantively on measures of student growth from other schools in the same county or city when controlling for student characteristics and resources. As such, policies assuming that “chartering” in-and-of-itself (without regard for key resources) can improve outcomes are likely misguided. This is especially true where such policies do little more than reshuffle low and lower income minority students across schools within city boundaries.