WIDA™ RESEARCH REPORT

EXAMINING DISTRICT-LEVEL GROWTH USING ACCESS FOR ELLS

PREPARED BY

Narek Sahakyan, Wisconsin Center for Education Research,
University of Wisconsin–Madison

H. Gary Cook, Ph.D., Wisconsin Center for Education Research,
University of Wisconsin-Madison

AUGUST 2014

Wisconsin Center for Education Research
SCHOOL OF EDUCATION
UNIVERSITY OF WISCONSIN–MADISON

WIDA advances academic language development and academic achievement for linguistically diverse students. WIDA was formed as the result of a federal grant to comply with the requirements of the No Child Left Behind Act. It is a consortium of states and districts working together to promote achievement of English language learners. The organization has created a comprehensive system that includes English Language Development Standards, Spanish Language Development Standards, English language proficiency assessments, professional development for educators of ELLs, and research on all aspects of English language learning.

RESEARCH

The WIDA Research Department seeks to provide timely, meaningful, and actionable research that promotes educational equity and academic achievement for linguistically and culturally diverse students. Its annual research agenda is developed under the guidance of the WIDA Consortium Board Research Subcommittee and includes topics in the areas of academic language, standards, professional learning, and policy.

Introduction

The following report provides a description of a study examining school districts in the WIDA Consortium whose English language learners (ELLs) exhibit consistently high growth on the ACCESS for ELLs (ACCESS) assessment. The study is herein labeled *High-Flying Districts.* The study has the two-fold objective to:

I) Identify school districts within the WIDA consortium where *consistent and relatively higher (lower)* growth in English language proficiency is observed; and

II) Find *common and contrasting* characteristics among the identified *high-flying* (i.e., ELLs consistently gaining at high rates) and *low-cruising* (i.e., ELLs consistently gaining at low rates) school districts that potentially underlie this consistent and relatively higher (lower) growth.

This report is organized into two parts. The first part describes the measure of growth used in the study and the data used in the analysis. The second part presents the methodology and econometric model applied to identify high-flying and low-cruising school districts (HFDs and LCDs). It also presents results from analyses with identified HFDs and LCDs. To support a closer look at HFDs and LCDs, additional district-level variables from external sources are introduced. This allows a closer comparative look at identified high-flyers and low-cruisers.[1] Due to a lack of literature or empirical evidence on issues around ELL growth, we do not make the assumption that faster growth is a desirable, or conversely, that slower growth is an undesirable attribute for a school district. Instead, we make use of the large student- and district-level samples available through ACCESS and other sources and let the data speak for itself. The second part of the report focuses on identifying consistent patterns by visualizing the contrast between the demographic factors and academic outcomes for the HFD and LCD groups.

This project is a largely empirical exercise, examining growth from a research perspective to ultimately inform policymakers, administrators and practitioners about some statistical findings and patterns that emerge from contrasting various characteristics and outcomes of the fastest and slowest growing districts. This study does not examine growth in WIDA's districts for accountability purposes.[2] It is intended to serve as but one example of how to perform meaningful and reliable comparisons across school districts when using ACCESS for ELLs test scores as measures of ELL performance. However, the same methodology can easily be generalized to individual- , school-, and state-level comparisons, as well as extended to perform similar analyses on similar large-scale assessment instruments. We hope that the findings from this report will generate discussion and encourage similar analyses contributing to the very limited literature on the growth of English language proficiency.

---

[1] The external sources are public and include the National Center for Education Statistics (nces.ed.gov) and selected WIDA states' Department of Education web portals and accountability reporting systems.

[2] This report contains no identifiable district-, school- or individual-level information. This research has been approved by the University of Wisconsin-Madison Education and Social/Behavioral Science Institution Review Board (IRB) in accordance with federal regulations, state laws, and local and University policies.

Part 1: Identifying school districts within the WIDA consortium where consistent and relatively higher (and lower) growth is observed

Data and Descriptive Statistics

To generate the outcome variable in this study, ACCESS test scores were used to calculate annual district-level measures of ELL growth.[3] Even though ACCESS assesses multiple domains of English fluency (i.e., listening, speaking, reading, and writing), for the purposes of this study, overall Composite Scale Scores were chosen as the most simple, complete and representative measure for ELL's performance.[4] Using ELL test score data for the period of 2007–2012, five district-level aggregated Composite Scale Score gains (CSSGs) were calculated for the five adjacent growth cycles within this time period.[5] All districts were longitudinally connected for this period; however, because states joined the WIDA consortium at different times, not all districts had data for all five growth cycles. Table 1 presents the distribution of WIDA's school districts by state and growth cycle.

[3] The WIDA ACCESS for ELLs is a unique dataset that contains detailed test score and demographic information on over 1,500,000 ELs across 35 states.

[4] Vertical scaling is one of the useful properties of composite scale scores, making it possible to aggregate and compare students' outcomes across grades, districts and states. Namely, a composite scale score of 400 on ACCESS is constructed to reflect the same degree of English language proficiency, regardless of the level or grade of the students. For more details behind the calculation of composite scale scores, as well as other technical information on ACCESS, we refer the readers to the WIDA "Annual Technical Report for ACCESS for ELLs English Language Proficiency Test," Series 203, 2011–2012 Administration.

[5] Due to concerns with small and non-representative sample size, districts with less than 30 enrolled ELLs were excluded from the analysis. Additionally, while ACCESS data was also available for the 2012–2013 growth cycle, no district-level data was available for this period for the additional variables necessary to perform the secondary analysis. Therefore, the district sample was restricted to include only the 2007–2012 period.

Table 1
Distribution of WIDA School Districts by State and Growth Cycle

State	2007–2008	2008–2009	2009–2010	2010–2011	2011–2012
AL	57	60	61	59	57
DC	4	5	7	9	9
DE	16	15	15	15	15
GA	74	77	80	89	100
HI				16	16
IL		205	212	124	225
KY	23	29	34	37	42
ME	11	13	15	16	17
MO					78
MS			33	37	38
NC			96	100	100
ND	12	14	14	12	12
NH	12	11	12	13	13
NJ	126	122	127	126	144
NM				68	75
OK					4
PA	108	108	114	126	124
RI	17	19	15	17	19
SD			9	14	15
VA		16	58	60	64
VT	10	9	7	7	6
WI	100	106	108	111	116
WY				13	13
Total	**570**	**809**	**1017**	**1069**	**1302**

We start the analysis by looking at some basic facts about ELL growth in WIDA. Table 2 presents some descriptive statistics on district-level CSSGs by growth cycle.

Table 2
Descriptive Statistics on District-level Growth (CSSGs)

Growth Cycle	N (Districts)	Mean	Std. Dev.	Minimum	Maximum
2007–2008	570	26.99	6.85	5.13	50.02
2008–2009	809	27.78	6.70	3.67	49.95
2009–2010	1017	30.16	8.33	-7.59	58.83
2010–2011	1069	29.59	8.44	-8.34	63.98
2011–2012	1302	30.18	8.23	-13.38	76.09

Table 2 shows that the CSSGs exhibit an increasing trend, ranging from 27 to 30 within the 2007–2012 period. We also observe a higher variability in district-level CSSGs, as reflected in a decrease in the minimum, and increase in the maximum values for average district-level CSSGs in the later cycles as new states, districts, and ELLs join WIDA.

The first objective of this study is to find consistent and higher (lower) gains in district-level EL progress on ACCESS. To evaluate whether such gains exist and are sustained from one growth cycle to the next, we first assess the magnitude of district-level cross-year correlations for CSSGs. Table 3 presents the correlation coefficients between district-level CSSGs across the five growth cycles. Of particular interest are the correlation coefficients in boldface, indicating correlations for adjacent growth cycles.

