Getting More From Less

Measuring Efficiency in Connecticut High School Districts

BY DENNIS HEFFLEY AND
CAN BEKAROGLU*

With nearly 58% of all town expenditures in Connecticut going to education, recent fiscal pressures have forced teachers and administrators to make do with less. But this pressure to perform is not new. The push for "assessment" and "accountability," and the mandates and penalties of "No Child Left Behind," have left many educators feeling more like beleaguered test tutors than well-trained professionals. Yet despite the pressure and professional doubts, many Connecticut high school districts seem to be preparing students for the SAT Reasoning Test in a relatively efficient way.

Educators generally recognize the need for common yardsticks, but when the yardsticks define the curriculum, often to the near exclusion of all else, and when students' most frequent question is "Will this be on the test?," it may be time to take a harder look at the yardsticks, or at least the way we read them.

Our state's own yardsticks are the Connecticut Mastery Tests (CMT)—now given in grades 3-8—and the grade 10 Connecticut Academic Performance Test (CAPT). The other important yardstick is the SAT Reasoning Test (SAT), typically taken near the end of 11th grade for college applications. The CMT and CAPT are mandatory and taken by nearly all students. The SAT is optional, but Connecticut's participation rate (83%) ranks 4th among states, topped only by Maine (90%), New York (85%), and Massachusetts (84%); the U.S. average is only 46%.

So what's wrong with standardized testing? It's been around for a long time, and most of us survived the process with few scars. But the process has changed. No longer just a way to gauge individual performance or academic potential, standardized testing also has become the primary "assessment tool" for evaluating curriculum, teachers, administrators, schools, and districts. Unfortunately, test performance also depends on external factors beyond educators' control. Previous articles in these pages (Spring issues of 1997, 1998, and 2002, and Summer 2007) found that much of the variation in test scores, at both high school and lower-grade levels, reflects socioeconomic disparities rather than differences in school resources. So does a district's lower score on an exam reflect its teachers' failure to teach, inadequate resources, or adverse local or personal conditions that make learning difficult?

Whatever the cause, failure to acquire needed information and skills hampers success in school and later in the workplace. But when it comes to identifying and correcting deficiencies, understanding the source of the problem matters plenty. Focusing on the school alone can lead to policies that ignore the underlying barriers to learning. By contrast, recognizing that family, community, and educational inputs all contribute to academic success underscores the need for a multi-pronged approach to the problem.

A DIFFERENT APPROACH

Our analysis departs from the earlier studies in two ways. First, rather than using a single measure of performance or "output," we simultaneously consider performance in three critical areas: math, reading, and writing. The technique we use—data envelopment analysis (DEA)—allows for multiple outputs and inputs. Second, rather than treating the results of one exam, at a given point in time, as the output, we define the outputs in "value-added" terms: differences in performance, relative to other districts, on two important standardized tests, administered roughly 3 years apart.

We evaluate each district's improvement in average math, reading, and writing scores between the 8th grade Connecticut Mastery Test (CMT8) and the average math, reading, and writing scores on the SAT Reasoning Test. By focusing on essentially the same cohort—8th graders in 2006 who then took the SAT as high school juniors in 2009—we hope to control for "student quality" and see whether their training in the intervening years (grades 9-11) improved or diminished their relative performance. Socioeconomic factors might still play a role during this period, so we'll also take those differences into account. But using value-added measures should help to isolate each district's contribution to its students' SAT scores, a primary determinant of college acceptance in many parts of the country.

The procedure entails several steps. First, for each exam part or subject (math, reading and writing) we construct an output measure that captures the improvement in relative performance between the two exams. The numerator of the value-added measure of output is the district's average SAT score on that subject, relative to the average across all 119 districts. The denominator is the district's average CMT8 score on the same subject, again relative to the average across districts. A larger ratio signals greater improvement in students' performance between the two tests.

