Review of Kansas Cost Studies

Bruce D. Baker, Rutgers University

Prepared on behalf of Schools for Fair Funding, Inc.

March 23, 2018

Executive Summary

In the attached report, I summarize and critique a) past efforts to measure the costs of meeting Kansas' constitutional obligation that the legislature "make suitable provision for finance of the educational interests of the state," b) Dr. Jesse Levin's reports which summarize methods for conducting cost analysis in education, and his review of the recent cost analysis by WestEd and Dr. Lori Taylor, and c) I provide additional critique and analysis of the report by WestEd and Dr. Lori Taylor.

As a general overview:

- Previous studies by both Augenblick and Myers, and William Duncombe and John Yinger in collaboration with the Legislative Division of Post Audit provided reasonable guidance, leading to reasonable reforms to the state school finance formula, which were never fully realized;
- Spending in high poverty districts has slipped below 2006 levels, adjusted for competitive wage growth. If these spending levels weren't sufficient in 2006 to meet 2006 standards, they cannot possibly be sufficient now;
- Teacher wages have slipped substantially relative to the wages of similarly educated, same age non-educators in Kansas, making it difficult if not entirely infeasible to recruit and retain a teacher workforce of similar quality to that which existed in 2006.
    - The quality of the teacher workforce is of utmost importance in determining the quality of schooling provided to Kansas children.
    - Reducing the gap between teacher and non-teacher wages to even those levels which existed in 2006 would require a significant increase in funding for Kansas districts. Any estimate suggesting such increases are unnecessary simply aren't credible.
- The new, WestEd/Taylor study provides reasonable guidance for moving forward on state school finance policy reform, with a few caveats noted in the body of this report.

Dr. Jesse Levin's Preliminary Review

Dr. Levin's report has been characterized in local and regional media as levying harsh criticism on prior efforts to determine the cost of Kansas' constitutional obligations regarding school funding (Hawver's Capitol Report, March 10). Indeed, Dr. Levin did raise concerns regarding the 2002 Augenblick and Myers study and its translation into

policy recommendations (most notably, the combination of a successful schools derived base figure with weights from the professional judgment analysis).

Dr. Levin had much less to say, and few criticisms to offer regarding the cost model estimated in 2006 by William Duncombe and John Yinger (DY), but did critique how that model was translated into policy recommendations by the Legislative Division of Post Audit.

Dr. Levin's initial report provides useful guidance for checking the sensibility, reliability and validity of findings generated by cost studies. In light of his recommendations, I show in this report that:

- While Dr. Levin raises concerns regarding the A&M study methods, previously published academic articles comparing the findings of the A&M study to a) the LPA DY cost model and b) other cost studies suggest that the A&M findings were reliably correlated with other studies and validly associated with student outcomes.
- In fact, the one prior cost study which deviates most significantly from the body of studies available in the mid-2000s, in terms of relating adequacy gaps to existing outcomes (validity check) and in terms of sensitivity to poverty (reliability check), is Dr. Taylor's cost model of Texas school districts.

WestEd & Dr. Lori Taylor Cost Model Methods

We (researchers including Dr. Levin, Dr. Taylor and myself) have all learned a great deal about how to refine data, methods and models for estimating education costs since we first engaged in such endeavors. Dr. Taylor produced cost estimates for Texas school districts in the early to mid-2000s using a highly non-linear model, setting aside concerns over endogeneity (not using a two-stage approach), and controls for inefficiency (not including indirect predictors of inefficiency). That model produced smaller need adjustments than other cost models estimated around that time (see Baker, Taylor and Vedlitz, 2008).[1]

By contrast, the 2006 Duncombe and Yinger (DY) model estimated for Kansas did use a two-stage model and did include indirect controls for inefficiency, as per the usual method of these authors. Over time, I have become convinced that the Duncombe and Yinger approach more adequately isolates the relationship between inputs and outcomes, and costs associated with improving outcomes for low income students (e.g. poverty weights).

To summarize WestEd and Dr. Taylor's new Kansas cost model:

- The current model applies methods more similar to that of William Duncombe and John Yinger, including:

[1] Baker, B. D., Taylor, L. L., & Vedlitz, A. (2008). Adequacy estimates and the implications of common standards for the cost of instruction. *National Research Council*.

- Consideration that the outcome measures of interest are endogenous and use of instrumental variables (2SLS) estimation, blending this approach with Dr. Taylor's preferred method for cost modeling – stochastic frontier modeling.
- Inclusion of indirect controls for inefficiency to account for omitted variables bias in the spending measure (spending not associated with outcome variation, but predictable as a function of fiscal capacity, competition and public monitoring characteristics of districts).

- Dr. Taylor's model has the advantage over the DY model of using multiple outputs, not aggregated into a single index.
- Taylor's approach continues to differ in one particularly relevant regard from the DY approach, and that is in the use of a polynomial (2^{nd} order) U-shaped curve to represent variations in costs associated with economies of scale (where DY uses a series of district size categories).
- Taylor also did not test for or acknowledge potential variation in poverty related costs in relation to poverty concentration, urbanicity or population density.
- Taylor does not seem to have used a cross-validation (predictive validity) method for her selected model.

Manifestation of modeling differences in cost estimates

Dr. Taylor's model yields largely rational results and cost estimates but for some problematic distortions resulting from the U-shaped economies of scale weight.

- The use of a second order polynomial term to generate an economies of scale weight generates the inappropriate assumption that large districts (>20,000) have higher uncontrollable costs than midsize districts (2,000-5,000). Hypothetically, a district with 20,000 students could be reorganized into 4 to 10 districts with 2,000 to 5,000 students to operate at lower cost (greater efficiency).
  - Thus, the proposed scale weight has the effect of a) depressing cost estimates for mid-size districts and b) inflating cost estimates, especially for otherwise very low need very large districts.
  - Overestimating the costs per pupil for low need very large districts – like Blue Valley, Shawnee Mission and Olathe puts these districts current spending below supposed needed spending to achieve desired outcomes, despite their already very high outcomes. This adds as much or more than $50 million in the total cost of meeting Taylor's adequacy targets (for Scenario A), for these three districts alone.
  - Overestimating costs of low need large districts (simply because they are large) and underestimating the costs for high need midsize districts (simply because they are midsize) also compromises weak validity checks on the model. Because there are low need large districts that are high performing, but estimated to face adequacy gaps, and higher need midsized districts that are low performing, but estimated to currently exceed their adequate funding levels, the correlation between funding gap and outcomes is reduced. These correlations are lower for the WestEd Taylor study than for the prior Duncombe and Yinger Study.

- This problem cannot be fixed by simply bottoming out the economies of scale weight at the current minimum or raising it to the large district plateau. Changing the structure of the scale term would affect other factors in the model. The appropriate solution would be to re-estimate the model with district size categories, as done by Duncombe and Yinger, wherein large districts serve as the baseline group.
    - Applying this change, Dr. Taylor might find that there is indeed a relationship between poverty and population density (as in the DY model) which may not appear in the current model due to the large district weight created by the U-shaped size curve.

Adoption & Moving Forward

As noted in the WestEd/Taylor report, it is reasonable for the legislature to consider phasing in the additional funding required to meet cost targets established in accordance with the accountability goals. Phase in requires consideration of two important factors:

- Continued changes in the competitive wages for school employees, most notably teachers. That is, the *inflation factor* which should be used in adjusting cost targets for out years is a comparable wage inflation factor,[2] not a consumer price index. The cost of providing comparable education services over time depends on the wages necessary to continue recruiting and retaining a similarly qualified teaching workforce, and not on changes to the price of a loaf of bread or gallon of gasoline (as per a CPI).
- The legislature should be aware that if they and/or the Kansas Board of Education decide to raise outcome standards further, the costs of achieving those standards will be higher, and the funding targets must be accordingly adjusted.

Finally, cost studies are rarely if ever translated directly into state school finance policy – adopted "as is" so-to-speak (Appendix B). The 2006 Post Audit study included a cost model estimated by Duncombe and Yinger, but then Post Audit staff translated that study into a structure and series of estimates for adoption in policy, making many reasonable changes, and some objectionable (noted in following report) ones.

The most reasonable path forward might be to seek ways to introduce new funding into the formula structure adopted in 2007 and make adjustments to weights to better align with Taylor's cost estimates, rather than attempting to adopt an entirely new formula.

The present WestEd Taylor study applies rigorous methods to high quality (higher than previously) data to arrive at reasonable estimates of the cost of achieving the legislature's constitutional mandate. The findings of the study are highly correlated with those of the two previous studies. Taken as a whole, the present study, and two which came before it, provide reasonable, empirically based evidence for reforming and funding the state school finance system to meet constitutional demands.

[2] http://bush.tamu.edu/research/faculty/Taylor_CWI/

Previous Cost Studies

On March 2, 2018, Jesse Levin of the American Institutes for Research issued his preliminary summary and critique of prior cost studies performed on behalf of the Kansas Legislature. Those studies included:

1. Analyses by Augenblick and Meyers (A&M) released in 2002 which including base cost estimates derived via *Successful Schools* analysis (average spending analysis) and base and additional costs (related to student needs, etc.) derived via *Professional Judgment* analysis.
2. The 2006 study prepared by the Legislative Division of Post Audit, which included *Cost Function* model-based estimates prepared by William Duncombe and John Yinger of Syracuse University, input based "base cost" analysis (cost of basic curricular mandates) prepared by LPA staff, and a hybrid funding formula proposal guided in part by the DY cost model, with additional assumptions introduced by LPA staff.

The first of these studies (A&M) provided guidance to the court during *Montoy v. Kansas* for determining the legislature's constitutional obligation to "make suitable provision for finance of the educational interests of the state." But, the A&M study never served to directly inform reforms to the School District Finance Act.

The second of these studies provided the basis for reforms to the School District Finance Act to be phased in from 2007 forward. Specifically, legislation adopted relied on recommendations provided by LPA staff, based only in part on the DY cost model estimates. Estimates from the LPA/DY study also informed subsequent judicial analysis during the course of the Gannon litigation.

As I explained in a brief prepared on behalf of Schools for Fair Funding, Inc. in 2006, the modifications made by LPA staff – among which was the choice to assume that federal aid would cover a significant portion of student need weighting for low income students and English language learners – served to significantly undercut the provision of constitutionally adequate funding for the state's highest need districts. Below is an excerpt from my 2006 review of the LPA adaptation of the Duncombe and Yinger estimates.