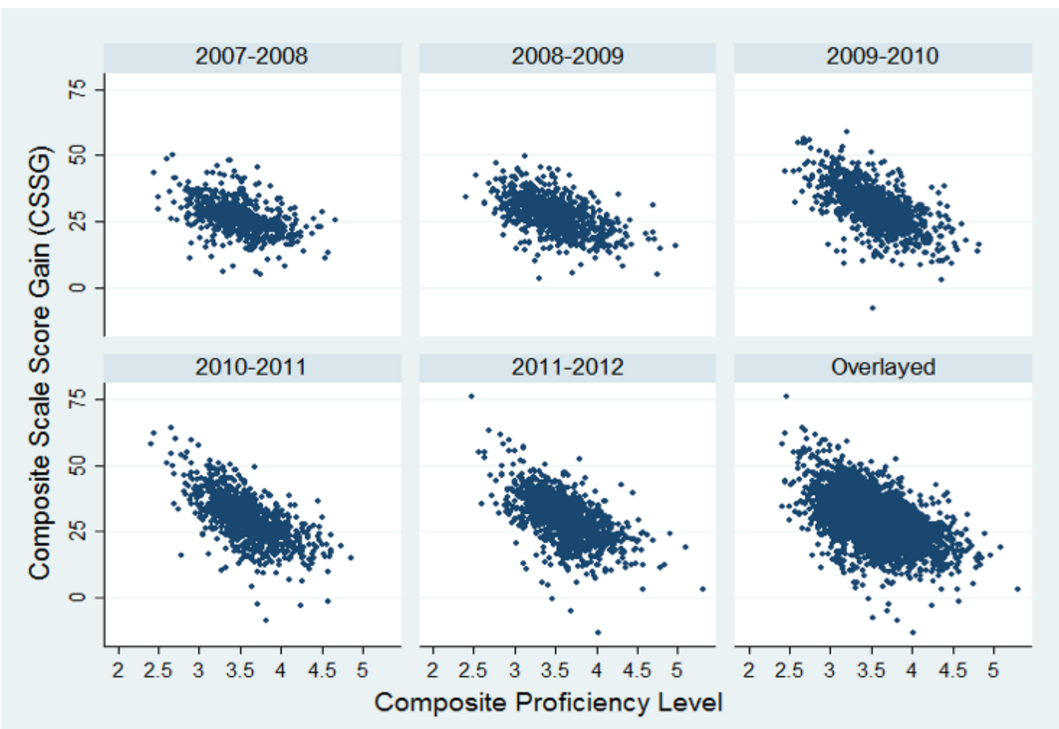
Table 3
Cross-cycle Correlation Coefficients for District-level CSSGs

Growth Cycle	2007–2008	2008–2009	2009–2010	2010–2011
2007–2008	1	-	-	-
2008–2009	**0.58**	1	-	-
2009–2010	0.52	**0.62**	1	-
2010–2011	0.57	0.58	**0.65**	1
2011–2012	0.55	0.57	0.63	**0.67**

The reported across-cycle correlation coefficients provide preliminary evidence that within this sample of school districts there are indeed some districts that are recording consistent and relatively faster (and slower) growth rates. In other words, generally if within a particular cycle a school district has performed well in terms of growth (high CSSG), we can expect it to perform similarly the next cycle. The data also shows that this trend increases with time, starting at 0.58 in the 2007–2008/2008–2009 growth cycle and reaching 0.67 in the 2010–2011/2011–2012 growth cycle.[6]

Since the first step of this research project calls for finding consistent outliers with respect to district-level growth, a quick, yet naïve, method to identify these high-flyers and low-cruisers would be to look for overlapping school districts within the top and bottom ends (25th and 75th percentiles, for example) of the CSSG distribution across the five growth cycles. However, this approach is overly simplistic and is complicated by the districts' heterogeneity in their respective ELL populations. Specifically, the literature on English language proficiency assessments has illustrated that it is increasingly harder to achieve growth in higher-grade and higher-proficiency ELLs. Putting it differently, when it comes to growth in ELL performance as defined by ACCESS test scores, *lower is faster, higher is slower* (LFHS).[7] Figures 1.1 and 1.2 present the distribution of CSSGs by districts' average proficiency level and grade, providing some preliminary visual evidence in support for the LFHS rule. The data reveals a consistent, downward-sloping trend in the distribution, which gets more defined in the later cycles. This is indication that the starting proficiency level and grade are important factors in evaluating ELL growth at the district level as well.

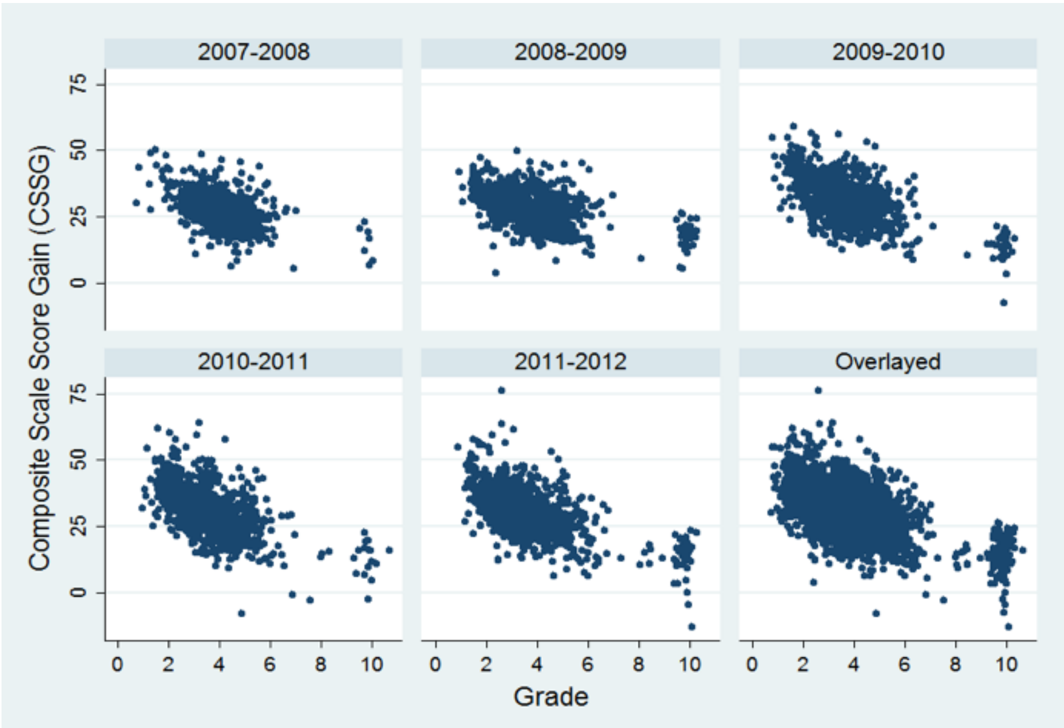
Figure 1.1. District-level growth (CSSG) by proficiency level.



[6] This potentially indicates the presence of a first order auto-regression, or intuitively speaking – an annual *inertia* effect. The autoregressive structure of district-level growth is addressed explicitly in the estimation of the econometric model.

[7] While this may be true for other large-scale assessment instruments, Cook, Boals, Wilmes and Santos (2008) and Cook and Zhao (2011) provide evidence for LFHS for ACCESS data at the individual student level.

Figure 1.2. District-level growth (CSSG) by grade.



Figures 1.1 and 1.2 make it clear that differences in the districts' starting points at the beginning of each growth cycle, i.e. the districts' ELL composition with respect to proficiency level and grade should not be neglected. Otherwise, despite the vertical scaling, a simple comparison of CSSGs would favor districts that happen to serve a higher proportion of lower-proficiency and lower-grade ELLs. In order to take into account the proficiency level and grade of ELLs, we need to explicitly include these variables into any measure assessing ELL growth.[8] Given the exploratory nature of this study, the longitudinally connected structure of the data, and the need to control for multiple variables, an ordinary least squares (OLS) fixed-effects regression model is applied, often used in similar repeated-observation type frameworks.[9] The next section provides a brief description of the model used to derive an ELL growth measure that is free of the distorting effect of LFHS.

[8] An alternative, yet very involved solution to the LFHS issue is to perform the analysis separately for different grades and levels.

[9] For more details behind the estimation of the model please consult Greene (2011), Raudenbush and Bryk (2002), or any standard regression procedure textbook.