We adopt this relative measure for two reasons. First, the CMT8 and SAT are scaled differently, so a meaningful comparison requires a "normalization" of each exam score. We do that, for both exams, by using each district's score relative to the average score for all districts in the study. Second, output measures used in DEA must be positive. A value-added measure of output based on a simple difference in performance between the two exams could be negative and would present technical problems that defining output as a ratio avoids. Across the 119 districts, the value-added measure ranges from 0.832 to 1.075 for math; 0.861 to 1.092 for reading; and 0.852 to 1.067 for writing.

The second step in the analysis requires identifying certain "inputs" that might affect the three value-added district outputs. The four inputs used here are teachers, administrators, and computers—each measured per 100 students—and annual hours of instruction. The table shows the variation in the four inputs and the three value-added output measures across the 119 districts.

The final and most complex step involves the use of data envelopment analysis (DEA) to measure each district's efficiency in producing the three outputs from the four inputs. DEA essentially compares each district's observed mix of inputs and outputs with those of other districts and determines whether the same (or an even larger) output bundle might be produced with the same (or an even smaller) input bundle by adopting some combination of the input-output mixes seen in other districts. The particular DEA model used here generates a separate efficiency score for each input and each output. If the district could expand one or more outputs without using extra inputs, its "output efficiency" score will be less than 1. Similarly, if one or more inputs might be reduced without sacrificing outputs, the district's "input efficiency" score will be less than 1.

EFFICIENCY SCORES

For each district, the columns in the table show the District Reference Group or DRG (assigned by the CT Department of Education, based on differences in socioeconomic characteristics); total district enrollment in grades 9-12; the four inputs and three value-added outputs; an output efficiency score (averaged across the three outputs); and an input efficiency score (averaged across the four inputs). Figures in the bottom three rows of each column show the minimum, average, and maximum values.

The good news is that the average output efficiency score is quite high (0.970), and 21 of the 119 districts are fully efficient with respect to both outputs and inputs: Avon, Brookfield, Coventry, Glastonbury, Meriden, Monroe, New Britain, New Milford, Newtown, Simsbury, Southington, West Hartford, Windham, and Regional Districts 4, 5, 8, 9, 12, 16, 18, and 19. Two other districts, Greenwich and Weston, were output-efficient, but dropped down the list in terms of input efficiency.

This group provides the benchmark for other districts, but with one caveat: because DEA only compares each district with the most efficient districts in the sample, it's conceivable that yet higher performance levels are possible, but the available data cannot reveal such possibilities. Nevertheless, we find that a number of well-regarded districts not only post high levels of performance on the SAT, but also rank high in value-added (improvement in relative performance from CMT8 to SAT); and they achieve these results in an efficient way.

FURTHER IMPLICATIONS

The results also suggest several other points. First, not all of the "fully efficient" districts are the customary high-achievers. Meriden, New Britain and Windham are good examples. In terms of raw SAT performance (measured by their combined math, reading, and writing scores), none of the three districts ranks among the top 100 in the state. But their relatively high value-added measures of output and conservative use of inputs result in high measures of efficiency. This outcome reinforces the notion that not all "premier districts" are particularly efficient, and not all "troubled districts" are inefficient, after controlling for factors beyond a district's control—accomplished here by relating *value-added* measures of performance to standard inputs.

Second, 8 of the 17 regional districts post very strong results. Since regional districts typically draw students from several towns, one might expect that community differences in expectations or preferred levels of funding would make it more difficult for schools to do their job. Within this group, however, that does not seem to hold. Regional districts, when well managed, might strengthen the bonds between neighboring communities, encourage greater local support for education, and avoid the more wasteful forms of competition between neighbors.

Third, average output efficiency (0.970) significantly exceeds average input efficiency (0.823). The separate components of the average input efficiency, not shown in the table, are 0.865 for teachers, 0.804 for administrators, 0.644 for computers, and 0.979 for hours of instruction. The particularly low figure for computers is interesting, but may simply indicate that schools, like some businesses, carry equipment on the books beyond its useful life. Older, perhaps unused, computers are still being counted as inputs, even if they contribute little to the measured outputs, thereby dragging down the input efficiency scores.