Table 1 displays the effects of LPA's modifications to Duncombe and Yinger's cost estimates across the state's largest districts. Notably, the districts most harmed by the LPA modifications are those with very high rates of limited English proficient students, including Kansas City, Garden City and Dodge City. Liberal, too small to appear on this list, is similarly harmed. **Even if the LPA Appendix 16 cost estimates were fully funded by SB 549, these districts would fall $500 to $700 per pupil below their actual estimated needs to achieve State Board of Education mandated outcome levels.** Because LPA added back in such factors as new and ancillary new facilities weight, Olathe ends up with a cost per pupil estimate in LPA's Appendix 16 nearly $500 per pupil higher than the cost estimate in Duncombe and Yinger's Appendix F.

Table 1

Actual Costs of Outcomes (D&Y) for 2006-07 Compared to Post Audit Version of Outcome-Based Costs (excl. sped, trans, voc)

District	Percent Free Lunch	D&Y '07 Cost per Pupil (Appendix F)	Appendix 16 Post Audit (excl. Sped, Voc., Trans.)	Unmet Obligation (rel. to LDPA Appendix 16 excl. Sped, Voc., Trans.)
Dodge City	60%	$7,215	$6,451	$764
Kansas City	66%	$8,254	$7,624	$630
Garden City	48%	$6,697	$6,186	$511
Derby	23%	$5,590	$5,429	$161
Shawnee Mission	12%	$5,415	$5,260	$155
Lawrence	22%	$5,604	$5,452	$152
Salina	36%	$5,884	$5,736	$148
Wichita	59%	$7,375	$7,257	$118
Auburn Washburn	16%	$5,084	$5,082	$2
Blue Valley	2%	$5,194	$5,202	-$8
Topeka	56%	$7,075	$7,269	-$194
Junction City	35%	$5,867	$6,126	-$259
Maize	7%	$5,084	$5,345	-$261
Olathe	12%	$5,354	$5,828	-$474

Dr. Levin's Overview and Critique

Dr. Jesse Levin provides an overview and critique of several aspects of the studies noted above. Dr. Levin also provides general guidance regarding costing out methods:

- Dr. Levin disregards *Successful Schools* methods generally as not meeting basic requirements for "cost" analysis, as it fails to address various factors know to influence the costs associated with achieving desired outcomes. On this point I concur. While successful-schools analyses informed the policy recommendations made by A&M in their original report, successful schools analyses have not played an ongoing role in either informing legislated reforms or judicial evaluation of the school finance system.
- Dr. Levin raises several concerns regarding the LPA input-oriented approach of studying the expenditures on programs and services associated with complying with state statutes and regulations. This input-oriented analyses has also had little (or no) bearing on subsequent legislation or judicial analysis, except perhaps to provide guidance on setting spending levels on those categories of spending not included in the DY cost model.
- Dr. Levin raises concerns that he and I, and Dr. Taylor and I raise regarding the precision of using *Evidence Based* models for determine the costs of meeting state specific (including constitutional, statutory or regulatory) standards. That is, that the outcome measures included in studies from which the evidence basis is drawn may not be aligned with the standards in question. Nonetheless, an evidence-based approach can provide a template for identifying and costing out the inputs/resources for a basic school prototype (much the same as in professional judgment analysis) which may then be reconciled with cost model estimates based on a state's own standards and measures.
- Dr. Levin describes *Cost Function Modeling* as follows: "a comprehensive education cost function model considers spending as a function of a) measured outcomes, b) student population characteristics, c) setting characteristics (economies of scale, population sparsity), d) regional variation in input prices including competitive wages, and e) **factors affecting spending that are not associated with outcomes** ("efficiency" per se)." (emphasis added) Dr. Levin also notes that "inefficiency" per se, as identified via a cost model merely indicates that some spending is not associated with the measured outcomes in the model, but not that the spending is necessarily unimportant. In fact, that spending might be associated with important outcomes or standards not included in measures used in the model. Dr. Levin further explains that: "Factors that contribute to this type of measured "inefficiency" are also increasingly well understood. For one, local public school districts with greater fiscal capacity – greater ability to raise and spend more – are more likely to do so, and may spend more in ways that do not directly affect measured student outcomes." This declaration is of non-trivial significance in cost model estimation.
    - The cost modeling approach used by Duncombe and Yinger explicitly accounts for factors which indirectly influence school district efficiency – factors associated with "fiscal capacity" and with "public monitoring." In their Kansas cost model, DY use the following measures (Fiscal Capacity: Consolidated Districts, per pupil Income, per pupil Property Values, Tax Aid Income Ratio; Public

Monitoring: Local Tax Share, % Adults College Educated, %65 or Older, % Owner Occupied Housing). Notably, some measures overlap categories.

- Importantly, these measures help to account for spending variation which is not associated with outcome variation, but is nonetheless predictable. Excluding these measures yields a model of spending which suffers from *Omitted Variables Bias.* When making cost predictions, these "efficiency variables" can be set to specific, constant levels (e.g. what would the district have spent if it had average "fiscal capacity" and/or "public monitoring" characteristics?) to remove the spending variation associated with these factors. DY set their efficiency measures to the 67%ile – so as to predict "costs" for districts that are at the top third in efficiency characteristics.
- Alternative approaches to cost modeling used by some authors do not include these factors known to contribute to spending variation, instead leaving that variation in a random error term, where a portion of that random error term is presumed to represent efficiency (based on a pre-determined statistical distribution). But that error term is, in fact, not random as it includes the omitted variables bias noted here, and thus cost projections based on such a model may be inaccurate.

Dr. Levin raises several specific and handful of broader concerns regarding the two prior studies done on behalf of Kansas Legislators. Dr. Levin raises concerns that the studies are now dated. This concern relates to the methods, data and findings of the A&M study, and to the underlying data and findings (though not the methods) of the DY cost model. Regarding PJ methods (discounting SS altogether), Dr. Levin suggests that significant improvements have been made to these methods over time which serve to enhance their reliability and validity, and in some specific cases precision.

Notably, the A&M Kansas PJ study was among the earlier studies of its kind, and the first in which A&M convened panels to consider multiple prototypes of different sizes in order to better understand costs associated with economies of scale. Dr. Levin notes that newer studies have used alternative and redundant panel configurations in order to cross-check (blind comparisons) resource recommendations. Dr. Levin also refers to "weak" validity tests of the kind he and colleagues used in New Mexico for evaluating adequacy cost estimates, such as comparing adequacy/cost gaps to existing outcome gaps. Dr. Levin explains:

> "If the model is working as intended so that adequate funding is provided in an equitable manner that affords all students an equal opportunity to achieve regardless of their needs or location, then we should see a systematic relationship between a district's relative need (how much more/less they need to provide a sufficient education) and student outcomes such as achievement on standardized tests.
>
> As an example, previous studies have performed this type of validation analysis for large-scale costing-out studies in New Mexico (Chambers et al., 2008a) and New York (Chambers et al., 2004a; Chambers, Levin & Parrish, 2006). The analysis involves calculating the funding shortfall or *Adequacy Gap*, which is a

district-level measure defined as the relative difference between the projected necessary per-pupil funding to provide a sufficient education and actual per-pupil funding."

Dr. Levin identifies a number of additional concerns:

- Dr. Levin implies that, due to district aggregation of FRL counts, the PJ method may not have been sufficiently sensitive to child poverty concentrations across districts.
- Dr. Levin raised concerns about steps taken by LPA to convert the DY cost model into formula recommendations, including the removal of federal funds from student need weights. While Dr. Levin illustrates that this step did not alter the relative differences between At Risk and Bilingual weights, Dr. Levin notes that the presumptive formula by which federal aid must be allocated to close the gaps left by removing it, may not be feasible or compliant with specific federal regulations.
- Dr. Levin discusses a recent formula change in California – the Local Control Formula – as a basis for considering the possibility of poverty concentration affecting costs, as opposed to poverty by intersection with population density as found in the DY model.

Comments on Dr. Levin's Preliminary Review

I will highlight a few key points here that are largely consistent with Dr. Levin's underlying arguments. But first, it is important to disregard outright and references to or comparisons with California's LCF formula as a whole or with regard to specific weights, design or magnitude. The LCF was not based on any empirical analysis of cost and the LCF poverty concentration weight not based on any modeled effect of the costs associated with poverty concentration. I concur however, that one might reasonably identify and estimate the magnitudes of such costs via rigorous methods.

An especially important issue raised by Dr. Levin is that of reliability and validity of cost study findings, and advancements made in the period following the original Kansas A&M study. In fact, the first academic literature which addresses these questions emerges at the time of the second Kansas study – the LPA DY cost model study. As explained above, Dr. Levin proposes an approach – a weak validity test – which involves comparing "adequacy gaps" with "outcome gaps." This test is drawn from three articles published in 2006:

- Chambers, J., J. Levin, and T. Parrish. 2006. "Examining the Relationship Between Educational Outcomes and Gaps in Funding: An Extension of the New York Adequacy Study.** Peabody Journal of Education 81(2): 1-32.
- Baker, B. D. (2006). Evaluating the reliability, validity, and usefulness of education cost studies. *Journal of Education Finance, 32*(2), 170-201.
- Duncombe, W. (2006). Responding to the charge of alchemy: Strategies for evaluating the reliability and validity of costing-out research. *Journal of Education Finance*, 137-169.

Interestingly, two of the three articles actually apply Dr. Levin's recommended test to the two Kansas cost studies, along with additional more rigorous checks on reliability and validity.

First, in Table 6 from my article, I show that adequacy ratios (current spending as % of adequacy target) in both Kansas Studies are positively associated with outcome measures, with the DY cost model having stronger correlations (around .6). Other studies have weaker and even negative correlations (Taylor, Texas A&M cost model) between adequacy gap estimates and actual outcomes.

