Improving Efficiency in School Budgeting and Spending

Bo Yan, Jefferson County Public Schools
Matthew Linick, Cleveland Metropolitan School District

Executive Summary

Facing the challenge of limited resources, districts want to know how to effectively and equitably use financial resources to best meet the needs of their students. Two agencies, Jefferson County Public Schools and Cleveland Metropolitan School District, are leading the charge to do so. Strategic Data Project (SDP) Fellows from these agencies asked: How can districts improve the efficiency of school budgeting and spending by linking financial, school operations, and student achievement data?

The SDP Fellows in Cleveland Metropolitan School District evaluated the effectiveness of individual academic programs for achieving accountability goals. In the absence of a randomized controlled trial, the SDP Fellow matched students who used a program with similar students who did not, and compared the difference in outcomes. As a result, Cleveland's principals are provided with a report card for each program evaluated, so they can make data-driven decisions that result in the selection of programs that should improve outcomes for kids.

The SDP Fellow in Jefferson County Public Schools created a new financial tool called cycle-based budgeting that ties budget requests to measurable goals, and includes a timeline for program evaluation. Jefferson County's use of the SDP Fellow's budgeting model allows the district to track spending on programs and test their alignment with district priorities.

Strategic Data Project Fellowship Capstone Reports

Strategic Data Project (SDP) Fellows author capstone reports to reflect the work that they led in their education agencies during the two-year program. The reports demonstrate both the impact fellows make and the role of the SDP network in supporting their growth as data strategists. Additionally, they provide recommendations to their host agency and may serve as guides to other agencies, future fellows and researchers seeking to do similar work. *The views or opinions expressed in this report are those of the authors and do not necessarily reflect the views or position of the Center for Education Policy Research at Harvard University.*

Introduction

Local education agencies (LEAs) need a better understanding of the link between financial data, school operation data, and student achievement data in order to determine how can be allocated and utilized in an effective and equitable way. LEAs also need tools to share these findings with district and school administrators to facilitate budget decision making. Jefferson County Public Schools and Cleveland Metropolitan School District came together to address two challenges related to budgeting than can improve decision-making: (a) the effectiveness of school district academic resources and (b) how the budgeting process can be used to foster continuous improvement.

The first case study focuses on the Cleveland Metropolitan School District, which recognized that school and district leaders need information on the effectiveness of programs in their LEAs, and they need this information in an easily consumable format. The second case study looks at Jefferson County Schools, which sought to connect student achievement data with financial data through a new budgeting model and use it to set expectations and improve accountability.

Cleveland Metropolitan School District:

Determining the Effectiveness of Academic Resources

The mission of the Cleveland Metropolitan School District (CMSD) is comprehensive: "The CMSD envisions 21st Century Schools of Choice where students will be challenged with a rigorous curriculum that considers the individual learning styles, program preferences and academic capabilities of each student, while utilizing the highest quality professional educators, administrators and support staff available" (CMSD, 2017). Meeting this goal is not always easy,

as the National Center for Children in Poverty (2017) reported that 53.9% of Cleveland's children live in poverty, giving it the second highest childhood poverty rate in the country.

Facing challenges associated with high levels of childhood poverty, budgetary challenges for the CMSD, and academic performance concerns across the city, members of Cleveland's business community, the mayor, the CMSD administration, Governor Kasich, local philanthropies, and many others banded together to design and launch the Cleveland Plan and pass House Bill 525 (see CMSD, 2012; Jackson, 2012). The citizens of Cleveland also passed a significant levy providing the district with greater financial capacity and an ability to strategically invest additional resources. Part of the strategy to achieve the goals of the Cleveland Plan involves transferring authority and resources to schools and investing in high-leverage system reforms. The primary questions that guided this case study are:

1. How effective are the academic resources used throughout CMSD at impacting performance on the metrics by which CMSD is held accountable?

2. How can CMSD provide information on program quality to school leaders in a manner that will allow them to make well-informed, quality decisions about how to invest their resources?