SOCIOECONOMICS AND SIZE

Earlier we noted that socioeconomic factors may still affect a district's

EFFICIENCY IN TRANSFORMING EDUCATIONAL INPUTS INTO OUTPUTS

			INPUTS				OUTPUTS			EFFICIENCY SCORES	
District	Ref Group	Total HS students	Teachers per 100 students	Admin per 100 students	Computers per 100 students	Annual hours of instruction	Value-added (math)	Value-added (reading)	Value-added (writing)	Output efficiency	Input efficiency
Ansonia	H	761	7.36	1.71	30.57	962	0.930	0.997	0.947	0.966	0.802
Avon	B	1036	7.56	1.54	22.88	990	1.046	1.003	1.026	1.000	1.000
Berlin	D	1061	8.88	1.32	22.10	1023	0.955	0.961	0.955	0.957	0.841
Bethel	D	1034	9.26	1.45	29.30	1009	0.958	0.956	0.936	0.949	0.781
Bloomfield	G	704	10.91	2.56	39.44	990	0.832	0.861	0.853	0.844	0.638
Bolton	C	275	12.65	1.82	36.33	1060	0.990	1.049	1.044	0.993	0.748
Branford	D	1146	9.21	1.13	39.07	1059	0.920	0.946	0.931	0.943	0.778
Bridgeport	I	5453	6.96	1.76	27.89	989	0.896	0.889	0.852	0.873	0.841
Bristol	G	2806	7.20	1.50	21.78	1007	0.972	0.975	1.009	0.990	0.889
Brookfield	B	1020	7.34	1.18	34.11	930	0.934	0.943	0.964	1.000	1.000
Canton	C	515	9.81	1.75	28.11	1022	0.956	1.003	0.981	0.979	0.732
Cheshire	B	1635	7.57	1.16	26.12	1001	0.983	0.951	0.969	0.970	0.898
Clinton	D	638	9.97	1.57	38.10	994	0.978	1.014	0.970	0.984	0.731
Colchester	D	982	8.56	1.32	32.46	1006	0.929	0.942	0.937	0.947	0.793
Coventry	E	559	9.23	1.43	30.51	1023	1.075	1.027	1.028	1.000	1.000
Cromwell	D	560	10.18	1.61	43.90	986	0.943	0.940	0.868	0.900	0.722
Danbury	H	2932	7.55	1.64	18.59	963	0.948	0.957	0.953	0.961	0.890
Darien	A	1180	9.86	2.20	41.91	1011	1.040	1.055	1.028	0.999	0.758
Derby	H	430	11.16	2.09	30.66	1003	0.861	0.946	0.897	0.901	0.673
East Granby	D	258	12.36	2.71	43.91	1038	0.992	0.939	0.990	0.970	0.588
East Haddam	E	406	8.67	2.22	53.57	1008	0.952	1.003	0.972	0.958	0.674
East Hampton	D	580	8.10	1.72	27.27	1021	0.950	0.961	0.976	0.961	0.780
East Hartford	H	2411	7.28	1.62	29.45	1057	0.963	0.981	0.952	0.974	0.796
East Haven	G	1130	8.27	1.59	36.91	953	0.931	0.969	0.986	0.973	0.762
East Lyme	D	1304	8.51	0.92	28.07	1003	0.989	0.972	1.030	0.985	0.957
East Windsor	E	455	9.74	1.32	34.99	1008	0.970	0.956	1.014	0.981	0.774
Ellington	C	770	8.81	1.30	41.61	1007	0.949	0.974	0.955	0.970	0.761
Enfield	E	2083	8.31	1.34	31.10	993	1.005	0.989	0.966	0.981	0.829
Fairfield	B	2611	9.50	1.65	41.04	933	0.999	0.975	0.991	0.994	0.825
Farmington	B	1389	8.94	1.15	19.59	994	1.003	0.965	0.969	0.973	0.938
Glastonbury	B	2065	7.79	1.31	11.57	996	0.982	1.004	1.017	1.000	1.000
Granby	B	749	8.60	1.47	44.51	1000	0.917	0.945	0.983	0.932	0.765
Greenwich	B	2693	9.34	1.97	26.43	1027	1.042	1.045	1.047	1.000	0.906
Griswold	E	792	8.66	1.01	42.84	1021	0.962	0.928	0.948	0.951	0.824
Groton	G	1414	7.85	2.12	29.69	1032	0.993	0.984	1.021	0.993	0.757
Guilford	B	1154	8.30	1.65	25.63	986	0.925	0.951	0.934	0.939	0.799
Hamden	G	2194	7.73	1.37	31.37	963	0.956	0.936	0.934	0.954	0.825
Hartford	I	5884	8.57	2.35	41.06	952	0.937	0.904	0.888	0.921	0.687
Killingly	G	834	10.29	1.92	42.81	1021	0.910	0.935	0.909	0.901	0.686
Lebanon	E	587	9.81	1.36	34.19	1011	0.928	0.982	1.015	0.975	0.768