The Fixed-Effects Model

Formally, the estimating model is characterized by Equation 1 below:

(1) $y_{it} = X_{it}\beta + d_i + e_{it}$, *for t = 1,…,T and I = 1, …, N, where*

- y_{it} is district i's composite scale score gain in growth cycle t (i= 1, …, 1302; t = 1,…5)
- X_{it} is a matrix of independent variables, i.e. district i's average *composite proficiency level, grade* and *size* in growth cycle t (as measured by the number of enrolled ELLs, included to control for potential size effects)
- β is a vector of coefficients to be estimated by the model, capturing the effect of c*omposite proficiency level, grade* and *size* on growth. Since the model is linear, these β coefficients can be directly linked to the downward slope of the distribution of CSSGs as presented in Figures 1.1 and 1.2.
- d_i is the unobserved time-invariant individual effect for district i (e.g. historical and institutional factors). Intuitively, di can also be thought of as district dummy variables. The underlying assumption for a fixed-effects versus a random-effects model is that the time-invariant district effect d_i is correlated with (is not independent of) X_{it}
- e_{it} is the district- and time- dependent error term, assumed to be *iid* (independent and identically distributed), normally distributed, with a mean of zero and a homoscedastic distribution.[10]

The assumption of nonzero correlation between d_i and X_{it} implies that the district's time-invariant (fixed) effect is not independent of the district's average proficiency level, grade and size. In other words, if there is something institutional about the school district that is fixed across time and that correlates with high (low) growth, the assumption is that it would also be correlated with that district's starting point with respect to its ELLs' average proficiency and grade, as well as the district size. Generally speaking, the random-effects assumption is a stricter one, as it implies that district-level time-invariant factors are completely independent of the time-varying district characteristics.[11]

Intuitively, by applying the model specified in Equation (1), we are assessing whether variations in district-averaged starting composite proficiency level (CPL) and grade around each district's mean value (across time) are related to variations in composite scale score gains around each district's mean value (across time). By including district fixed-effects, we control for the average differences across districts in both observable (CPL, grade and district size) and unobservable predictors. These fixed-effect coefficients (d_i) absorb all the across-district heterogeneity, leaving only the within-district variation for the predictors to explain, thereby greatly reducing the potential threat of omitted-variable bias.[12]

[10] The latter assumption is relaxed during the model estimation via heteroskedastisity-robust standard errors.

[11] The assumption of fixed effects versus random effects is tested and confirmed in favor of the fixed-effects method using the Hausman (1978) specification test

[12] Allison, P. D. (2005). Fixed effects regression methods for longitudinal data using SAS. SAS Press.

A limitation of the fixed-effects model is that it cannot accurately assess the effect of variables that have little or no within-group (district) variation. For example, if we were interested in whether district-level institutional factors, such as teacher/pupil ratios, SES, funding, program and/or school type, were statistically significant predictors of district-level growth, we would not be able to reliably estimate these using the fixed-effects regression approach, because it is natural to expect that most, if not all of the variation for these variables would be across, and not within districts. For this reason, as well as due to the difficulties associated with collecting secondary data for all the WIDA districts, the analysis for some of these variables is conducted separately for a selective sample of districts (HFDs and LCDs) using district-level predicted CSSGs, which have been adjusted for the effect of initial CPL, grade and district size via the fixed-effects model parameter estimates. The next section provides the main results from the estimation of the model.

Estimation Results

The estimation of Equation 1 was performed using the PROC MIXED statement in SAS and replicated in STATA using *xtregar.*[13] Table 4 presents the estimation results.

Table 4
Estimation of the Effect of District-level Proficiency and Grade on Growth (LFHS)

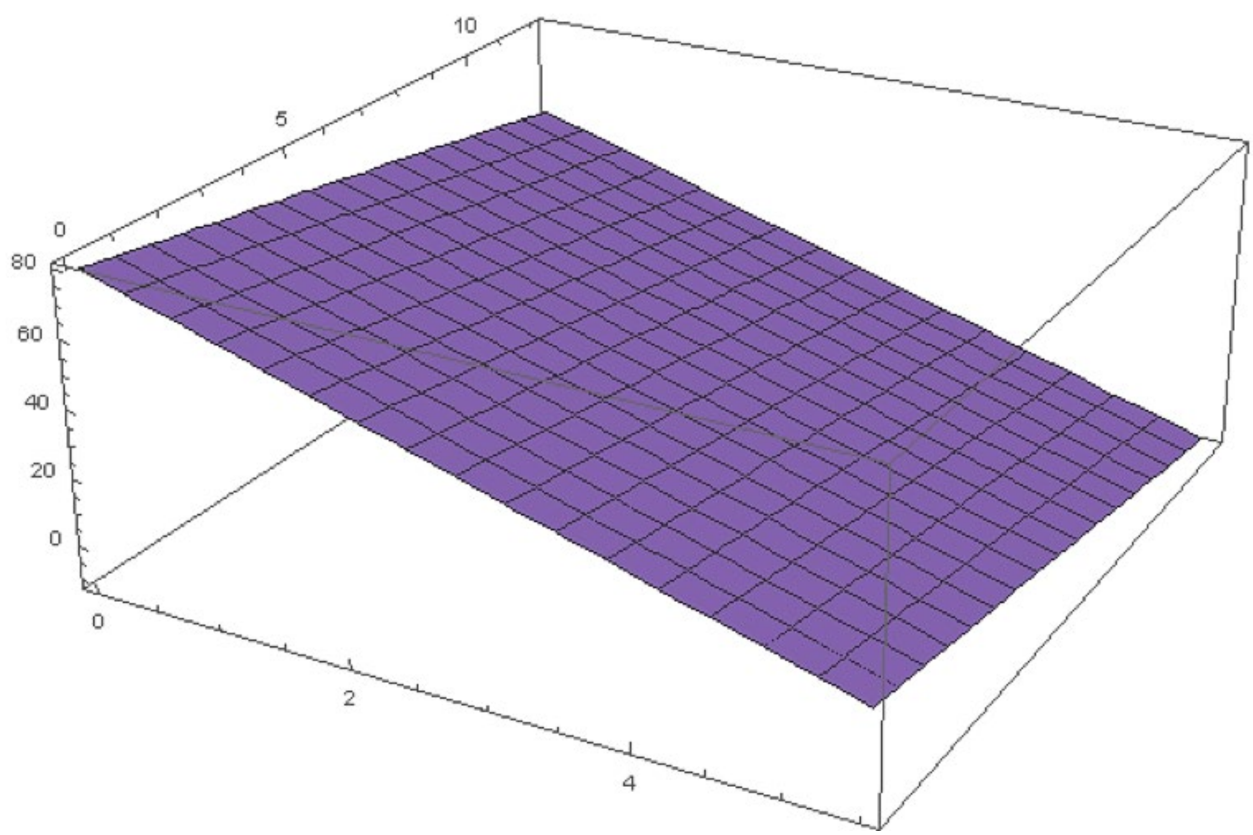
```
FE (within) regression with AR(1) disturbances  Number of obs      =      3283
Group variable: distnumber                      Number of groups   =      1186

R-sq:  within  = 0.3867                         Obs per group: min =         1
       between = 0.5653                                        avg =       2.8
       overall = 0.5163                                        max =         4

                                                F(3,2094)          =    440.07
corr(u_i, Xb)  = -0.4988                        Prob > F           =    0.0000
```

Effect on CSSG	Coefficient	Standard Error	t Value	Pr> |t|
Composite Proficiency Level	-11.59	0.28	-41.49	<.01
Grade	-2.41	0.07	-33.66	<.01
Number of ELLs	-0.0002	0.00	-2.47	0.01
Intercept	79.52	0.97	82.05	<.01

[13] While the model parameters differed slightly using SAS or STATA, the set of identified *high-flying* and *low-cruising* districts obtained was identical using either method. Appendix A presents the SAS and STATA code used to perform the estimation.