Literature Review

As Cook (2004) noted, "Ultimately, every evaluation is concerned with describing 'what works,' with identifying program factors that can be deliberately varied to bring about change in a valued outcome" (p. 88). Quasi-experimental designs allow researchers to make limited causal inferences provided certain assumptions are met (Murnane & Willett, 2011; Shadish & Cook, 2009). These methods allow the researcher to make more plausible inferences than if he or she examined trends only using descriptive data (Murnane & Willett, 2011). The ability to make

causal inferences can help those involved in education understand the effects of educational interventions, policies, and programs.

Shadish, Cook, and Campbell (2002) outlined the requirements for making a causal inference established by philosopher John Stuart Mills: cause must precede effect, systematic changes in a cause must correspond to a systematic change in the effect, and the researcher must be able to eliminate all other potential explanations for the effect aside from the investigated cause. Murnane and Willet (2011) asserted that a well-executed randomized experiment fulfills all of requirements set forth by Mills and allows social scientists to make causal inferences and pursue research.

Field experimentation in the social sciences first became popular in the 1960s; such studies included the New Jersey Negative Income Tax Experiment, The National Institute for Mental Health Collaborative Depression Project, the Head Start and Follow Through evaluations, and the Manhattan Bail Bond Experiment (Shadish & Cook, 2009). Experiments allow researchers to make causal inferences about the effects of a given treatment on a unit or research interest; however, it is important to note that the measured effects of a given treatment are always in reference to another treatment, usually a control, or absence of the measured treatment (Holland, 1986). Randomized controlled trials (RCTs) are used to approximate the effects of a treatment on a unit compared to a control. It is impossible to measure the actual effects of a treatment compared to a control, however, because a unit cannot be both in a state of treatment and a state of absence of treatment at the same time—unless the unit is Schrödinger's cat.[1]

---

[1] Schrödinger's Cat refers to a thought experiment devised by Austrian physicist Erwin Schrödinger in 1935.

The observation of what would have happened had the unit not received the treatment is referred to as the counterfactual; through a RCT, a researcher can reasonably approximate the counterfactual (Shadish, Cook, & Campbell, 2002). If a researcher is able to effectively approximate the counterfactual, estimating the effect and significance of the treatment is a matter of simple statistics. The effect of the treatment, after subtracting the effect of the counterfactual, is referred to as the individual treatment effect (ITE). The average ITE is referred to as the average treatment effect (ATE) for all participants. A simple t-test can then inform the researcher if the ATE is significant (Murnane & Willet, 2011).

The Tennessee STAR Experiment, perhaps the most well-known field experiment in education, demonstrated the benefits of reduced class size, especially for minority students (Finn & Achilles, 1990). In recent years, there has been an increased focus on the use of experimental methods to make causal inferences about educational programs. Evidence of this new focus can be seen in the Department of Education's creation of the Institute of Education Sciences (IES), and IES's preferential awarding of funding to experimental and quasi-experimental studies (Shadish & Cook, 2009).

Many times, perfectly implementing an RCT in a highly politicized, public, and personal field like education can be a very difficult task, especially because it is difficult to force a person to accept treatment or control when assigned to either group. There are many ways that researchers compensate for these difficulties, such as estimating the intent to treat (ITT), in which the researcher analyzes outcomes based on units' original assignments to treatment or control groups regardless of units' actual treatment statuses. Another possible method examines the effect of the treatment on the treated (TOT), wherein the researcher examines the effect of

the treatment on those who self-select into the treatment group as compared to those who self-select into the control group (Shadish & Cook, 2009).

While experimental designs reasonably approximate the effect on the theoretical counterfactual, quasi-experiments reasonably approximate the effects of an experiment, which is important in the social sciences where correctly executing a well-designed RCT may not be possible. Quasi-experimental methods provide an opportunity for researchers to approximate the causal inferences seen in a well-executed RCT (Linick, 2012; Murnane & Willet, 2011; Shadish & Cook, 2009).