Figure 1

Table 6. Correlations of Poverty and Student Outcomes with Adequacy Ratios (K–12 districts enrolling >2,000 pupils)

State	Method	Correlation with Adequacy Ratio (Actual/Adequate)			
		Poverty[a]	Reading	Math	Graduation
New York	Cost function: Duncombe and Yinger 2002	–0.524	0.359[b]		
Texas	Cost function: Reschovsky and Imazeki 2004	–0.509	0.646	0.610	
New York	Professional judgment: American Institutes for Research and Management Analysis and Planning 2004	–0.449	0.293[b]		
Nebraska	Professional judgment: Augenblick and Myers 2002	–0.417	0.676	0.604	0.320
Nebraska	Cost function	–0.401	0.702	0.641	0.366
Minnesota	Cost function: Data Envelopment Analysis	–0.390	0.544[c]	0.498[c]	
Kansas	Professional judgment: Augenblick and Myers 2002	–0.317	0.445	0.372	
Kansas	Cost function: Duncombe and Yinger 2006	–0.613	0.605	0.572	
Texas	Professional judgment: Management Analysis and Planning 2004	–0.102	0.201	0.141	0.244
Texas	Cost function: Texas A&M University 2004	0.257	–0.311	–0.240	–0.160
Arkansas	Evidence based: Odden et al. 2003	0.331	–0.395	–0.455	

[a]School year 2000 subsidized lunch rate (NCES Common Core of Data, Fiscal/Non-Fiscal Longitudinal File).
[b]200-point comprehensive index.
[c]Index score (rather than percentage proficient or higher).

Baker, B. D. (2006). Evaluating the reliability, validity, and usefulness of education cost studies. *Journal of Education Finance, 32*(2), 170-201.

In my article, I also showed in Table 4 that the cost function results were very highly correlated with the A&M PJ results, with a correlation between district level cost estimates across the two studies of .879 for all districts and .734 for large districts. The reliability across these studies is greater than that for other states where multiple studies have been done, including where alternative cost functions have been estimated.

Figure 2

Table 4. Correlations Between Cost Estimates Provided by Alternative Methods in the Same State

	All K–12 Districts			Large K–12 Districts (>2,000)		
	Cost Function 1	Cost Function 2	Current Resources[a]	Cost Function 1	Cost Function 2	Current Resources[a]
Nebraska[b]						
Cost function			0.472			0.227
Professional judgment (Augenblick and Myers 2002)	0.538		0.551	0.784		0.020
Kansas						
Cost function (Duncombe and Yinger 2006)			0.635			0.215
Professional judgment (Augenblick and Myers 2002)	0.879		0.742	0.734		0.074
Texas						
Cost function 1 (Texas A&M University 2004)			0.613			0.505
Cost function 2 (Texas A&M University 2000)	0.871		0.596	0.717		0.623
Cost function 3 (Reschovsky and Imazeki 2004)	0.683	0.733	0.379	0.736	0.797	0.504
Professional judgment 1 (Management Analysis and Planning 2004)[c]	0.666	0.599	0.299	0.815	0.752	0.555
New York						
Cost function 1 (Duncombe, Lukemeyer, and Yinger 2004)			0.517			0.478
Professional judgment (American Institutes for Research and Management Analysis and Planning 2004)	0.546		0.836	0.732		0.681

[a]Current operating expenditures per pupil (includes expenditure of federal funds).
[b]As reported in Bruce D. Baker (2005), "Nebraska's School Finance System Fails to Provide Equal Opportunities for Nebraska School Children," prepared for plaintiff districts in the case of *Douglas County School District v. Heineman*, p. 47.
[c]Includes district-level costs for only a selected group of districts.

Baker, B. D. (2006). Evaluating the reliability, validity, and usefulness of education cost studies. *Journal of Education Finance*, *32*(2), 170-201.

While Dr. Levin raises concerns that the 2002 A&M study did not internally include reliability and validity checks, these findings provide convincing evidence that the study yielded reliable and valid results (though we did not know that until years later).

William Duncombe applied additional tests of reliability and validity to his cost model findings. First, Duncombe estimated district cost indices for each year of data in the study and compared their consistency over time. His Table 3 reveals a high degree of consistency among district cost indices from year to year – which in part explains why district cost estimates from a 2002 PJ study might remain so highly correlated with district cost estimates from a cost function estimated years later.

Figure 3

Table 3. Comparisons Between Cost Indices for Different Years for Kansas School Districts

	2000–2004	2000–2002	2000–2001	2003–2004
Correlations				
2000–2004	1			
2000–2002	0.985	1		
2000–2001	0.954	0.984	1	
2003–2004	0.947	0.984	0.969	1
Averages by census region				
Large central cities	124.1	131.2	115.0	129.3
Medium cities	92.3	93.6	98.8	91.4
Urban fringe of large cities	87.3	87.7	86.5	85.9
Urban fringe of medium cities	98.2	91.5	92.2	91.5
Large town	101.2	103.4	98.0	101.7
Small town	95.7	97.2	95.3	94.8
Rural metro	105.2	104.6	105.6	106.7
Rural nonmetro	94.3	94.1	95.2	93.0

Duncombe, W. (2006). Responding to the charge of alchemy: Strategies for evaluating the reliability and validity of costing-out research. *Journal of Education Finance*, 137-169.

Duncombe's most compelling analysis, which goes beyond that suggested by Dr. Levin, is a predictive validity test which he uses to select the optimal cost model. For this test, Duncombe estimates 4 different versions of the cost model to data for years 1-5 and uses that model to predict actual spending for year 6.

When blindly predicting the subsequent year of data, two issues are of interest. First, on average, how much prediction error is there? (expressed as absolute value of the percent error). Second, is there bias in the predictions (more over or under prediction)? Answering these questions across four models a) allows a general determination of validity of the method and b) allows the researcher to identify which specific model, among models is preferable (most valid).

This specific test is what led their team to select the model which included an interaction term between poverty and population density to capture urban poverty related costs. That is, the poverty-density interaction term was selected by a rigorous cross-validation technique.

Figure 4

Table 6. Estimates of Forecasting Error
(difference between predicted and actual as a percentage of actual)

Distribution	Naive Forecast	Base Model	Model 2	Model 3	Model 4
Bias (percentage error)					
Mean	−7.1	1.1	4.8	0.4	1.2
Median	−6.8	0.2	4.3	−0.6	1.2
Minimum	−50.7	−31.3	−31.6	−27.5	−34.8
5th percentile	−35.2	−17.4	−16.7	−15.7	−16.5
10th percentile	−26.5	−11.3	−9.1	−11.6	−9.9
25th percentile	−16.3	−5.9	−2.9	−6.1	−5.0
75th percentile	2.0	7.6	12.0	6.7	7.6
90th percentile	12.5	16.1	21.5	14.2	15.7
95th percentile	18.7	20.3	28.0	18.0	17.8
Maximum	27.9	39.5	46.7	35.2	34.8
Accuracy (absolute percentage error)					
Mean	13.1	8.5	10.4	7.9	7.9
Median	10.6	6.7	7.7	6.4	6.2
Minimum	0.0	0.1	0.0	0.0	0.0
5th percentile	0.9	0.6	0.6	0.8	0.7
10th percentile	1.6	1.4	1.2	1.4	1.2
25th percentile	4.6	3.1	3.5	2.9	2.6
75th percentile	19.1	11.6	14.9	11.6	11.1
90th percentile	27.4	19.5	23.9	17.6	17.6
95th percentile	35.2	22.8	28.3	19.3	20.9
Maximum	50.7	39.5	46.7	35.2	34.8

Note: Naive forecast is based on the log of per-pupil base spending regressed on the log of the performance index.

As discussed previously, three alternative specifications of the cost model are estimated, and their forecasting bias and accuracy statistics are reported in Table 6. Removing the interaction of free lunch share and pupil density (Model 2) appears to increase forecasting error and led more frequently to overestimates of spending. Using a different functional form for enrollment (Model 3) and including squared efficiency variables (Model 4) appears to marginally improve forecasting accuracy for about a quarter of the districts.

Duncombe, W. (2006). Responding to the charge of alchemy: Strategies for evaluating the reliability and validity of costing-out research. *Journal of Education Finance*, 137-169.

Additional analyses of the Kansas cost studies appear in a paper by me, Lori Taylor and Arnold Vedlitz of Texas A&M University for the National Research Council in 2008. Specifically, “Table 2” from that study compares the implicit poverty adjustments from various cost studies, including the two Kansas studies. The Table reveals that the DY cost model had a stronger poverty effect than the A&M PJ analysis. But, the PJ finding was consistent with PJ findings in Washington and Pennsylvania. The DY cost model findings were also consistent with other cost models – with poverty adjustment smaller than in some other studies (Minnesota, Missouri[1] and Texas) and marginally larger than others (Missouri[2]). The notable outlier among cost function studies here is the very low poverty adjustment from the Taylor, Texas A&M Cost model (.395).

The Taylor Texas Cost model differs from the Kansas, Missouri and other Texas model in that it a) does not include indirect controls for efficiency and b) does not account for the endogeneity of the outcome measure. The Taylor/Texas model and Rhode Island (my own) models both use Stochastic Frontier analysis accounting for inefficiency in an “error term.”

Figure 5

Table 2: The Implicit Adjustments for Student Poverty

State	Study Type	Implicit Poverty Adjustment	Baseline Cost Estimate
Arkansas	Evidence Based	0.225	$6,115
Kansas	Cost Function	0.965	3,982
Kansas	Professional Judgment	0.681	6,172
Minnesota	Cost Function	1.679	4,932
Missouri	Cost Function	0.992	4,013
Missouri	Cost Function	0.802	4,900
New York	Cost Function	1.346	5,511
New York	Professional Judgment	0.915	7,196
Pennsylvania	Professional Judgment	0.616	6,436
Rhode Island	Cost Function	0.672	5,725
Texas	Cost Function	0.395	4,030
Texas	Cost Function	1.273	3,147
Washington	Professional Judgment	0.581	6,841

Note: The implicit poverty adjustments are coefficient estimates from a regression of the district-level cost of an adequate education (in logs) on the log of enrollment, the log of enrollment squared, the share of students in poverty and the NCES Comparable Wage Index. In all cases, the coefficient estimates are significantly different from zero at the 1-percent level. Complete regression tables available upon request.

Baker, B. D., Taylor, L. L., & Vedlitz, A. (2008). Adequacy estimates and the implications of common standards for the cost of instruction. *National Research Council*.

