MDRC Working Papers on Research Methodology

A Regression-Based Strategy for Defining Subgroups in a Social Experiment

James J. Kemple and Jason C. Snipes

with

Howard Bloom

Manpower Demonstration
Research Corporation

MDRC

June 2001

This is part of a series of MDRC working papers that explore alternative methods of evaluating the implementation and impacts of social programs and policies.

An earlier version of this paper was presented at the 22nd Annual Research Conference of the Association of Public Policy Analysis and Management (November 2000).

This work was supported by the U.S. Department of Education, Office of Educational Research and Improvement and National School-to-Work Office; and The Pew Charitable Trusts. It is based on data from the Career Academies Evaluation, which is funded by:

DeWitt Wallace–Reader's Digest Fund
Ford Foundation
U.S. Department of Education
U.S. Department of Labor
The Commonwealth Fund
Charles Stewart Mott Foundation
William T. Grant Foundation
The Pew Charitable Trusts
The Rockefeller Foundation
The George Gund Foundation
The Grable Foundation
Richard King Mellon Foundation
American Express Foundation
Alcoa Foundation
Russell Sage Foundation
Center for Research on the Education of Students Placed At Risk (CRESPAR)
Westinghouse Foundation
The Citigroup Foundation
Bristol-Myers Squibb Foundation, Inc.

Dissemination of MDRC publications is also supported by the following foundations that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: the Ford, Ewing Marion Kauffman, Ambrose Monell, Alcoa, George Gund, Grable, Starr, Anheuser-Busch, New York Times Company, Heinz Family, and Union Carbide Foundations; and the Open Society Institute.

The authors would like to thank Hans Bos, Charles Michalopoulos, and Winston Lin for their insights and advice on this paper.

For information about MDRC and copies of our publications, see our Web site: www.mdrc.org.

Two prominent themes often emerge from evaluations of education and social program evaluations: (1) the interventions being studied serve diverse populations, even if they are intended to target groups with particular characteristics; and (2) the interventions' impacts vary across groups within the population being served. Thus, most evaluations of education, employment and training, welfare-to-work, and health-related interventions are not only interested in the question "What works?" but in the question "What works for whom?" This leads to an extensive investment in subgroup analysis. This paper describes a "regression-based" strategy for defining subgroups that is more systematic than traditional "accumulation" strategies. It relies on data and methods used in MDRC's Career Academies Evaluation to illustrate the use of this technique and discusses its various advantages and limitations.[1]

I. <u>The Career Academies Evaluation</u>

In 1993, MDRC began an evaluation of the Career Academy approach as it had been defined in previous research and implemented in a broad range of settings across the country. The evaluation's primary goal is to provide policymakers and educators with reliable evidence about the impact that Career Academies have on students' success in high school and their transition to further education and the labor market. The evaluation will also offer lessons about how Career Academies operate and are sustained and about the pathways through which Academies affect student engagement and performance in school.

The Career Academies Evaluation is a rarity in the field of education research in that it has demonstrated the feasibility and benefits of implementing a large-scale, multi-site random assignment research design within an ongoing high school program.[2] This was made possible because each of the Career Academies in the study received applications from approximately twice as many students as it was able to serve. The analyses presented in this paper are based on a sample of 1,764 students who applied for one of the Career Academies selected for the study. Of these, 959 students were randomly assigned to the program group (referred to in this paper as the *Academy group*) and were accepted for admission to the Academies. The remaining 805 students were randomly assigned to a control group (referred to in this paper as the *non-Academy group*) and were not invited to participate in the Academies, although they could choose other options in the high school or school district.

The random assignment process ensured that there were no systematic differences between the two groups of students in terms of their observable and unobservable background characteristics, prior school experiences, and initial motivation and attitudes toward school. Any systematic differences that subsequently emerged between the groups can be attributed with confidence to differences in their access and exposure to the Career Academies.

Students in the study sample were identified at the end of 8th or 9th grade. The primary focus of this paper is on outcomes measured through the end of their scheduled 12th-grade year, just before they would have graduated from high school.

---

[1]See Kemple and Snipes, 2000, for a more detailed discussion of the findings from the evaluation.

[2]For a more detailed description of the random assignment design and how it was implemented, see Kemple and Rock, 1996.