Project Scope and Methods

The overarching goal of this project was to play a role in ensuring that district resources are used as effectively and efficiently as possible. In order to accomplish this goal, we needed to identify the programs that would be evaluated, conduct quasi-experimental evaluations (where appropriate), collect other program information, design a format by which to share the evaluation findings, and create a vehicle through which to share this information with principals and other stakeholders. Ultimately, this required cross-departmental efforts and several iterations in order to create the final product.

The quasi-experimental evaluations conducted in this case study relied on propensity score matching (PSM). PSM is a quasi-experimental design that seeks to reduce bias in naïve analyses by matching treatment students to statistically identical comparison students (Rosenbaum & Rubin, 1983). PSM generates a propensity score between 0.0 and 1.0 that predicts a student's probability of participating in a program. Taylor (2013) described several matching options when conducting the matching technique including exactly or to nearest neighbor, matching one treatment student to one comparison student or many comparison

students, and replacing matched comparison students into the matching pool or only allowing students to be matched one time. In this evaluation, samples were constructed using one-to-one matching, without replacement, and nearest neighbor matching with a caliper.[2] Students were matched on a variety of measures.[3]

We used PSM to select a sample of nonparticipating comparison students with characteristics similar to students who participated in specific programs and determine the effectiveness of those programs for students who used the program relative to similar students who did not. CMSD evaluators included information about student demographics and prior achievement to match treatment students to statistically similar students. Only students with pretest, posttest, and demographic data were included. In addition, only students found to be acceptable matches were retained for the matched samples. In order to account for the within-school variation of outcomes, we utilized hierarchical linear modeling (HLM) to cluster student-level data within each school and include school-level data to account for between school differences within the district.

Eventually, principals will be provided with a "report card" for each program in the approved vendor directory.[4] For this pilot project, 10 programs will have report cards that include principal comments from surveys as well as information on outcomes of interest and an overall grade based on the grades of each supporting element. Programs are graded on:

---

[2] In "one-to-one matching," each treatment student is matched to one comparison student. "Without replacement" means that once a comparison student was matched to a treatment student, that comparison student was removed from the pool of potential matches. "Nearest neighbor matching with a caliper" means that each treatment student was matched to a comparison student with the nearest propensity score. The use of a caliper prevents treatment students from being matched to nearest neighbors that fall outside the restricted caliper. Guidance suggests that the maximum caliper be calculated as one-quarter of the standard deviation of the treatment variable: $Caliper = SD_{Treatment} * 0.25$.

[3] Pre-test score, race, gender, gifted status, special education status, limited English proficiency status, grade, and age were included in the matching algorithm.

[4] This is a compilation of vendors that have gone through the RFP process with the academic resources team. Principals are not required to pick a program from this list, though it is strongly encouraged.

- **CMSD evaluation (minimal exposure).** This grade is based on the performance of participating students relative to their matched peers. All students identified as having met the minimal threshold for participating in the program (typically at least one session) are included in the treatment group.

- **CMSD evaluation (above average exposure).** This grade is based on the performance of participating students relative to their matched peers. All students identified as having above average "dosage" in the program are included in the treatment group.

- **External evaluation.** This grade is based on external evaluations. It is based on the nature of the findings and the rigor of the methods used in the evaluation.

- **Principal feedback (implementation).** This grade is based on principal feedback provided through surveys about the quality of the implementation of the program at individual schools.

- **Principal feedback (program performance).** This grade is based on principal feedback provided through surveys about the outcomes, value, and deliverables associated with the program at their schools.