Comments on WestEd/Taylor Study

Dr. Taylor's updated cost model adopts (and advances) what I consider to be the most credible and useful methods for estimating costs associated with meeting specific outcome standards. Importantly, the methods used differ from previous cost models estimated by Dr. Taylor and are more similar to, and an extension of methods applied in Kansas back in 2006 by William Duncombe and John Yinger of Syracuse University. Notably, Dr. Taylor's model has the advantage over the DY model of data quality improvements over the past decade and use of multiple outcome measures. Two key features of the Taylor model, which replicate (and/or extend) the strategy taken by DY are:

- Consideration that the outcome measures of interest are endogenous and use of instrumental variables (2SLS) estimation (though combining the two-stage approach with a stochastic frontier approach (for the second stage model);
- Inclusion of indirect controls for inefficiency to account for omitted variables bias in the spending measure (spending not associated with outcome variation, but predictable as a function of fiscal capacity, competition and public monitoring characteristics of districts).

There are, however, a few non-trivial differences:

- Taylor's approach continues to differ in one regard from the DY approach, and that is in the use of a polynomial (2nd order) curve to represent variations in costs associated with economies of scale (where DY uses a series of district size categories).

- Taylor also did not test for or acknowledge potential variation in poverty related costs in relation to poverty concentration, urbanicity or population density.

Additionally, Dr. Taylor does not report how the final model parameters were selected and/or whether a specific validation method was used for model selection. DY, in a series of related academic articles using their Kansas model, elaborated that model selection involved prediction accuracy tests – specifically, fitting different models to selected years of data, and predicting district spending for subsequent years of data not included in model fitting. This procedure is what led DY to adopt the model which include an interaction term between poverty and population density, which eventually led the legislature to include a high-density district poverty weight in the revised formula.

Dr. Levin reiterates in his review the importance of reliability and validity checks on cost estimates, whether from cost modeling or alternative methods. In the sections that follow, I will apply methods suggested by Dr. Levin to Dr. Taylor's findings.

Key Findings

Figure 6 presents a modified version of Dr. Taylor's Table 13, displaying the estimates from the cost model. Three outcome measures are included. A variety of geographic factors including economies of scale, wage variation, population density and a rural indicator. Student need factors the usual subset of a poverty measure, a measure of English language learners and a measure of the share of children with disabilities. The model also accounts for grade ranges served (a useful alternative is to account for the share of children within certain grade ranges). And, importantly, the model accounts for a variety of factors which may predict variation in spending which is unassociated with outcomes – inefficiency/efficiency factors. The basic elements of the model are quite similar to those of the DY model which is provide in Appendix A.

Figure 6

Table 13. Cost Model Coefficient Estimates

Variable	***Coeff. (SE)***
Outcomes	
Normal Curve Equivalent	5.295*** (-0.607)
Graduation Rate	1.244*** (-0.262)
Graduation Rate * High School	0.696*** (-0.0995)
Geographic Factors (Scale & Wage Variation)	
District Enrollment	-1.444*** (-0.0568)
District Enrollment squared	0.0991*** (-0.00378)
Salary index (log)	1.373*** (-0.279)
Population Density	0.166*** (0.018)
Population density* Salary Index	-0.510*** (-0.0414)
Rural indicator	0.0505*** (-0.0112)
Student Needs	
% Economically Disadvantaged	**0.886*** (-0.078)**
% English Language Learner	**0.226*** (-0.0667)**
% Special Education	**2.157*** (-0.226)**
% English Language Learner, sq	-0.623*** (-0.109)
% Special Education, sq	-6.135*** (-0.674)
School/District Structural Characteristics	
Elementary grades served	-0.129*** (-0.016)
High school grades served	-0.508*** (-0.0909)
Efficiency & Endogeneity Controls/Corrections	
AYP Schoolyear = 2016	-0.0364*** (-0.00591)
First stage Residuals, NCE	-5.102*** (-0.609)
First stage residuals, Graduation	-1.454*** (-0.271)
Herfindahl Index, log	0.797*** (-0.249)
Border metro	2.320*** (-0.372)
% Owner occupied	7.293*** (-1.321)
% Over 60	-2.316 (-1.496)
% College	-12.06*** (-1.542)
Constant	9.644*** (-0.357)
Usigma	-7.214*** (-0.958)
Vsigma	-4.095*** (-0.0418)
Observations	2,310

Robust standard errors in parentheses

**** p<0.01, ** p<0.05, * p<0.1*

Outcomes and Costs

In their 2006 model, Duncombe and Yinger found:

> "We found a strong association between the amounts districts spend and the outcomes they achieve. In the cost function results, a 1.0% increase in district performance outcomes was associated with a 0.83% increase in spending—almost a one-to-one relationship. This means that, all other things being equal, districts that spent more had better student performance. The results were statistically significant beyond the 0.01 level, which means we can be more than 99% confident there is a relationship between spending and outcomes."

The WestEd/Taylor model echoes this conclusion, with new and different outcome measures, thus revealing differences in the magnitude of the relationship. The authors note:

> "Table 17 presents coefficient estimates and standard errors from the cost function analysis. As the table illustrates, the analysis finds a strong, positive relationship between educational outcomes and educational costs, once differences in scale, need and price are taken into account. Consider first the Condition NCE scores. The estimation indicates that a one percentage point increase in academic performance is associated with a 5 percent increase in cost. Similarly, a one percentage point increase in the graduation rate is associated with an 1.2 percent increase in cost at lower grades and a 1.9 percent increase in cost at the high school level." (p. 61)

Put bluntly: Money matters. It costs more to achieve higher outcomes, and as further explained in the report, it will cost more to achieve the states desired outcomes which are higher than present outcome levels for many children in many districts and schools.

Efficiency

Regarding efficiency, the Taylor model finds that Kansas school districts are highly efficient in their current production of outcomes, given their current spending levels. Specifically:

> "In Model 1, the average cost efficiency score was 0.956, indicating that buildings were producing nearly 96% of their potential output, on average. Given that inefficiency in this context means unexplained expenditures, not necessarily waste, and that many buildings may have been producing outcomes that were not reflected in test scores, the average efficiency level was quite high."

The policy implication of this finding is that the legislature cannot expect to simply *squeeze* even higher outcomes from Kansas schools and districts at current spending levels. That current spending levels are in fact insufficient to achieve desired outcome levels, by improving efficiency alone.

Student Needs

The Taylor model also reveals logical relationships between student need factors and costs associated with achieving common outcome goals. Specifically, the authors find a poverty coefficient of .89, which they note is much higher than the current formula weight (which was derived in part from the 2006 Post Audit interpretation that federal funding could be removed from the estimated cost when setting state policy – an assumption which raised some concern in Dr. Levin's report).

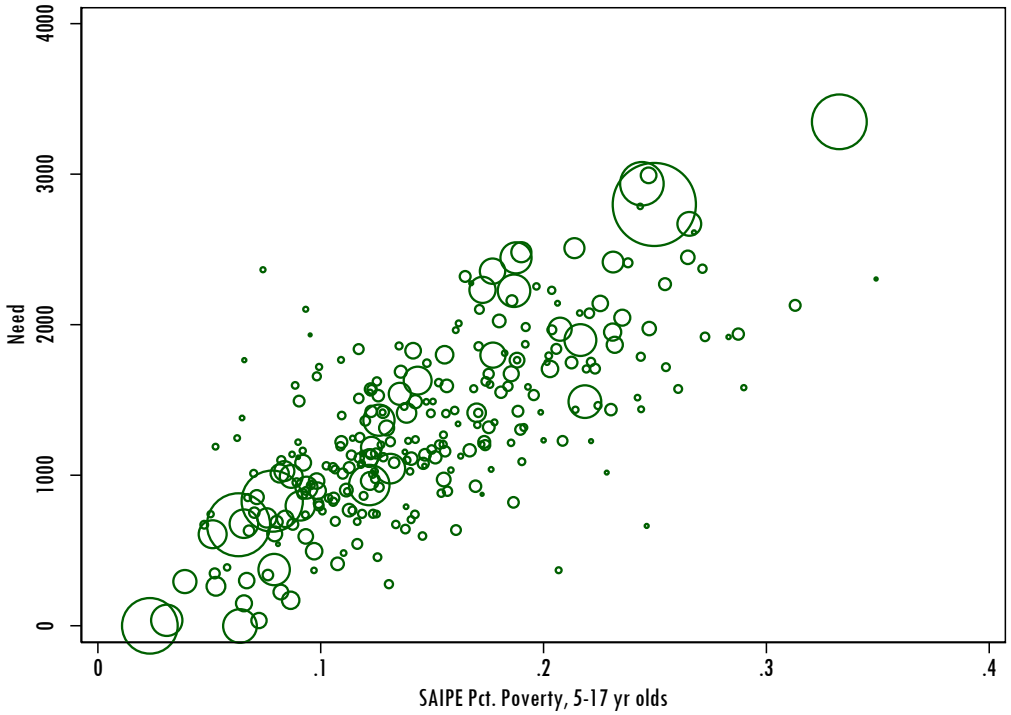
Taylor, like Duncombe and Yinger had difficulty isolating a significant ELL weight (due to conflation with other factors), but did find a coefficient around .2.

The WestEd/Taylor report notes in the text, a negative special education effect and then attempts possible explanations for that effect, but Taylor's Table 13 (figure 6 above) actually shows a positive special education cost effect, of logical magnitude.

Figure 7 shows the relationship between districts' student need funding generated by the "Student Need Index" from Taylor's Table 27, and U.S. Census Poverty (2016) rates for districts. Need funding is calculated by multiplying the Student Need Index for each district times the unique base cost for each district. Figure 7 shows that districts above 30% census poverty would receive over $3,000 per pupil in additional need related support. This is logical and consistent with prior Kansas studies, and, studies conducted in other state settings (more to follow).

Figure 7

Student Need Weight (Dollars per Pupil Generated) by Census Poverty Rate



Economies of Scale

Figure 8 shows the district size, or economies of scale index in relation to district enrollments. Taylor's cost model fits a U-shaped curve in relation to district size and spending (logged). Using this approach, Taylor's model infers that costs "bottom out" for districts between around 1,000 and 5,000 students, but then rise quite substantially as a function of size alone, for much larger districts, generating for these districts between $3,000 and $4,000 per pupil in additional funding (equivalent to the highest student need adjustment).