As indicated by the t-values reported in Table 4, the estimated coefficients for CPL and Grade are highly statistically significant. The negative sign of the parameters confirms the *Lower is Faster, Higher is Slower* principle at the district-level.[14] Intuitively, given the parameter estimates reported in Table 4, a one-point increase in the district-level starting CPL (thus the negative sign) on average will lower the districts' expected CSSG by 11.59 composite scale score points. Similarly, if a district's ELL composition increased by one unit with respect to the district's average starting grade, the district's growth would decrease by 2.41 CSSGs. The estimated correlation of - 0.49 between the district fixed-effect and *predicted CSSG* ($X\beta$) is evidence in favor of our selection of the fixed-effects versus the random-effects model. Based on the coefficients estimated by Equation 1, Figure 2 below presents a three-dimensional representation of *LFHS* with respect to district's starting CPL and grade.

Figure 2. District-level growth: *Lower is faster, higher is slower.*



[14] A negligible, yet statistically significant effect was found for "district size" – districts with more enrolled ELLs reported slightly lower CSSGs.

Building on the preliminary evidence presented in Figures 1.1 and 1.2, results reported in Table 2 confirm that we should expect the highest ELL growth, as defined by the annual difference between the district-averaged composite scale scores, in districts with lower starting CPL and grades. Figure 2 shows that as we move up the proficiency level/grade ladder, it becomes increasingly more and more difficult to record high growth (CSSGs).

Next, given the parameter estimates produced by the model characterized in Equation 1, for each district and growth cycle we calculate *adjusted CSSGs,* which will now be free of the potentially biasing effect of district composition with respect to its starting CPL and grade, as well as district size. This will give us a fair, *apples-to-apples* comparison of district performance in terms of ELL growth.[15]

Identifying high-flyers and low-cruisers

Since the district-level heterogeneity in its size, starting CPL, and grade have now been accounted for, we are free to proceed with the comparison of *adjusted CSSGs* across districts. To find high-flyers and low-cruisers we look at the top and bottom quartiles of the distribution of this district-level growth measure, with the ultimate goal of finding some of the same districts in each of the groups across the five growth cycles.[16] Table 5 presents the distribution of identified districts according to top and bottom rankings.

[15] It is important to note that for the calculation of adjusted CSSGs both the fixed-effect and time-varying errors effect of the district (total residual) were used. While using the district fixed-effect or the time-varying errors alone to generate HFDs and LCDs could be an informative exercise, we leave this possibility for future research.

[16] The criteria of top and bottom quartiles for district classification was subjective and was applied to obtain a manageable sample size for the secondary analysis. Altering these criteria to obtain a larger or smaller sample of districts is straightforward and does not affect the generality of the results.

Table 5

Distribution of WIDA Districts by HFD and LCD Status

Times Identified /Total Time	High-Flying (≥ 75 percentile)		Low-Cruising (≤ 25 percentile)	
	Frequency	Percent	Frequency	Percent
0/5	850	57.3	775	52.2
1/5	106	7.1	106	7.1
1/4	47	3.2	56	3.8
1/3	43	2.9	74	5
1/2	61	4.1	73	4.9
2/5	57	3.8	64	4.3
1/1, 2/4	112	7.6	95	6.4
2/3	19	1.3	54	3.6
3/5	55	3.7	46	3.1
2/2	28	1.9	67	4.5
3/4	15	1	18	1.2
3/3	10	0.7	26	1.8
4/5	**38**	**2.6**	**19**	**1.3**
4/4	**13**	**0.9**	**4**	**0.3**
5/5	**30**	**2**	**7**	**0.5**

Table 5 shows that over 50% of WIDA districts (57% for HFDs and 52% for LCDs) are never classified as either high-flying or low-cruising. However, we are more interested in the districts that were repeatedly found to be high-flying or low-cruising. These districts are found at the bottom of the Table 5, formatted in boldface. Thirty (30) districts were identified in the top 75 percentile of the growth distribution for each of the five growth cycles. Similarly, seven school districts that were consistently at the bottom 25 percentile of the CSSG distribution. To maintain a balance between the two groups, districts that were found to be low-cruising in four out of five most recent growth cycles were also classified as LCDs. Given these criteria, a final list of 43 HFDs and 30 LCDs was identified.[17] Additional data from national and state educational agencies, as well as other public data sources was collected and compiled for these districts to perform across- and within-group comparisons. Part 2 of this research report presents the results of the secondary analysis.

Part 2: Finding common and contrasting characteristics for high-flying and low-cruising districts

District-level secondary data compiled from federal and state educational institutions and other public sources.

For the 43 HFDs and 30 LCDs identified through the analysis in Part 1, additional district-level data was collected to compare and contrast the outcomes of the two groups. The district-level variables of interest identified for the secondary analysis can be categorized into demographic factors and academic outcome variables. The demographic variables are district sizes, as measured by the total number of enrolled students, number of full-time equivalent teachers (FTE), pupil/teacher ratios, numbers and proportions of students classified as Individualized Educational Program (IEP), Limited English Proficient (LEP) and Free or Reduced-price Lunch (FRL). Academic outcome variables collected from federal and state online sources are district-level graduation rates and district-level achievement in other academic content areas, as measured by the proportion of advanced and proficient students in state mathematics, reading, and science tests.[18]

The variables of interest were divided into two groups: (a) demographic factors, reflecting the *inputs* of the district and (b) district-level academic outcome variables.[19] Figure 3 presents a diagram of the variables and relationships we are interested in exploring in the context of these two district groups.

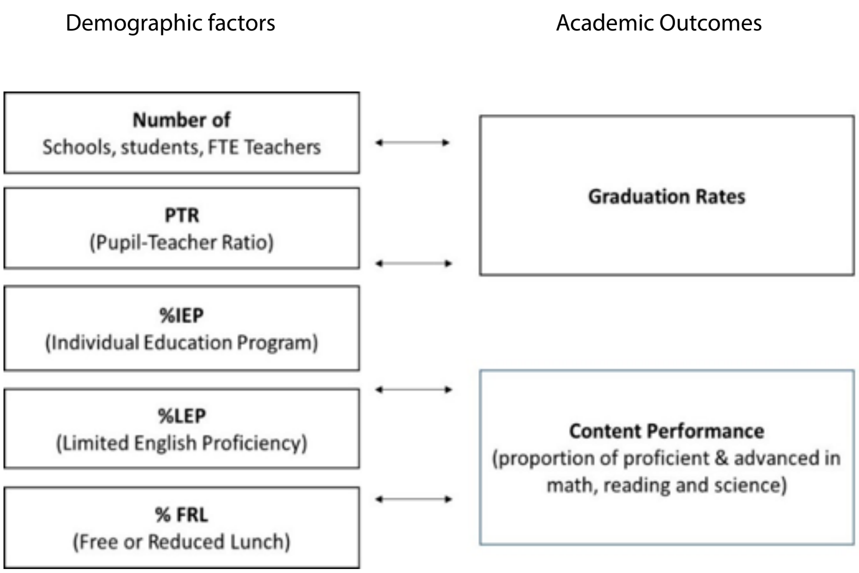
---

[17] It should be noted that by construction this classification of HFDs and LCDs precludes districts in newer WIDA states from being identified as such. In other words, it is impossible for a district to be classified as HFD or LCD if they have not been a part of WIDA for at least four years. However, this apparent deficiency stems from the focus of the research question to find *consistent* over- and under-performers in terms of ELL growth, rather than the research methodology.

[18] These variables were selected through consultations with ELL researchers at WIDA. Availability of federal and state data on district-level demographic and academic outcomes also affected both the scope and scale of the project. The authors thank Sookweon Min for her help in collecting and compiling the secondary data.

[19] Due to either lack or poor quality of the (secondary) district data for the earliest growth cycle, the 2007–2008 data were not included in the analysis.

Figure 3. District-level demographic factors and academic outcomes variables.



Contrasting demographic variables: High-flying versus Low-cruising Districts

Recalling that the objective of the research project is to explore and identify consistent patterns from the district data, we start by visualizing the differences across the two groups for district-level demographic variables. Figure 4 provides a contrast between the HFDs and LCDs for the average number of enrolled students (in thousands), number of schools and number of FTE teachers (in hundreds).

Figure 4. Average number of students, schools and FTE teachers in a LCD versus a HFD.