Timeline and Current Results

This project has been underway for over a year. The project began with questions from academic leadership about the effectiveness of CMSD's academic resources and has evolved into a cross-departmental effort to support and structure this work. This work has intersected with a redesign of the RFP process, the development of an online vendor directory, and stronger implementation monitoring from the academic resources office. Currently, the work associated with this project could be further impacted through partnership with the information technology department as the district moves to a single-sign-in environment. This will greatly facilitate the

collection of dosage data associated with online interventions. The program evaluation timeline has evolved in five phases over the last year:

- **Phase 1 (complete).** School data at-a-glance reports were provided to vendors and school leaders that showed the school-level data trends at each school over the prior years. This was not done to link program participation to results, but to begin discussion around data trends.

- **Phase 2 (complete).** Rigorous evaluations were conducted to identify the impact of programs on participating students. School and district leaders received reports on these findings, as well as one-page summary documents.

- **Phase 3 (complete).** Surveys were developed and findings collected from externally conducted evaluations to provide principals with multiple data points when considering how to allocate their resources.

- **Phase 4 (in progress).** An online tool is being developed to provide principals with easily understood, important information about the effectiveness of available programs. The first version has been created and will be updated and improved iteratively. The current version of the tool can be found at http://clevelandmetroschoolsvendordirectory.org.

- **Phase 5 (in planning).** Program usage data will be linked to existing data systems to create dashboards for school leaders, district leaders, and vendor partners to monitor implementation, usage, and attached metrics.

Lessons Learned in Cleveland

I have had the privilege to learn many lessons through my engagement in the SDP Fellowship and this project. I have had the freedom to direct this project and the needed guidance

to adjust it to be a useful and valuable tool for school and district leaders. I have received feedback from multiple experts, colleagues, and stakeholders. Of the many lessons that I have learned, two will guide much of my future planning and work in this space.

First, I learned the importance of collaboration and cross-departmental work. While I initially muscled my way through many of the challenges and obstacles, it was not until I developed an authentic partnership with a colleague in the academic resources department that a real vision for this work was created and progress was made. Without input from multiple teams, and buy-in from across the district department, this work would not have progressed with nearly the speed or value that it has.

Second, I learned the importance of flexibility. Throughout this process, I have been forced to iterate and adjust plans for the work. Once I came to appreciate the opportunity to build on or adjust existing plans, the overwhelming nature of such a large project became less problematic. Through flexibility and collaboration, I have contributed to a much stronger product than I initially envisioned or would have been able to produce on my own.

Jefferson County Public Schools:

Using the Budgeting Process to Foster Continuous Improvement

An organization's budget should mirror its strategic vision and goals by funding programs that most support these priorities and meet the greatest needs. Accordingly, an effective budgeting process should be engaged with both allocation of new spending and reallocation of existing spending to meet those goals. Those are the underlying premises of this case study, which took place in Jefferson County Public Schools.

Literature Review

Generally, an organization's budget can be categorized into two parts. One part is *norm-based spending*, which is usually determined by a set of rules based on student enrollment. For example, class size caps at different grade levels dictate how many teachers a school can employ based on student enrollment and, in some districts, school size determines how many counselors, specialists, custodians, or principals a building can have. In contrast, districts use *flexible spending* to run various programs and launch initiatives. While norm-based spending constitutes the bulk of a district's annual budget, budget discussions and decision making often center on the smaller-portion flexible spending that allows districts to implement strategic priorities and execute improvement plans. How to help a district make the best use of this part of its annual budget is the focus of this project.

In the K–12 setting, incremental budgeting and zero-based budgeting are probably the two most widely adopted models. With incremental budgeting, the budget used for the current fiscal year becomes the base for incremental increases or decreases for the next fiscal year. For zero-based budgeting, budget development starts with nothing in terms of budgeted dollars and every spending item needs to be justified for approval.

Budget decisions should be aligned with strategic priorities and tied to outcomes. Existing programs that are closely aligned with the district strategic plan and have been proven to be effective should be continued or expanded with funding support; ineffective programs or those that do not focus on the district's priorities should be altered or discontinued, with the savings re-allocated. Spending on new programs and initiatives should be justified by alignment, evidence of effectiveness, and well-developed implementation plans. Budget decisions serve as a compass, directing a district's attention and improvement effort on areas of greatest needs, and also as a force that drives people to be more oriented to cost impact by spending public money

efficiently and effectively. Unfortunately, both incremental budgeting and zero-based budgeting are limited in achieving these two goals.