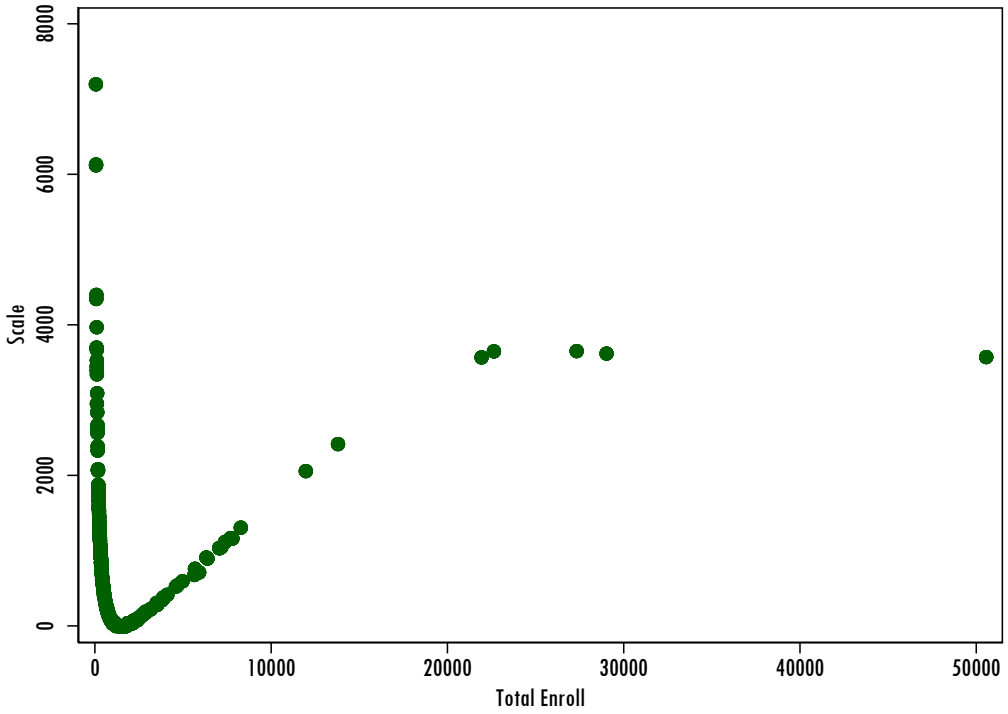
There are two potential problem areas here.

- Fitting the scale term in this way leads to the inference that large, very low need districts, need substantial additional funding simply because they are large, despite the fact that at least theoretically, a district of 20,000 students could operate as 10 districts of 2,000 students to achieve comparable cost efficiencies.

- Fitting the scale term in this way leads to a sharp dip in spending predictions for districts with 1,000 to 5,000 students, potentially driving down their predicted cost estimates below needed levels. The overall "curve" may fit the data reasonably well (Taylor Figure 11), but with these few distortions leading to the overestimation of costs for some and underestimation of costs for others (which might be revealed with DY-style prediction accuracy tests for forecast bias)

Figure 8

Scale Weight (Dollars per Pupil Generated) by District Enrollment



Reliability and Validity Checks

Here, I run a series of checks on the Taylor model findings based on those checks recommended by Dr. Levin in his preliminary report and in prior academic work by myself and William Duncombe in 2006, as well as work with Lori Taylor in 2008.

NOTE: The following analyses calculate district costs per pupil as per the explanation provided in WestEd/Taylor's Figure 8 (p. 65). That is, the various need/cost adjustments are assume "additive." Or:

Cost per Pupil = Base + (Base x Regional Index) + (Base x Scale Index) + (Base x Need Index)

However, this approach generates lower total cost estimates than applying a multiplicative (more common) approach to the need and cost weights, which nearly approximates the reported total cost estimates in the WestEd/Taylor report. That is:

Cost per Pupil = Base x Regional Index x Scale Index x Need Index

Clarification is required before taking steps toward adoption or adaptation into state school finance reforms.

Comparison to Other Cost Study Weights from Baker, Taylor & Vedlitz (2008) NRC Report

First, Lori Taylor, Arnold Vedlitz and I, in a paper for the National Research Council, compiled district cost per pupil estimates from several cost studies and fit regression models to those studies using common measures of child poverty and of competitive wage variation. Because all studies report their poverty and other adjustments differently, we used this method to equate the magnitude of those adjustments in the Table I included previously as Figure 5. Most cost function poverty estimates in that table fell from .80 to 1.2 (or higher). The Kansas DY cost function landed at .965 and the A&M PJ model at .681. Table 1 below shows that applying the same method to the Taylor Scenario A and B cost estimates yields poverty effects that are slightly smaller than for other cost function studies, but right between the A&M PJ and DY cost model for Kansas. Taylor reports a weight (based on the model coefficient itself) of .9. These are reasonable estimates of the relationship between child poverty and the costs of achieving common outcomes.

Table 1

Regression model determination of implicit poverty weight

	Scenario A		***Scenario B***	
	coef	**se**	**coef**	**se**
SAIPE Pct. Poverty, 5-17 yr olds	**0.709*** **	**0.119**	**0.726*** **	**0.121**
NCES CWI (extended)	-0.154**	0.068	-0.151**	0.069
ln_enroll	-0.614***	0.048	-0.611***	0.049
ln_enroll2	0.040***	0.003	0.040***	0.003
Intercept	11.497***	0.194	11.507***	0.199
Number of observations	277		277	
R2	0.538		0.529	

*note: *** p<0.01, ** p<0.05, * p<0.1*

Reliability: Comparison to Prior KS Cost Studies

Table 2 summarizes the correlations between district cost estimates from the three Kansas cost studies, a) across all districts, not weighted for student enrollment, b) across districts, weighted for student enrollment, and c) across districts with 2,000 or more pupils, weighted for student enrollment. The correlation tells us whether generally, those districts estimated as having higher or lower costs per pupil in one study, had higher (or lower) costs per pupil in the other studies.

Especially when applying weighting for district enrollment, or when looking at scale efficient districts, the correlations between the cost estimates from the DY study and the Taylor model are quite high – between .80 and .90. That is, we see a high degree of reliability across Kansas studies. The correlations between the DY and A&M study and the correlations between the Taylor and A&M study are also reasonable high.

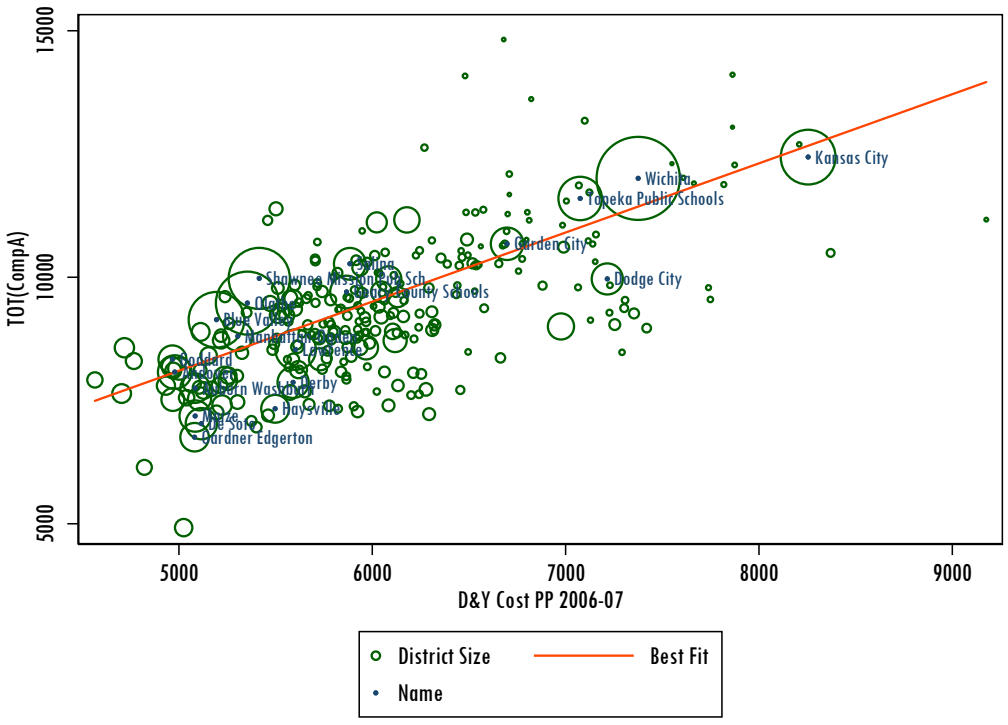
The take home point here is that the Kansas legislature now has a third study which largely tells the same story of which districts face higher and lower per pupil costs of achieving common outcome goals.

Table 2

	All		*All (Weighted)*		*Large (Weighted)*	
	DY Cost Model	A&M PJ	DY Cost Model	A&M PJ	DY Cost Model	A&M PJ
DY Cost Model	1.00		1.00		1.00	
A&M PJ	0.88	1.00	0.73	1.00	0.82	1.00
Taylor Maintenance	0.77	0.76	**0.87**	0.56	**0.90**	0.67
Taylor Scenario A	0.65	0.63	**0.81**	0.48	**0.86**	0.62
Taylor Scenario B	0.65	0.63	**0.83**	0.51	**0.88**	0.65

Figure 9 visually displays the clarity of the relationship between per pupil cost estimates from the DY study (horizontal axis) and Taylor Scenario A (vertical axis). Total cost figures are not adjusted for inflation, so the DY estimates are lower. The point of Figure 9 is to show that generally, districts receiving the highest per pupil cost estimates in 2006 received the highest in the current study and vice versa. Notable deviations include a somewhat lower (below the red line) estimate for Dodge City in the Taylor model, as well as higher estimates (above the red line) for Shawnee Mission, Blue Valley and Olathe (likely a function of the large district size weight).

Figure 9



Weak Validity Test (relating spending gaps to outcome gaps)

Here, I run a weak validity check explained and illustrated by Jesse Levin in his preliminary report – that is, to what extent are current funding gaps related to current outcomes. One would expect, for example that districts identified

as needing substantially more funding to achieve desired outcomes, would have relatively low outcomes, and vice versa. As Dr. Levin suggests and as I had done in my 2006 article, I estimate the correlations between district adequacy ratios (Current Spending per Pupil/Adequacy Cost per Pupil) and a variety of relevant outcome measures.