Ledyard	D	1057	8.37	1.23	27.90	981	0.933	0.952	0.977	0.965	0.843
Litchfield	E	419	14.77	1.67	19.56	1080	0.967	1.021	1.018	0.988	0.757
Madison	B	1238	8.63	1.45	27.96	1015	0.997	0.996	0.977	0.984	0.810
Manchester	G	2153	8.32	1.58	35.54	1027	1.007	1.014	1.024	0.996	0.794
Meriden	H	2535	8.01	1.51	26.03	945	1.041	0.993	1.014	1.000	1.000
Middletown	G	1336	7.83	1.72	25.74	978	0.975	1.006	0.991	0.993	0.810
Milford	D	2133	8.81	1.45	37.06	950	0.949	0.941	0.959	0.961	0.777
Monroe	B	1402	7.47	1.00	37.71	975	1.013	1.007	1.011	1.000	1.000
Montville	E	858	9.11	1.52	41.84	973	0.945	0.950	0.957	0.934	0.762
Naugatuck	G	1410	8.75	1.63	23.80	967	0.927	0.929	0.920	0.932	0.807
New Britain	I	3129	6.85	1.15	22.57	972	1.037	1.024	0.995	1.000	1.000
New Canaan	A	1223	8.82	1.96	24.83	988	1.024	0.961	0.967	0.970	0.799
New Fairfield	B	938	8.90	1.49	41.92	974	0.923	0.929	0.929	0.910	0.769
New Haven	I	5476	9.71	2.36	36.12	995	0.886	0.921	0.880	0.895	0.669
Newington	D	1528	6.98	1.24	30.89	1013	0.998	0.977	0.966	0.975	0.881
New London	I	805	8.94	2.36	38.44	1012	0.902	0.881	0.854	0.879	0.677
New Milford	D	1591	7.36	1.19	21.57	991	0.998	1.001	1.058	1.000	1.000
Newtown	B	1719	7.15	1.11	18.07	952	0.962	1.004	1.000	1.000	1.000
North Branford	E	729	8.07	1.65	38.98	1051	0.930	0.949	0.963	0.931	0.763
North Haven	D	1240	7.93	1.13	25.56	975	0.982	0.981	0.979	0.984	0.904
North Stonington	E	245	15.10	2.86	49.24	1020	1.056	0.967	0.977	0.970	0.625
Norwalk	H	3314	8.07	1.66	31.60	997	0.974	0.975	0.967	0.972	0.779
Old Saybrook	D	483	9.54	1.24	38.81	1037	0.962	0.914	0.901	0.937	0.756
Plainfield	G	916	7.53	1.42	49.94	1007	0.982	0.986	0.960	0.958	0.784
Plainville	E	898	8.57	1.56	38.24	971	1.023	0.985	0.979	0.983	0.776
Plymouth	E	553	7.85	1.99	22.12	978	0.941	0.942	0.987	0.961	0.804
Portland	E	362	9.78	1.93	50.16	987	0.957	0.983	1.009	0.965	0.687
Putnam	G	394	9.14	2.03	41.97	1005	0.948	0.943	1.002	0.952	0.701
Ridgefield	A	1750	8.34	1.26	41.87	966	1.001	0.999	0.999	0.995	0.806
Rocky Hill	D	755	9.17	1.32	34.26	1036	0.959	0.954	0.932	0.960	0.765
Seymour	E	861	8.07	1.63	29.85	939	0.933	0.965	0.965	0.964	0.877
Shelton	D	1687	8.19	1.36	13.14	987	0.971	0.953	0.993	0.974	0.961
Simsbury	B	1535	8.44	1.50	30.18	921	1.014	0.979	0.985	1.000	1.000
Somers	C	567	9.51	1.23	26.68	1062	0.994	0.958	0.969	0.970	0.822
Southington	D	2164	8.51	1.48	19.33	948	0.967	0.959	1.021	1.000	1.000
South Windsor	B	1651	7.67	1.39	23.52	1035	0.962	0.964	0.975	0.968	0.852
Stafford	G	521	8.39	2.11	25.63	1013	0.967	0.988	1.002	0.985	0.746
Stamford	H	4599	8.22	1.78	24.31	999	0.990	0.974	0.995	0.983	0.797
Stonington	D	811	8.34	1.60	26.52	998	1.015	1.002	1.029	0.997	0.850
Stratford	G	2366	7.20	1.82	29.63	944	0.933	0.950	0.945	0.970	0.888
Suffield	C	866	8.65	1.15	40.09	1033	0.979	0.957	0.957	0.969	0.797
Thomaston	E	378	13.44	1.85	61.93	970	1.040	1.010	0.939	0.975	0.635
Thompson	E	400	10.23	2.50	39.24	986	0.936	0.972	1.019	0.974	0.652
Tolland	C	849	7.89	1.41	35.47	1046	0.998	0.985	1.036	0.993	0.819
Torrington	G	1296	8.72	1.77	28.11	1044	0.992	0.973	1.035	0.990	0.778
Trumbull	B	2094	7.95	1.15	23.10	988	0.950	0.958	0.879	0.938	0.903