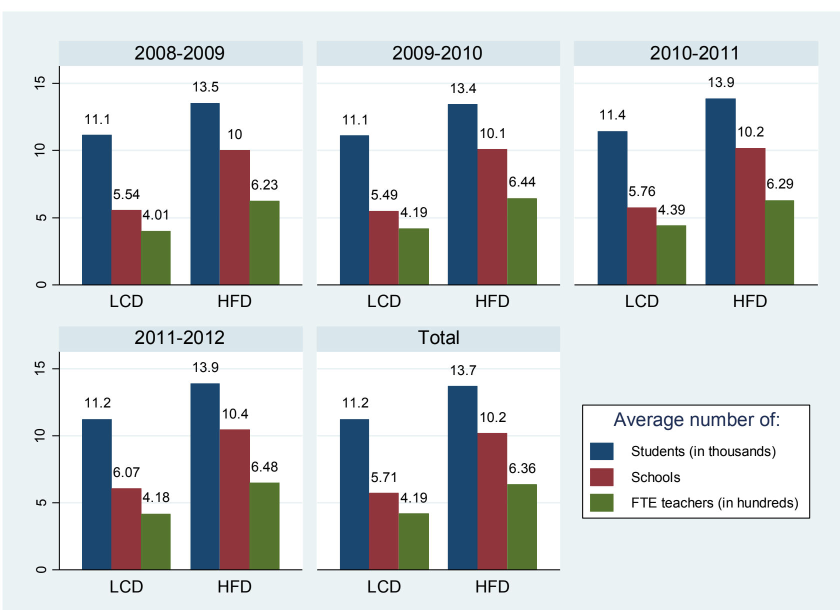


Figure 4 illustrates that typically LCDs have a consistently lower average number of enrolled students, schools, and FTE teachers compared to that of HFDs. This may suggest the presence of *scale effects,* i.e., that it is more difficult to achieve consistent and higher growth in smaller school districts and schools, which also happen to employ lower numbers of FTE teachers. The last sub-graph in Figure 4 combines all the cycles together to provide a general contrast for the average differences across the two groups. The differences revealed in Figure 4 may suggest a skewed teacher resource distribution for HFDs and LCDs. To explore this possibility, Figure 5 provides average pupil/teacher ratios for HFDs and LCDS.

Figure 5. Pupil to teacher ratios, LCD versus HFD.

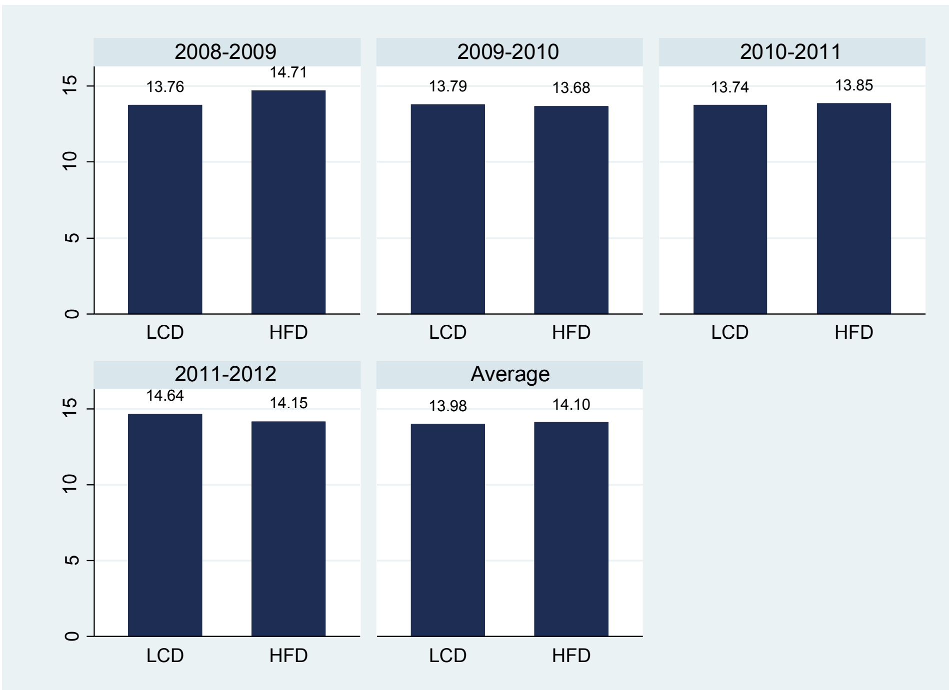
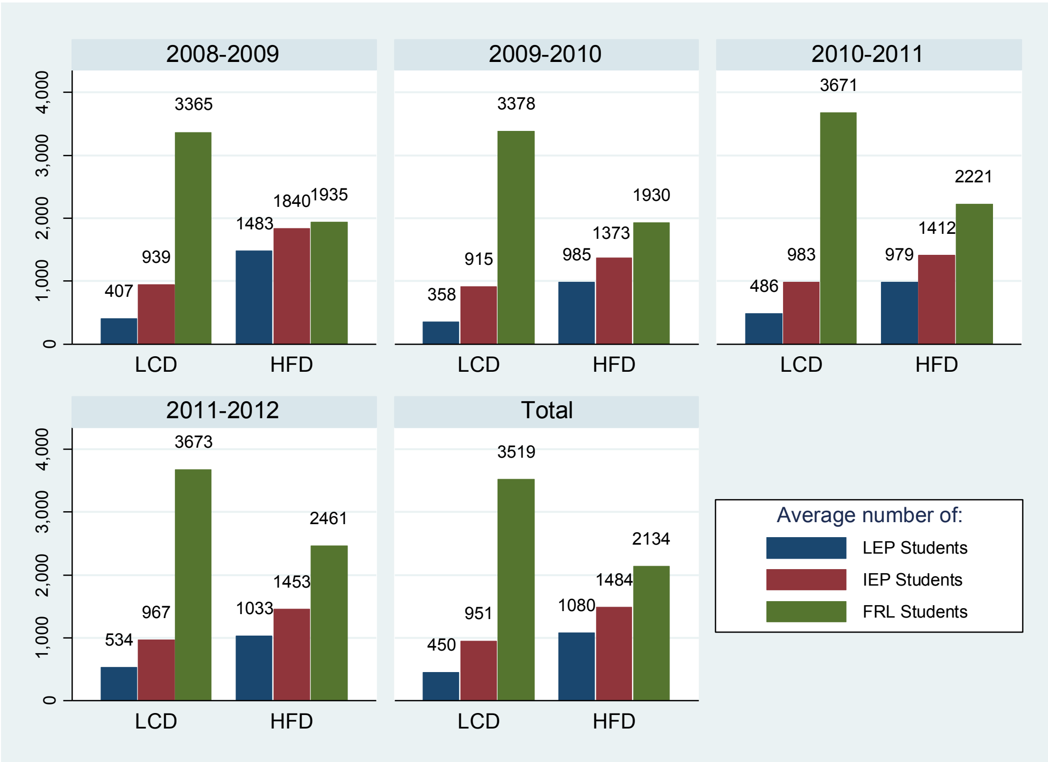


Figure 5 illustrates that contrasting the pupil/teacher ratios between HFDs and LCDs does not produce any tangible trends. Namely, while in the 2008–2009 and 2010–2011 growth cycles HFDs had marginally higher pupil/teacher ratios, the opposite is true in the 2009–2010 and 2010–2011 growth cycles. The average LCD-HFD differences given in the last sub-graph lead to the conclusion that the difference across the groups in pupil/teacher ratios is negligible. In sum, despite the higher number of FTE teachers in HFDs as uncovered and presented in Figure 4, there appear to be no systematic differences in pupil/teacher ratios between the two groups.

Next, we take a look at the student composition of the districts, again by contrasting the characteristics of the HFDs versus LCDs. Similar to Figure 5 above, Figure 6.1 compares the average number of students who are classified as LEP, IEP, and FRL and who are enrolled in high-flying versus low-cruising districts.

Figure 6.1. Average numbers of IEP, LEP and FLR students: LCD versus HFD.