Due to time constraints and data convenience, I use the federal measure (F-33 Census Fiscal Survey) of current spending per pupil (subtracting food and transportation) from 2015 as the current spending comparison basis. Table 3 correlates adequacy ratios with re-scaled outcome measures from 2015 from the Stanford Education Data Archive (combining ELA and Math into a single index). Table 4 correlates adequacy ratios with a) rates of children scoring in Category 1 on Kansas State Assessments and, b) rates of children scoring in Category 3 or 4 on Kansas State Assessments.

Note that in my previous published work, I found that the adequacy ratios using the DY cost model were correlated at .605 with state language arts results and .572 with state math results. I found that the adequacy ratios using the A&M PJ study were correlated at .445 with language arts and .372 with math.

Table 3 shows somewhat lower correlations between adequacy ratios constructed using Taylor's Scenario A and Scenario B cost targets and ELA and Math scores from the Stanford Education Data Archive. Weighted, and for large districts only, also weighted, the correlations are between .310 and .474. In Table 4, using Kansas assessment data from 2017, correlations are even smaller.

Table 3

SEDA[3] Combined Outcome Index

	All	*All (Weighted)*	*Large (Weighted)*
Maintenance	0.196	0.409	0.474
Scenario A	0.160	0.310	0.365
Scenario B	0.177	0.343	0.406

Table 4

2017 KS Proficiency Rates

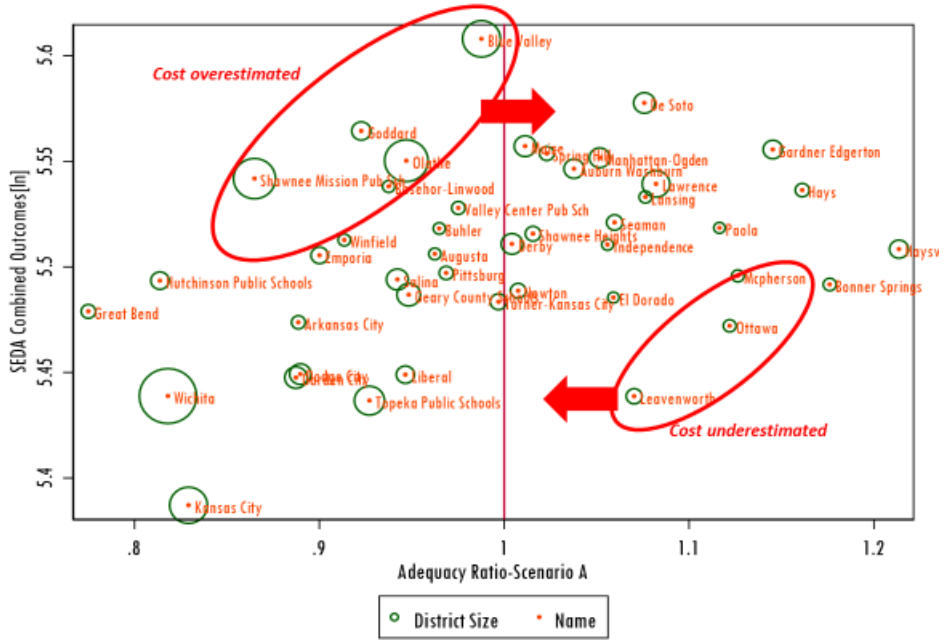
	ELA		*Math*	
	Scenario A	Scenario B	Scenario A	Scenario B
All Districts				
% Level 1	-0.029	-0.019	0.040	0.000
% Level 3 or 4	-0.038	-0.013	-0.061	0.002
All (Weighted)				
% Level 1	-0.260	-0.282	-0.247	-0.284
% Level 3 or 4	0.131	0.169	0.113	0.162
Large (Weighted)				
% Level 1	-0.342	-0.377	-0.358	-0.397
% Level 3 or 4	0.274	0.321	0.276	0.325

[3] Sean F. Reardon, Andrew D. Ho., Benjamin R. Shear, Erin M. Fahle, Demetra Kalogrides, & Richard DiSalvo. (2017). Stanford Education Data Archive (Version 2.0). http://purl.stanford.edu/db586ns4974.

These findings raise some questions about the cost predictions generated by the Taylor model, which are partly explained in Figure 10. Figure 10 shows the relationship, for districts with 2,000 or more pupils, between adequacy ratios using Taylor's Scenario A and my combined outcome index (log of ELA + Math NAEP scaled state assessment scores from SEDA). We would expect a reasonably tight diagonal from the bottom left to upper right corner of the figure. Wichita and Kansas City fall within those expectations (lower left), as do De Soto and Gardner-Edgerton (upper right). Wichita and Kansas City are estimated to need substantially more funding to achieve desired outcomes and, in fact, yield relatively low outcomes. De Soto and Gardner Edgerton are estimated as spending more than they would need to achieve desired outcomes, and in fact perform quite highly.

But, due largely to the district size factor – u-shaped curve – districts like Shawnee Mission and Olathe are estimated to need more resources to achieve desired outcomes, and Blue Valley is estimated as having only approximately what it would need to achieve desired outcomes. These three large, relatively low need districts, however, already have very high achievement levels, suggesting either that they are very efficient, or that the large district funding boost is overestimating their needs (this does not apply to Goddard or Basehor, which also fall in my oval). It is likely that these districts should be shifted to the right in the figure, actually having more than (Blue Valley) or similar to (Shawnee Mission) what they would need to achieve desired outcomes. It seems highly unlikely that Shawnee Mission's funding deficits would be near those of Kansas City and greater than those of Dodge City or Topeka, or that Olathe's needs would be similar to those of Topeka.

Figure 10



Just as the U-shaped curve may inappropriately boost cost estimates for these very large low need districts, districts with 2,000 to 5,000 pupils are caught in the dip of the curve. These include, for example, Leavenworth, Ottawa and Turner as shown in Figure 11. It seems likely that being caught in the dip of the scale curve is leading to an underestimation of cost for Leavenworth. The alternative interpretation is that Leavenworth is simply a less efficient district and should be performing much better at its current spending level. This seems unlikely. If Leavenworth and Ottawa were shifted to the left, and Blue Valley, Shawnee Mission and Olathe to the right along the horizontal axis, we would likely see a higher correlation between funding gaps and outcomes. That is, we would have stronger evidence that the cost estimates are valid.

Figure 11

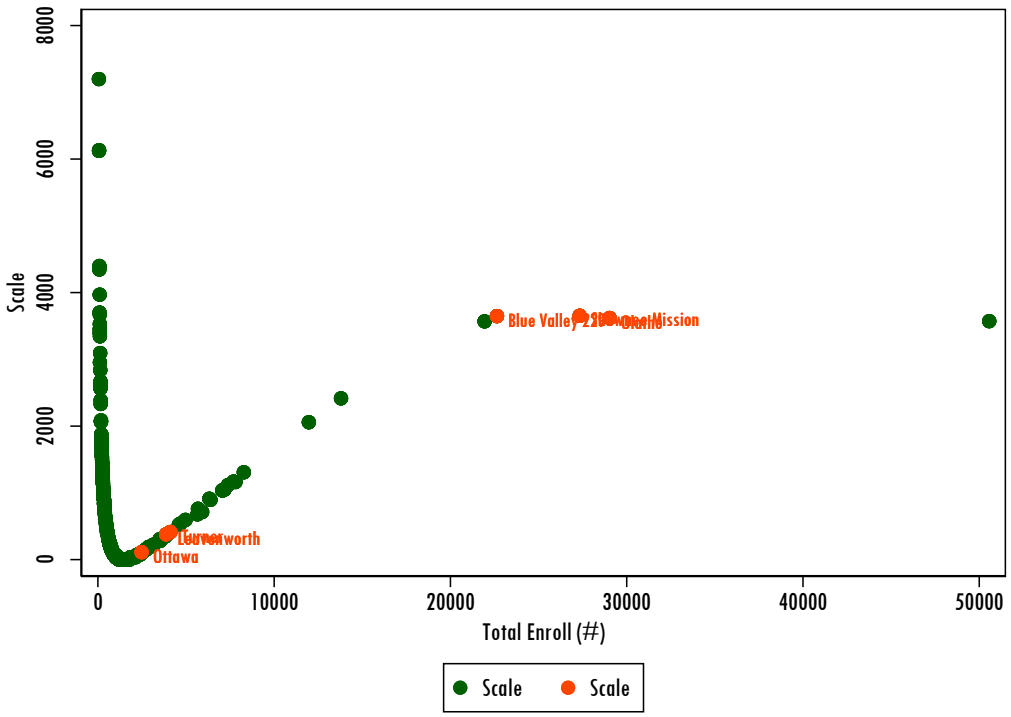


Table 5 provides a more detailed comparison of the three low-need very large districts benefited by the district scale term and three higher need mid-size districts caught in the dip. Perhaps the most useful contrast is between Shawnee Mission and Turner. These districts are immediately adjacent, and thus logically face similar labor costs. Shawnee Mission is much larger, but with less than half the rate of low income children, have the rate of ELL children and slightly lower rate of children with disabilities. Most cost analyses would find that Turner's per pupil costs, driven by student needs, exceed, at least marginally Shawnee Mission's per pupil costs, assuming economies of scale level off, rather than climb for large districts. The DY cost model estimated for 2007, Turner's per pupil cost at $5,968 and SMSD at $5,415, or about a 10% margin in favor of Turner. The Taylor model estimates the two to have similar maintenance costs, and Shawnee Mission to have higher costs of either performance improvement scenario. Current per pupil spending for the two districts is similar. Turner is estimate to face only a small deficit, if any, to achieve desired outcomes, and Shawnee Mission a much larger deficit, solely as a result of the economies of scale curve.