The problem with incremental budgeting is obvious for the disconnection between budget decisions and outcomes. Under this model, a program automatically becomes permanent once it is approved. With this entitlement status, people expect to receive more or less what they received in the previous year and largely spend it in the way in which it has been spent—often regardless of how it has improved teaching and learning. This creates an environment for people to become complacent since they are neither motivated nor pressured to continuously improve. Since incremental budgeting does not provide a process through which existing programs can be reviewed for adjustments or discontinuation, this budgeting model tends to focus on the allocation of new spending. The re-allocation of existing spending is largely not exercised except when there is a budget crisis.

The key advantage of zero-based budgeting is the inherent annual process that forces people to justify both existing and new spending according to certain criteria. This can also be used as an opportunity to reflect on program implementation for improvement. It is, however, rather tedious and time-consuming to implement a real zero-based budgeting process, especially for large districts with tens of thousands of budget items. There are also situations where most of the budget items are continuously approved each year, producing a result equivalent to that from an incremental budgeting model. In such cases, the zero-based budgeting process is in name only.

The annual review and decision cycle, which is the fundamental strength of zero-based budgeting, can also be a disadvantage. This is especially true in education where many programs take at least one year to get fully implemented and more time to produce an effect. Often times, a

decrease in student achievement might be observed at the end of the first implementation year of an effective program, as students and teachers are still adapting to the new method or strategy. With the annual decision cycle, such programs may not survive first-year scrutiny even though they will produce positive results in the long run.

The Development of Cycle-Based Budgeting

To overcome the limitations of the incremental budgeting and zero-based budgeting models, we developed a new budgeting model called cycle-based budgeting, which can be understood as a combination of grant application and zero-based budgeting. Similar to a grant application, each new program or initiative must submit a budget request using an online application form. In the application, apart from filling out budget items and amounts, the submitter needs to specify measurable goals such increasing the percentage of students meeting the state math standard by 3%, or decreasing discipline referrals by 5%, as well as the number of years they need to accomplish those goals. As an adaptation of zero-based budgeting, cycle-based budgeting seeks to be more flexible by extending the program review and budget decision cycle from annually to multiple years. Submitters propose the number of years needed to reach measurable goals they set on the application form, but the district's senior leadership team makes the final decision, which can either shorten or lengthen the cycle.

During the cycle of a program, implementation and performance data are monitored and reviewed by program staff for adjustment purposes, but not by the district's senior leadership team for budget decisions. That is, the program's funding is secure unless things go terribly wrong. At the end of the cycle, whether the program accomplishes the measurable goals set at the beginning and whether it continues to align with the district's priorities will be reviewed with budget consequences.

The cycle-based budgeting model enables a district to look at its budget from a different yet badly needed angle. Budget discussions and decisions on flexible spending usually revolve around programs. Once a program is approved, total spending is routinely broken down into budget items and recorded using a corresponding accounting code. This accounting system allows a district to track spending and answer questions such as how much money is spent on salaries, benefits, supplies, and services, or to examine budget allocation by district department and school building. However, it fails to link a district's spending to its focus areas and improvement strategies. It is very difficult if not entirely impossible to find out where money is spent in terms of those focus areas and whether spending on improvement strategies makes a difference.

Through the online application form, cycle-based budgeting fills the missing link by enabling a district to track spending around a district's strategic planning and execution, and answer questions such as how much money has been spent on literacy, math intervention, or discipline, or on certain subgroups. Most important, it allows a district to examine the academic return on investment by looking at whether spending has led to intended outcomes.

Altogether, cycle-based budgeting helps create the time and space necessary for a program or initiative to be thoroughly planned, carefully implemented, closely monitored, and regularly reviewed. Clear expectations are set at the beginning and accountability is demanded at the end of the program. As a result, limited financial resources can be re-allocated depending on program implementation and impact, thus making the flexible spending truly flexible.