District	Ref Group	Total HS students	INPUTS: Teachers per 100 students	INPUTS: Admin per 100 students	INPUTS: Computers per 100 students	INPUTS: Annual hours of instruction	OUTPUTS: Value-added (math)	OUTPUTS: Value-added (reading)	OUTPUTS: Value-added (writing)	EFFICIENCY SCORES: Output efficiency	EFFICIENCY SCORES: Input efficiency
Vernon	G	1215	8.81	1.65	29.04	994	0.991	1.032	0.977	0.976	0.830
Wallingford	D	2168	8.58	1.29	29.15	1022	0.941	0.985	0.994	0.984	0.810
Waterbury	I	4495	7.99	1.98	40.71	1025	0.934	0.933	0.901	0.906	0.735
Waterford	D	1013	8.18	1.38	37.24	1007	0.998	0.993	0.988	0.990	0.793
Watertown	D	990	8.46	1.41	27.14	976	0.939	0.920	0.909	0.928	0.820
Westbrook	E	310	10.84	1.29	33.00	1007	0.934	1.015	1.009	0.973	0.788
West Hartford	B	3030	8.11	1.62	31.12	951	1.022	1.032	1.067	1.000	1.000
West Haven	H	1751	8.08	1.60	25.21	951	0.972	0.942	0.961	0.964	0.846
Weston	A	801	8.89	1.87	49.73	998	1.038	1.012	1.048	1.000	0.781
Westport	A	1721	9.27	2.32	50.53	972	1.036	1.042	1.013	0.997	0.723
Wethersfield	D	1212	8.69	1.49	21.71	971	1.004	0.974	1.005	0.987	0.877
Wilton	A	1256	8.76	1.59	38.24	1007	1.022	1.054	1.037	0.997	0.861
Windham	I	1033	8.93	2.03	22.12	984	1.037	1.092	1.067	1.000	1.000
Windsor	D	1473	8.77	2.04	42.09	941	0.912	0.927	0.945	0.937	0.753
Windsor Locks	E	584	9.50	1.88	32.34	1039	0.935	0.999	0.961	0.964	0.707
Wolcott	E	915	8.24	1.31	20.54	1002	0.954	0.904	0.937	0.931	0.875
Regional Dist 1	E	560	10.48	1.04	51.99	984	0.956	0.972	0.993	0.963	0.776
Regional Dist 4	C	615	9.32	0.49	38.77	984	1.008	0.998	1.027	1.000	1.000
Regional Dist 5	B	1649	8.72	0.79	22.35	993	0.989	0.989	0.975	1.000	1.000
Regional Dist 6	E	408	13.38	1.96	35.10	1021	1.039	0.996	0.987	0.988	0.680
Regional Dist 7	C	794	8.93	0.88	55.25	1044	0.957	0.974	0.964	0.954	0.817
Regional Dist 8	C	1019	8.58	0.59	37.55	997	0.967	0.987	1.025	1.000	1.000
Regional Dist 9	A	962	9.26	0.83	50.56	969	0.991	1.045	0.997	1.000	1.000
Regional Dist 10	C	778	8.80	1.80	34.54	969	0.992	1.024	1.002	0.991	0.773
Regional Dist 11	E	173	26.59	1.16	69.16	1037	0.996	1.017	1.013	0.981	0.610
Regional Dist 12	C	393	10.23	1.53	39.71	1025	1.033	1.076	1.050	1.000	1.000
Regional Dist 13	C	587	9.34	1.87	39.12	996	0.959	0.973	0.993	0.986	0.693
Regional Dist 14	C	866	7.67	1.27	38.01	991	0.956	1.004	0.976	0.990	0.810
Regional Dist 15	B	1415	7.67	1.27	24.53	1014	0.972	0.974	0.987	0.979	0.873
Regional Dist 16	E	816	8.55	1.23	37.63	937	1.058	1.016	1.056	1.000	1.000
Regional Dist 17	C	716	8.59	1.54	24.16	994	0.974	0.977	0.968	0.973	0.810
Regional Dist 18	C	471	10.40	2.34	35.70	955	1.000	1.079	1.023	1.000	1.000
Regional Dist 19	C	1184	9.83	0.54	30.79	1049	1.038	1.005	1.026	1.000	1.000
	MIN	173	6.85	0.49	11.57	921	0.832	0.861	0.852	0.844	0.588
	MAX	5884	26.59	2.86	69.16	1080	1.075	1.092	1.067	1.000	1.000
	AVG	1337	9.05	1.56	33.47	997	0.973	0.977	0.978	0.970	0.823