Table 5

District Name	*Ottawa*	*Leavenworth*	*Turner-Kansas City*	*Olathe*	*Shawnee Mission*	*Blue Valley*
Total Enroll (#)	2479	3873	4110	29029	27333	22640
Percentage Poverty (%)	42%	49%	63%	21%	28%	5%
Percentage ELL (%)	1%	2%	24%	11%	12%	3%
Percentage Special Ed (%)	11%	16%	11%	13%	9%	10%
Teacher Cost Index	1.34	1.37	1.54	1.53	1.56	1.56
Economies of Scale Index	1.03	1.10	1.11	1.97	1.97	1.97
Student Need Index	1.38	1.52	1.71	1.18	1.22	1.00
Cost at Maintenance	$ 7,939	$ 8,489	$ 7,634	$ 8,731	$ 8,433	$ 7,974
Cost of Scenario A	$ 7,902	$ 8,717	$ 8,575	$ 9,477	$ 9,977	$ 9,140
Cost of Scenario B	$ 7,976	$ 8,641	$ 8,876	$ 9,589	$ 10,015	$ 9,140
NCES Current Spending (2015, Excl. Food & Transportation)	$ 8,865	$ 9,329	$ 8,548	$ 8,975	$ 8,629	$ 9,027
Adequacy Ratio-Maintenance	112%	110%	112%	103%	102%	113%
Adequacy Ratio-Scenario A	112%	107%	100%	95%	86%	99%
Adequacy Ratio-Scenario B	111%	108%	96%	94%	86%	99%

Table 6 explores the cost implications of overestimating spending targets for Shawnee Mission, Blue Valley and Olathe, comparing their 2015 current spending levels (fed data, excluding food and transportation) against the cost targets. If we assumed that none of the three need additional funds to achieve desired outcomes, the default gap would be $0. Shawnee Mission may, in fact still require some additional resources. However, if we apply Taylor's Scenario A cost estimates, including the large district boost, these districts require total additional funding exceeding $50 million. That said, $50 million remains a relatively small share of the statewide cost of meeting adequacy targets, and some of this $50 million would be offset by raising targets for those districts caught in the dip, like Leavenworth or Turner.

Table 6

District Name	*Olathe*	*Shawnee Mission Pub*	*Blue Valley*
Total Enrollment	29,029	27,333	22,640
Cost of Scenario A	$ 9,477	$ 9,977	$ 9,140
NCES Current Spending (2015, Excl. Food & Transportation)	$ 8,975	$ 8,629	$ 9,027
GAP	$ 502	$ 1,348	$ 113
Total Cost	$ 14,579,957	$ 36,842,135	$ 2,553,854

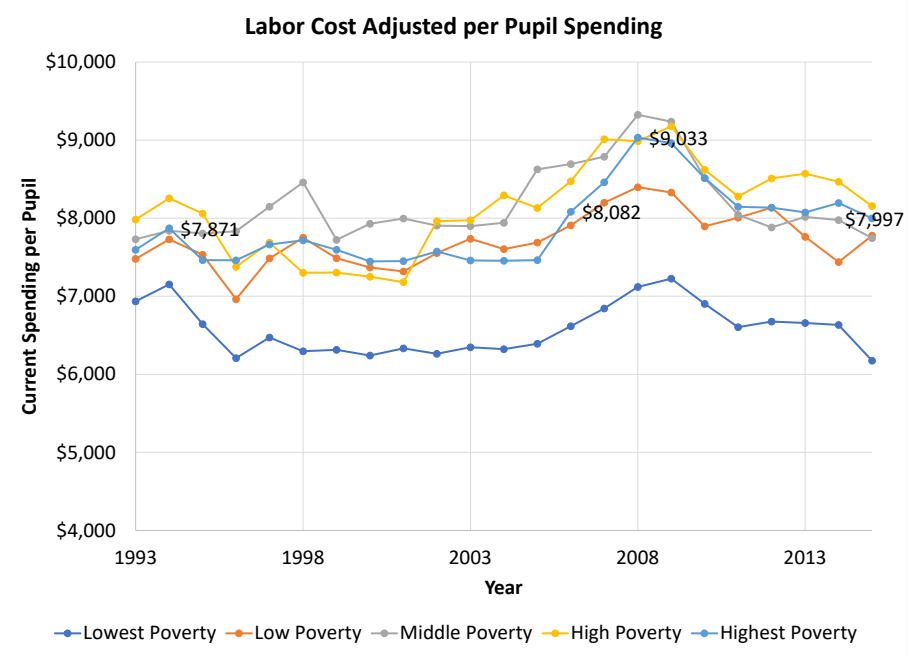
Why Have Total Costs of Adequacy Increased So Much?

A question that has been raised since release of the WestEd/Taylor study is – Why is the spending gap (to achieve adequate outcomes) so much larger, in the aggregate, now than it was in prior studies? There are many factors which can serve to explain the seemingly larger price tag estimated by Taylor, only a small margin of which can be attributed to possible overestimation of costs for low need large districts.

- First, 12 years after (and even more data years after) the DY and LPA studies, we are simply talking about larger dollar figures when not considering inflation adjustment.
- Second, Kansas like many states continues to raise and broaden its outcome expectations for kids, and higher outcomes cost more to achieve.

Further, current spending was declared inadequate in 2006, and was already measurably inadequate against either the A&M or DY/LPA targets. By 2007, inflation (comparable wage growth) adjusted spending per pupil in the highest poverty districts was just over $8,000 per pupil. That figure rose for the highest poverty districts as the Montoy remedy legislation was phased in, but later dipped to below 2007 levels. With higher outcome goals in play, and less spending than previously, the gap will necessarily be larger.

Figure 12

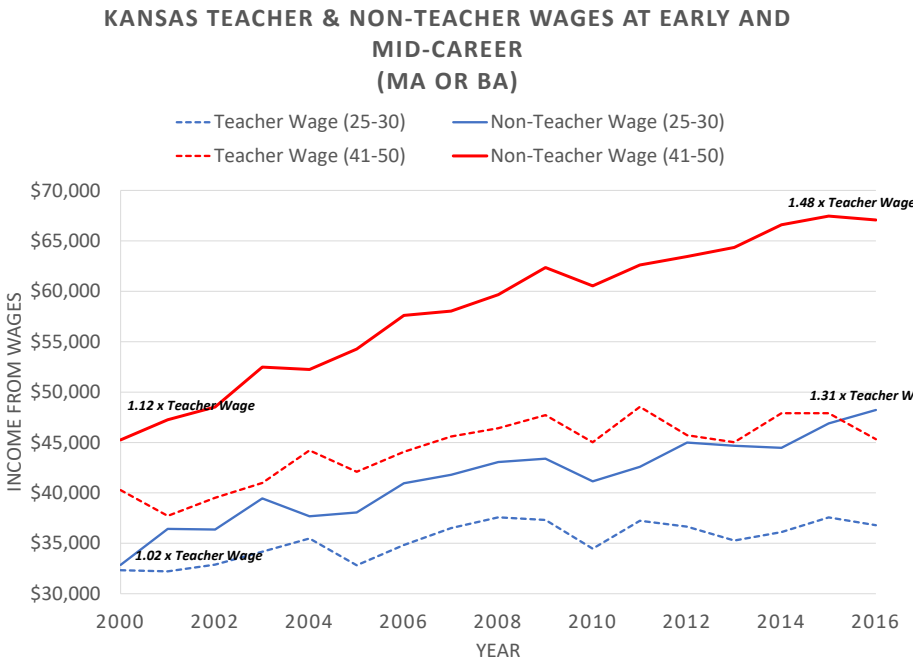


Baker, B.D., Srikanth, A., Weber, M.A. (2016). *Rutgers Graduate School of Education/Education Law Center: School Funding Fairness Data System.* Retrieved from: http://www.schoolfundingfairness.org/data-download

Figure 13 shows that over time, Kansas teacher wages have continued to fall further behind wages of similarly educated, same age non-teachers. Correcting this gap will require substantial infusion of funding, as implicated by

the Taylor cost model, which provides thorough consideration of labor costs and labor cost variation across district settings.

Figure 13

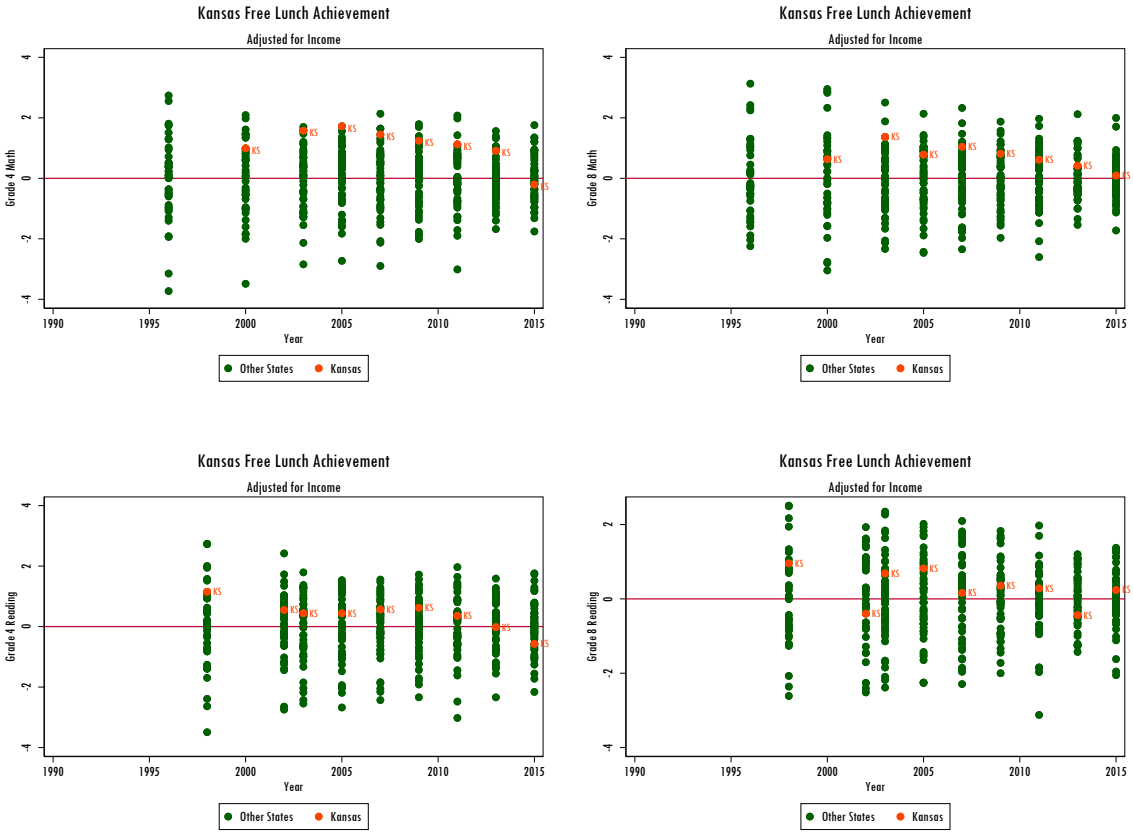


Baker, B.D., Srikanth, A., Weber, M.A. (2016). *Rutgers Graduate School of Education/Education Law Center: School Funding Fairness Data System.* Retrieved from: http://www.schoolfundingfairness.org/data-download

Finally, tracking NAEP scores for low income children, adjusted for differences in income between Kansas' low income children and low income children in other states, we can see that NAEP scores for Kansas low income children have dropped over time, on average, among states. Where Kansas low income 4th graders were among the highest scorers in grade 4 math by 2005, they are now slightly below average. Similarly, Grade 4 Reading has dropped precipitously to below average. Much of the drop has occurred on both tests since 2010. The same is true for Grade 8 math, but Grade 8 reading has jumped around a bit.