The trends depicted in Figure 6.1 provide some preliminary findings about WIDA's consistently and relatively faster and slower growing districts. First, looking at the average number of LEP and IEP students identified in blue and red in Figure 6.1, there are about 2–3 times as many LEP and IEP students enrolled in HFDs than that of LCDs. Second, tangible differences are observed across the two groups for the average number of FRL students. However, recalling from Figure 4 that HFDs, on average, had a higher total number of students, these differences may be somewhat exaggerated and misleading. Therefore, to further explore the differences in student composition between the high-flying and low-cruising districts, Figure 6.2 provides the same information presented in *proportions* of students who are classified as LEP, IEP, and FRL within a school district.

Figure 6.2. Average proportions of IEP, LEP and FLR students: LCD versus HFD.

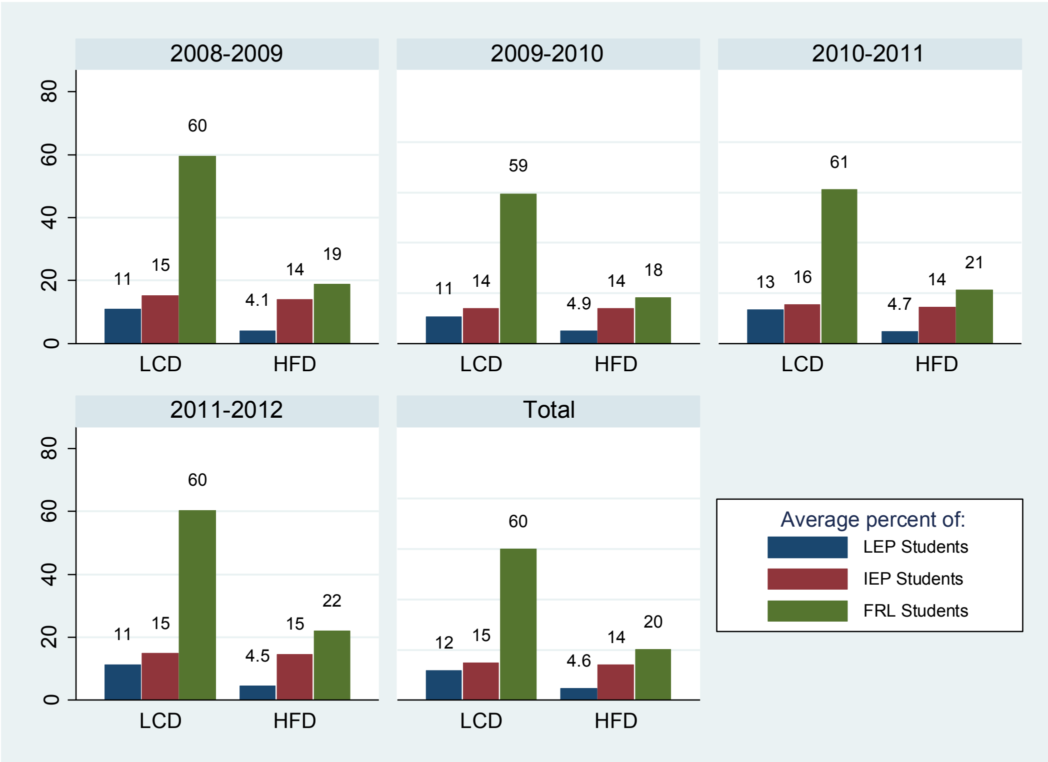


Figure 6.2 provides some interesting insights about the student body composition of the high-flying and low-cruising districts. First, across-group differences in IEP students seem to have flattened when using district-level proportions (instead of student counts). Second, comparing the height of the blue bars in the figure above, we can conclude that HFDs record a lower proportion of LEP students compared to LCDs.[20] Coupled with the finding on the differences in the total number of LEP students presented in Figure 6.1, this implies that the higher and more consistent growth is observed in school districts with a relatively higher number, but lower proportion of students classified as LEP.[21]

Finally, and perhaps most importantly, it is the consistently large and sustained differences in district-level FRL proportions across the two groups that are quite disconcerting. Namely, high-flying districts, on average, enjoy a district-level FRL rate

[20] This result should be treated with caution, recalling that the smallest districts (less than 30 ELLs) were removed from the analysis. However, it still holds true for the small, medium and large districts.

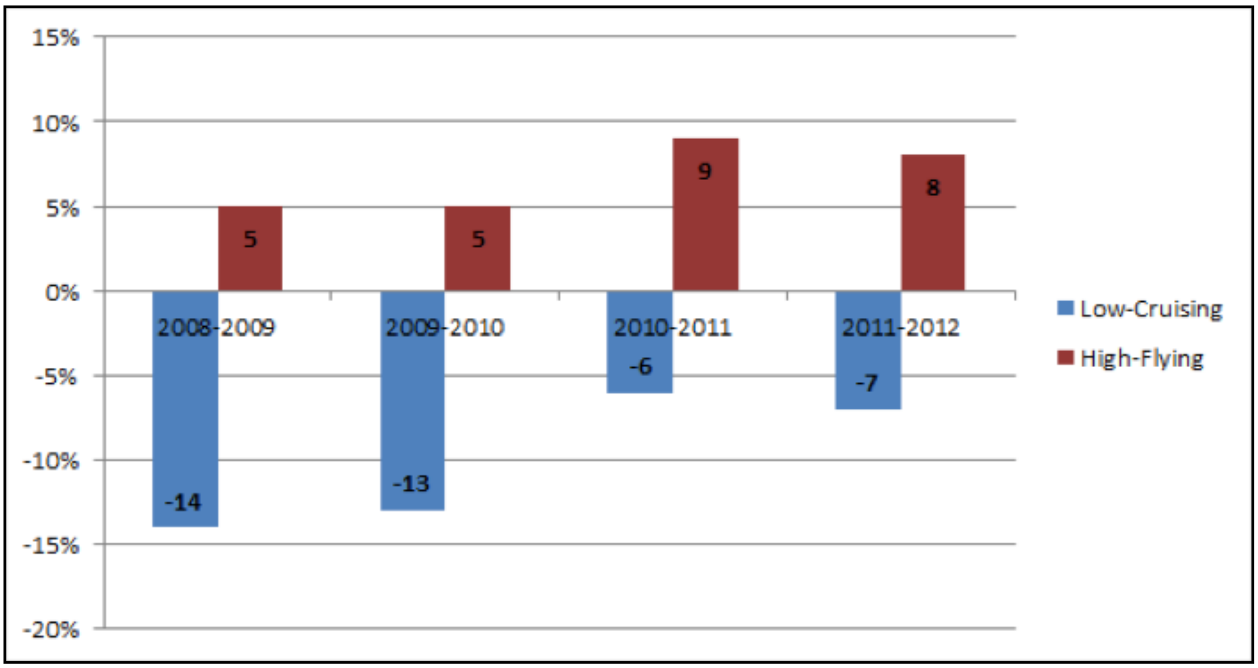
[21] While this is not the only possible explanation, we can speculate that these differences between HFDs and LCDs are partly due by the broader access to resources and higher availability of per-capita LEP support that is likely to exist in school districts that have a high enough number of LEP students qualifying them for Title III federal funding.

that is three times lower than that of low-cruising districts. Despite the fact that there may be a significant overlap across these three student classifications, Figure 6.2 makes it clear that among examined district-level demographic variables, the proportion of students who receive free or reduced-price lunch delineates the most salient difference between HFD's and LCD's growth (i.e., predicted CSSGs). Next, we turn to comparing the district-level academic outcome variables across the *High-Flying* and *Low-Cruising* groups.