SOURCE: *The Connecticut Economy*, based on Connecticut Department of Education data.

efficiency, even if we define output in value-added terms. To test that proposition, we regressed the DEA efficiency scores on the DRG classifications that appear in the second column of the table. Dummy variables denoted each group, B through I, with group A (the most advantaged in socioeconomic terms) serving as the omitted dummy. The regressions also controlled for district size (students) and allowed for a nonlinear size effect by including a quadratic term (students squared). This enabled us to see if there is some "ideal" district size when it comes to efficient use of resources.

The regressions showed that poor socioeconomic status has larger negative effects on output efficiency than on input efficiency. In both regressions, about a third of the efficiency differences was explained by the collective effects of the DRG dummy variables and the district size variables. In both regressions, district size was a significant factor. Controlling for the DRGs, the district size that maximizes output efficiency was 2,789 students. For maximizing input efficiency, the ideal district size was 2,782 students—remarkably similar, given the independent regressions. Both figures exceed the reported district average of 1,337 high school students by an order of magnitude.

A MATTER OF PERSPECTIVE

Properly used, standardized testing can help to establish and achieve high standards for students who are facing competition from an increasingly competent international workforce. Misused, it can demoralize students and teachers alike and drain the fun out of exploring new (or old) ideas and learning from the experience. More importantly, undue emphasis on testing can displace the most important skill—the ability to sift through raw material, untested ideas, and new concepts, and in the process figure out what's important and what's not.

If we are to rely on standardized tests, we need to use the results in ways that recognize the role of both educational inputs and community inputs. Our analysis suggests that high test scores, per se, are neither a reliable indicator of output nor a good measure of school efficiency. Performance measures of output either need to be defined in value-added terms, or adjusted for the socioeconomic conditions that hinder learning.

*Can Bekaroglu recently completed an M.A. degree in economics at the University of Connecticut.