Putting it all together, if spending has decline, wages have become less competitive and outcomes have dropped since the last time the state endeavored to estimate how much more it would cost to provide an adequate education, it stands to reason that the additional costs of achieving adequacy now will be greater, if not much greater than in the past.

Figure 14



Baker, B.D., Srikanth, A., Weber, M.A. (2016). *Rutgers Graduate School of Education/Education Law Center: School Funding Fairness Data System.* Retrieved from: http://www.schoolfundingfairness.org/data-download

Table 4. Cost Model Results[a]

Variables	Coefficients	P-value[d]
Intercept	-6.84027	0.19
Performance measure[b]	0.83013	0.00
Cost variables:		
Teacher salaries[b]	1.01765	0.02
Percent free lunch students	0.00636	0.00
Free lunch multiplied by pupil density	0.00065	0.06
Adjusted percent bilingual headcount[c]	0.00139	0.05
Enrollment categories:		
100 to 150 students	-0.12987	0.05
150 to 300 students	-0.29443	0.00
300 to 500 students	-0.38580	0.00
500 to 750 students	-0.44523	0.00
750 to 1,000 students	-0.45612	0.00
1,000 to 1,700 students	-0.52671	0.00
1,700 to 2,500 students	-0.57252	0.00
2,500 to 5,000 students	-0.56802	0.00
5,000 students and above	-0.55366	0.00
Efficiency-related variables:		
Consolidated districts	0.14780	0.00
Per pupil income[b]	0.13097	0.00
Per pupil property values[b]	0.05341	0.02
Total aid/income ratio	0.80593	0.00
Local tax share[b]	-0.02102	0.40
Percent of adults that are college educated (2000)	-0.00666	0.00
Percent of population 65 or older (2000)	-0.00347	0.02
Percent of housing units that are owner occupied (2000)	-0.00218	0.07
Year indicator variables:		
2001	-0.02209	0.31
2002	-0.01666	0.62
2003	-0.08637	0.14
2004	-0.13924	0.09
Adjusted R-square	0.4868	
Sample Size	1468	

[a]Estimated with linear 2SLS with the log of per pupil base spending as the dependent variable. Performance and teacher salaries are treated as endogenous with instruments based on variables for adjacent counties. See Appendix D for methodology. Data is for 1999-2000 to 2003-04.

[b]Measured as natural logarithm.

[c]Calculated by first regressing the share of bilingual headcount from KSDE on the Census measure of poor English (with no intercept). The predicted value from this regression is used as the estimate of the share of bilingual headcount, except in those districts where the share of bilingual headcount is greater than zero. See text for more details.

[d]Probability of being wrong if the hypothesis that the coefficient is equal to zero is rejected. P-values are based on robust standard errors, which correct for heteroskedasticity.

Appendix B: Use of Cost Studies to Inform State School Finance Policies

		NEW JERSEY	PENNSYLVANIA	KANSAS
CONTEXT		Achieve dismissal of long-running judicial oversight.		Comply with court-mandate (and achieve dismissal).
POLICY OBJECTIVE		Eliminate "Abbott" classification & achieve unified statewide formula (and spread aid across more districts).	Achieve unified, more equitable and adequate formula.	
ANALYSES				
	Cost Studies	Augenblick adapted by New Jersey Department of Education (2006)[1]	Augenblick, Palaich and Associates (2007)[2]	Augenblick and Myers (2002)[3] and Kansas Legislative Division of Post Audit (with William Duncombe, Syracuse University) (2006)[4]
	Methods	Successful Schools and Professional Judgment	Successful Schools and Professional Judgment	Augenblick and Myers - Successful Schools and Professional Judgment, LDPA – Education Cost Function and Evidence-Based
	Methodological Notes	NJDOE proposed initial resource configurations. Panels provided opportunity to adjust.[5] NJDOE produced summary report (three years after study completed).	Professional Judgment estimates based on achieving 100 percent proficiency in 2014. Included separate Philadelphia panel.[2]	Hired consultants (Duncombe & Yinger) explored interrelationship between poverty & population density finding significant cost effect.[6]
TRANSLATION TO LEGISLATION				
	Base Figure	Adopted $9,649 for 2009. Cost Study yielded $8,016 (Professional Judgment) to $8,493 (Successful Schools) in 2005.[7]	Adopted $8,355 for 2008-09. Cost Study yielded $8,003 (Professional Judgment) in 2006.[8]	Adopted $4,257 for 2007. Cost Function minimum estimate was $4,565 for 2007. General fund budget only.[9]
	Other Base Adjustments	Added grade level weighting. (Study included cost differences by grade range served).		Backed out federal funding and focused exclusively on "General Fund" expenses.
	Wage Adjustment	Estimated county level "comparable wage" adjustment (claiming NCES ECWI as precedent). Drives funds to high income counties.[10]	Location Cost Metric (largely based on Cost Study).[2,8]	Adopted special adjustment for 16 districts with highest housing prices. Provided additional taxing authority for wealthiest districts.[10]
	Economies of Scale Adjustment	None	District Size Supplement[8]	Carryover of prior legislation.[9]

Student Need Factors	Adopted sliding scale poverty concentration factor (from 47 to 57 percent) and constant ELL weight at 50 percent. Significantly reduced need weight by creating "combination" weight for children who are both low income and ELL (on basis of "redundant services").[5]	Adopted 43 percent low-income weight ($3,593/$8,355). Adopted variable ELL multiplier, which varied with district enrollment and ranged from 1.5 to 2.5 (smaller weight in larger districts, based largely on APA study).[2]	Adopted high density poverty weight (applied to select locations). Drives resources to high need, more "urban" districts. Also adopted non-proficient non-low income weight (not in study). Drives money to generally lower need suburban districts.[9]

[1] DUPREE, A., AUGENBLICK, J., SILVERSTEIN, J. (2006) REPORT ON THE COST OF EDUCATION (RCE) HTTP://NJ.GOV/EDUCATION/SFF/ARCHIVE/REPORT.PDF

[2] AUGENBLICK, PALAICH & ASSOCIATES (2007) COSTING OUT THE RESOURCES NEEDED TO MEET PENNSYLVANIA'S PUBLIC EDUCATION GOALS. PENNSYLVANIA STATE BOARD OF EDUCATION. HTTP://WWW.APACONSULTING.NET/UPLOADS/REPORTS/6.PDF

[3] AUGENBLICK, J., MYERS, J., SILVERSTEIN, J., BARKAS, A. (2002) CALCULATION OF THE COST OF A SUITABLE EDUCATION IN KANSAS IN 2000-2001 USING TWO DIFFERENT ANALYTIC APPROACHES. HTTP://SKYWAYS.LIB.KS.US/KSLEG/KLRD/PUBLICATIONS/SCHOOLFINANCEFINALREPORT.PDF

[4] KANSAS LEGISLATIVE DIVISION OF POST AUDIT (2006) COST STUDY ANALYSIS. ELEMENTARY AND SECONDARY EDUCATION IN KANSAS: ESTIMATING THE COSTS OF K-12 EDUCATION USING TWO APPROACHES HTTP://SKYWAYS.LIB.KS.US/KANSAS/KSLEG/KLRD/PUBLICATIONS/EDUCATION_COST_STUDY/COST_STUDY_REPORT.PDF. SEPARATE STUDY BY WILLIAM DUNCOMBE & JOHN YINGER (SYRACUSE, U.) EMBEDDED IN APPENDIX C OF THAT REPORT.

[5] BAKER, B.D. (2009C) EVALUATING THE "CONCRETE LINK" BETWEEN PROFESSIONAL JUDGMENT ANALYSIS, NEW JERSEY'S SCHOOL FINANCE REFORM ACT AND THE COSTS OF MEETING STATE STANDARDS IN ABBOTT DISTRICTS. EDUCATION LAW CENTER OF NEW JERSEY. HTTP://SCHOOLFINANCE101.FILES.WORDPRESS.COM/2011/10/BAKER-PJP-SFRA-REPORT-WEB.PDF.

[6] DUNCOMBE KS REPORT. SEE ALSO BAKER, B. D. (2011)

[7] NEW JERSEY DEPARTMENT OF EDUCATION. A FORMULA FOR SUCCESS: ALL CHILDREN, ALL COMMUNITIES. HTTP://NJ.GOV/EDUCATION/SFF/REPORTS/ALLCHILDRENALLCOMMUNITIES.PDF

[8] BASIC EDUCATION FUNDING WORKSHEETS: HTTP://WWW.PORTAL.STATE.PA.US/PORTAL/HTTP;//WWW.PORTAL.STATE.PA.US;80/PORTAL/SERVER.PT/GATEWAY/PTARGS_0_123706_1342399_0_0_18/FINANCES%20BEF%202008-09%20MAY2013.XLSX

[9] BAKER, B.D. (2011B) STILL WIDE OF ANY REASONABLE MARK: A REEXAMINATION OF KANSAS SCHOOL FINANCE. SCHOOLS FOR FAIR FUNDING. HTTP://WWW.ROBBLAW.COM/PDFS/P384.PDF (PAGES 65-69)

[10] BAKER, B. D. (2008). DOING MORE HARM THAN GOOD? A COMMENTARY ON THE POLITICS OF COST ADJUSTMENTS FOR WAGE VARIATION IN STATE SCHOOL FINANCE FORMULAS. JOURNAL OF EDUCATION FINANCE, 406-440.