Contrasting academic outcome variables for high-flying versus low-cruising districts

We now turn to comparing the outcome variables listed in Figure 3 across the high-flying and low-cruising groups. Figure 7 starts with a comparison of district-level graduation rates across the two groups. Bearing in mind that the districts in question could be located in different states, we calculate and use differences from the state-average graduation rates.[22]

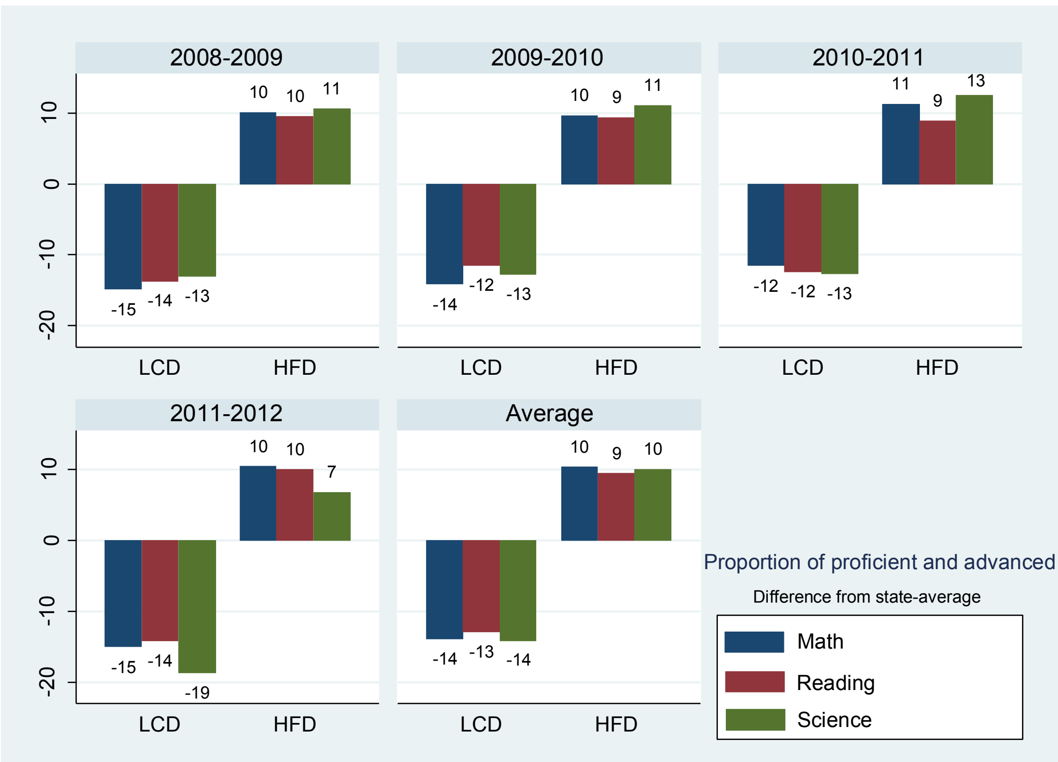
Figure 7. Graduation rates, difference from the state-average: LCDs versus HFDs.



[22] Using district graduation rates instead of difference from state-average rates yields very similar results.

A first look at Figure 7 reveals that high-flying districts have consistently higher graduation rates than their low-cruising counterparts. The differences in graduation rates range from 15-20%, depending on the cycle in question. Figure 7 supports the notion that faster ELL growth may go hand in hand with higher graduation rates. Next, we provide a similar comparison in Figure 8, looking at other academic content performance as the outcome variable. We use the proportion of students receiving an *advanced* or *proficient* score on state mathematics, reading, and science tests as another measure of academic performance. Again, due to the fact that the data from these assessments comes from different states, which are likely to have different criteria for classifying students, we use differences from state-average to visualize the contrasts between high-flyers and low-cruisers.[23]

Figure 8. Proportion of advanced and proficient students in mathematics, reading and science based on the differences in state averages: LCDs versus HFDs.

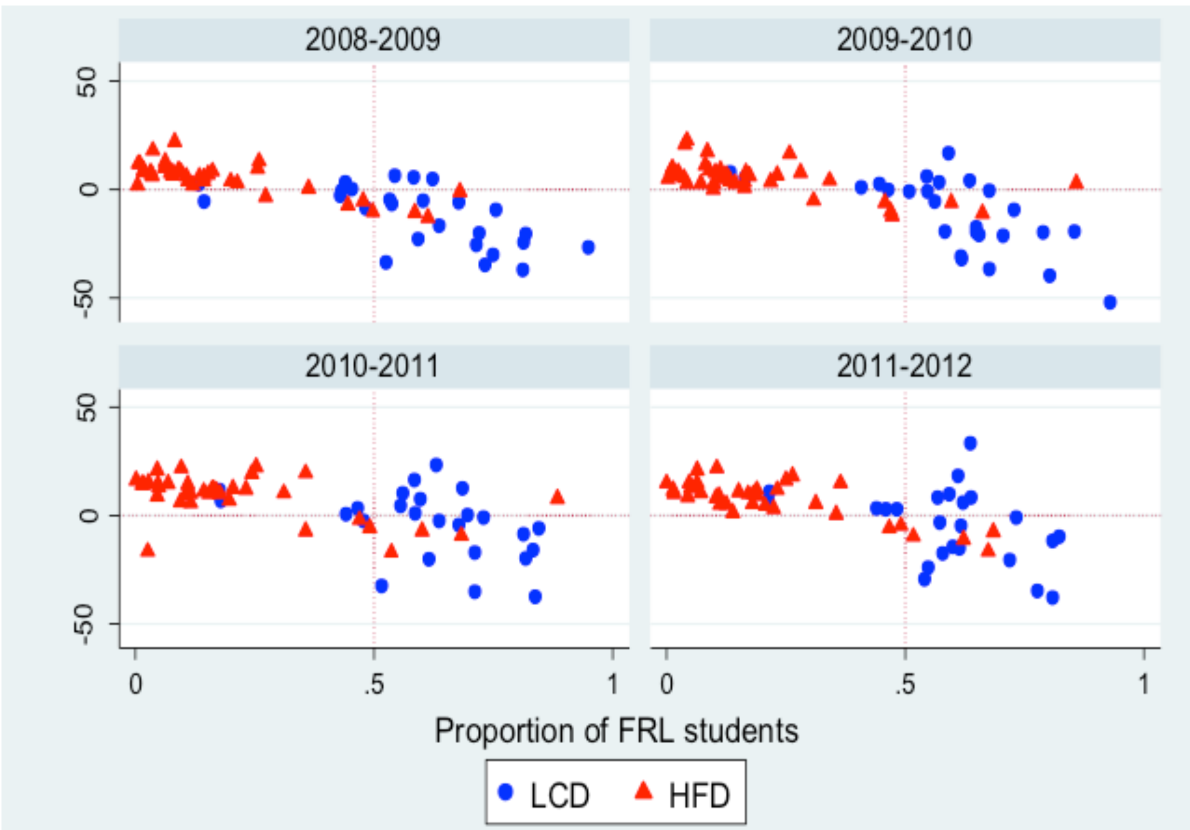


The trends presented in Figure 8 unequivocally complement the previous finding on HFD-LCD differences in average graduation rates. Along with higher graduation rates, HFDs also enjoy a higher proportion of students that score high on state mathematics, reading, and science tests.

[23] Again, similar to the result obtained with respect to graduation rates, using proportions (of advanced and proficient students) instead of differences from state-average yields very similar trends.

In sum, district-level academic outcome variables, as measured by (differences from state-average) graduation rates and proportions of students classified as advanced or proficient in state mathematics, reading, and science tests are consistently higher among HFDs. In other words, consistent and faster ELL growth within a district strongly associates with overall academic performance (measured for all students within a district). We must be careful not to directly relate high ELL growth to improved academic achievement in HFDs. There may be intervening variables that contribute to the observed high academic performance in HFDs. Indeed, Figure 6.2 shows LCDs have proportionately higher number of FRL students compared to HFDs. A large body of research has shown a direct link between poverty (here measured through district-level FRL levels) and academic achievement.[24] Given this relationship between academic outcomes and ELL growth, Figure 9.1 provides a scatterplot of the distribution of LCDs and HFDs along these two dimensions.

Figure 9.1. FRL and graduation rates: LCDs versus HFDs.