Lessons Learned in Jefferson County

One of the key reasons for the success of this project is that it was built on existing infrastructure instead of starting something brand new. Before this project, Jefferson County

Public Schools already had a process through which department heads and building administrators needed to submit budget requests for approval. This project improved that process by adding the additional components necessary to implement a cycle-based budgeting model.

It was clear from the work we did that the tools developed to support the cycled-based process must be user-friendly and easy to use. From a cost-benefit perspective, the electronic budget request form is easier to fill out than the paper form used in the past; a declined budget request can be easily modified and resubmitted the following year; and the finance team can monitor and track budget requests in real time. Instead of introducing additional work, the tools made things easier for the stakeholders, which led to a smooth transition.

Perhaps most important, leadership was crucial to the success of this project. Before the project started, both the director of planning and program evaluation and the director of finance wanted to improve the budget request and approval process they had put in place. This project was well aligned with their vision at the outset. As such, the work received continuous attention and support from the leaders, who helped communicate its purposes and goals to the district's other top leaders, allowing quick resolution of any issues that arose.

References

Cleveland Metropolitan School District. (2012, June 15). *Substitute House Bill 525: Summary of provisions.* Retrieved from http://www.clevelandmetroschools.org/cms/lib05/OH01915844/Centricity/Domain/98/SummaryHB525Provisions.pdf

Cleveland Metropolitan School District. (2017). *About us: Mission/vision*. Retrieved April 12, 2017, from http://www.clevelandmetroschools.org/Page/747

Cook, T. D. (2004). Causal generalization. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 88–113). Thousand Oaks, CA: Sage.

Finn, J. D., & Achilles, C. M. (1990). Answers and questions about class size: A statewide experiment. *American Educational Research Journal, 27*(3), 557–577.

Holland, P. (1986). Statistics and causal inference. (with discussions) Journal of the American Statistical Association, Vol. 81: 945–960.

Jackson, F. G. (2012). *Cleveland's plan for transforming schools: Executive summary.* Cleveland, OH: Office of the Mayor. Retrieved from http://www.clevelandmetroschools.org/cms/lib05/OH01915844/Centricity/Domain/4/ClevelandPlanExecutiveSummary.pdf

Levenson, N. (2012). *Smarter budgets, smarter schools: How to survive and thrive in tight times*. Cambridge, MA: Harvard Education Press.

Linick, M. A. (2012). *Preliminary dissertation defense.* Unpublished manuscript.

Linick, M. A. (2013). *Examining charter school policy and public school district resource allocation using multiple quasi-experimental designs* (Doctoral dissertation). Retrieved from ProQuest.

Murnane, R., & Willet, J. (2011). *Methods matter: Improving causal inference in educational and social science research*. New York, NY: Oxford University Press.

National Center for Children in Poverty. (2017). Researchers: Detroit, Cleveland, Buffalo, Cincinnati, Newark lead nation in child poverty. Retrieved from http://www.nccp.org/media/releases/release_153.html

Odden, A., & Picus, L. (2008). *School finance: A policy perspective (4th ed.)*. Boston, MA.: McGraw-Hill.

Rosenbaum, P. R., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41–55.

Shadish, W., & Cook, T. (2009). The renaissance of field experimentation in evaluating interventions. *Annual Review of Psychology, 60*, 607–629. doi:10.1146/annurev.psych.60.110707.163544

Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth Cengage Learning.

Sorenson, R., & Goldsmith, L. (2006). *The principal's guide to school budgeting*. Thousand Oaks, CA: Corwin.

Taylor, J. (2013). *Community college dual credit: Differential participation and differential impacts on college access and completion* (Doctoral dissertation). Retrieved from ProQuest.

Zockoff, N. (2012). *Smart school budgeting: Resources for districts.* Cambridge, MA: Rennie Center for Education Research & Policy.