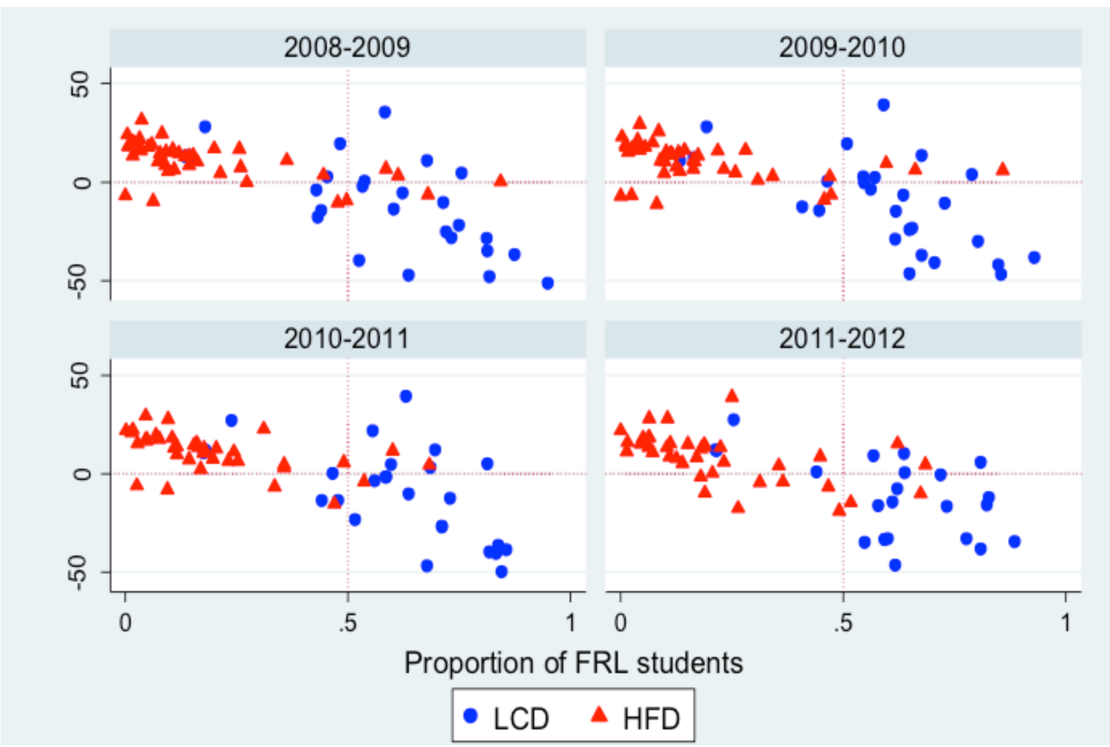


24 Caldas S. J. and Bankston III, C. (1997), Sirin, S. (2005).

In this figure, the y-axis is the district's difference from the state-average graduation rate, and the x-axis is the district's proportion of FRL students. The clustering of red triangles (HFDs) in the upper half of each sub-chart reflects the previous finding that higher graduation rates go with higher ELL growth. The left-right asymmetry in the distribution of red triangles (HFD) and blue dots (LCD) shows just how difficult it is for a district to be high-flying while having a high FRL proportion (>0.5), or conversely, for a low-cruising district to have a low FRL proportion (<0.5).

Given the high correlation between the district-level proportion of students classified as advanced or proficient in mathematics, reading, and science (ρ>0.8), as well as the similarities in patterns of HFD-LCD distribution with respect to these outcomes, a composite academic proficiency measure was constructed using the average of the three test domains. Similar patterns to Figure 9.1 are seen in the figure below in Figure 9.2, which uses the composite measure of state mathematics, reading, and science tests (y-axis) and compares it with the proportion of FRL students. (x-axis)

Figure 9.2. FRL and academic content performance: LCDs versus HFDs.



The distribution of HFD and LCD districts depicted in Figures 9.1 and 9.2 strongly suggests that the proportion of FRL students is the most likely determinant of both district-level ELL growth and performance in other academic areas. Perhaps more importantly, it demonstrates that despite the high FRL levels, a handful of exceptional districts are still able to *fly-high* and record graduation rates and academic performance that are close to or above their state-averages. Conversely, a few districts, despite having a relatively low proportion of FRL students and recording sufficiently high graduation rates, are classified as low-cruising due to consistent, low growth in English language proficiency.

These findings pose a significant challenge for school districts that have inherently high numbers and proportions of FRL students. However, we remain hopeful that a closer examination of these exceptional districts will shed light on some key factors and/or practices that enables them to *fly against the wind* despite the high FRL levels, or conversely, precludes them from doing so despite the low FRL levels. While a preliminary quantitative analysis yielded no tangible differences (to be expected with such a small sample size of eight districts), these findings call for a follow-up qualitative analysis in an effort to identify any unobserved attributes and/or practices that despite all odds drive the over- or under-performance in these districts.

Summarizing the uncovered trends for HFDs and LCDs

Summarizing the findings on the differences between the demographic characteristics and academic outcomes of HFDs and LCDs, we find that on average:

Demographic factors

- HFDs have a higher number of schools, students and FTE teachers.
- Pupil/teacher Ratios are not different across the groups.
- HFDs enroll a higher number but lower proportion of LEP students compared to LCDs.
- HFDs enroll a higher number of IEP students than LCDs; IEP proportions, however, are about equal across the two groups.
- There are large differences across the HFDs and LCDs in FRL numbers and proportions. HFDs enroll a significantly lower number, as well as a lower proportion of FRL students than LCDs.

Academic outcomes

- District-level ELL growth (CSSGs) associates with general academic performance. Namely:
  - Graduation rates are consistently higher among HFDs.
  - The proportion of advanced and proficient students in state mathematics, reading, and science tests is consistently higher for HFDs.
- Despite the high FRL levels, several districts still are able to fly high and record high graduation rates and academic performance that is close to or above their state-average.
- Similarly, a few districts, despite having a relatively low proportion of FRL students and average graduation rates, are classified as low-cruising due to consistent, low growth in English language proficiency.

References

Allison, P. D. (2005). *Fixed effects regression methods for longitudinal data using SAS.* SAS Press.

Caldas S. J., & Bankston III, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research, 90*(5), 269–277.

Cook, H. G., Boals, T., Wilmes, C., & Santos, M. (2008) *Issues in the development of annual measurable achievement objectives (AMAOs) for WIDA Consortium states.* Madison, WI: Wisconsin Center for Education Research.

Cook, H. G., & Zhao, Y. (2011, April). *How English language proficiency assessments manifest growth: An examination of language proficiency growth in a WIDA state.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Greene, W. H. (2011). *Econometric analysis.* Upper Saddle River, NJ: Prentice Hall.

Hakuta, K., Goto Butler, Y., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* University of California Linguistic Minority Research Institute Policy Report 2000-1.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica 46*(6), 1251–1271. ISSN 0012-9682. JSTOR 1913827.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Thousand Oaks, CA: Sage Publications.

Sirin, S. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research, *Review of Educational Research 75*(3), 417-453.

APPENDIX A
SAS and STATA Codes Used to Perform the Model Estimation

SAS

```
Proc Mixed METHOD=ML data=ACCESS_LONG MMEQ NOCLPRINT;
class District_Number Cycle;
model CompositeScaleScoreGain = CompositePL Grade Num_ELL /s outpred =
FE_predict RESIDUAL;
repeated/type=ar(1) subject=District_Number;
run;
```

STATA

```
Xtregar CompositeScaleScoreGain CompositePL Grade Num_ELL, fe rhotype(dw)
```

Wisconsin Center for Education Research
SCHOOL OF EDUCATION
UNIVERSITY OF WISCONSIN–MADISON

The Wisconsin Center for Education Research (WCER) is one of the nation's oldest university-based education research and development centers. WCER is based in the UW–Madison School of Education, which is consistently ranked one of the top schools of education in the country. With annual outside funding exceeding $47 million, WCER is home to centers for research on the improvement of mathematics and science education from kindergarten through postsecondary levels, the strategic management of human capital in public education, and value-added achievement, as well as the Minority Student Achievement Network and a multistate collaborative project to develop assessments for English language learners.

WIDA RESEARCH REPORT

Wisconsin Center for Education Research | University of Wisconsin-Madison
1025 West Johnson St. | Madison WI 53706 | www.wida.us