II. Analytic Importance of Subgroup Analysis

In order to understand the impact of Career Academies, it is important to recognize the heterogeneity of the student population and the fact that some groups of students benefited differently than others. When the impact results were averaged across the diverse groups of students the Career Academies served, it appeared that the programs produced only slight reductions in dropout rates, and modest improvements in students' progress toward graduation and increases in participation in youth development activities.[3] These aggregate results masked the high degree of variation in the Career Academies' potential to make a difference and in the actual differences they made for some students. In short, findings that were aggregated across the diverse groups of students served by the Academies did not reveal many of the most important effects that Academies had. Positive effects for some subgroups of students were offset or muted by small or zero impacts for other subgroups.

For example, an important goal of the Career Academies is to reduce dropout rates and increase students' engagement in school. As noted earlier in the evaluation Career Academies serve a broad cross section of students, many of whom enter the programs highly engaged in school.[4] It is unlikely that the programs will have an effect on dropout rates among these students, who are highly unlikely to drop out of school even if they do not attend an Academy. On the other hand, a number of students in the sample who applied for the Academies were relatively disengaged from high school and appeared to be at risk of dropping out of high school. To the extent that the Academies can have an effect on dropout rates, it is likely to be concentrated among these students. The magnitude of this effect could be diluted or even completely hidden if averaged with the lack of impact for the rest of the students in the sample.

In order to assess the effect of the Academies on student outcomes more sensitively, therefore, it was necessary to differentiate among students with different needs and trajectories at the time they entered the programs. The attempt to make distinctions among groups of individuals with different needs and characteristics, who might experience substantially different benefits from an intervention, is not uncommon to experimental research in general or to education research in particular. Toward this end, therefore, subgroups of the Career Academy research sample were defined with three goals in mind: (1) using pre-random assignment characteristics associated with the likelihood of dropping out of high school; (2) maximizing the contrast in outcomes among the subgroups (most importantly among those in the non-Academy group); and (3) maximizing the sample size for each subgroup.

The random assignment research design used in this evaluation provides a unique opportunity to identify subgroups of students who, without access to an Academy, were relatively highly likely to drop out of high school and to compare them with similar students who did have access to an Academy. Not only does such a design provide the unusual opportunity to establish which outcomes would have been observed in the absence of the Academy treatment, but it also provides an opportunity to observe the *relationships* between background characteristics and important outcomes in the absence of the intervention.

[3]See Kemple and Snipes, 2000.
[4]See Kemple and Rock, 1996.

There are several strategies for identifying subgroups. The following section of the paper describes a more traditional approach and highlights several limitations that led to the use of a strategy that provided greater insight into the variation in program effects.

III. Traditional Approach to Defining Subgroups: Risk-Factor Accumulation

One of the strategies most frequently used to define subgroups might be called "risk-factor accumulation." It entails first identifying a list of background characteristics typically associated with an important outcome or with the manner in which the program treatment is likely to be delivered. A critical outcome for many high school interventions, including the Career Academy's approach, is dropping out of high school. A number of education research studies have identified several background characteristics and prior school experiences that are associated with a high likelihood of dropping out of high school.[5] This includes prior school experiences such as poor attendance, low grades, or being held back in a previous grade. It also includes demographic characteristics such as being from a low-income family, having a sibling who dropped out, or having moved and transferred schools several times.

The risk-factor accumulation strategy classifies students into risk subgroups by counting the number of risk factors an individual has, weighting all the factors equally. For example, if one identified six characteristics associated with dropping out, individuals with two or more of these characteristics might be considered to be at "high risk" of dropping out; those with only one of the characteristics might be considered to be at "moderate" or "medium" risk of dropping out; and those with none of the characteristic factors might be considered at "low risk."

This strategy has the appeal of being straightforward in execution, and it can be translated directly into a strategy for targeting students to receive special services. For example, if a particular school intervention were found to prevent students in the high-risk subgroup from dropping out, teachers or administrators might wish to ensure that students with two or more of the risk characteristics be included in that program.

At the same time, the accumulation strategy has several important limitations. First, such an analysis gives equal weight to each of the risk-related background characteristics and prior school experiences examined. As a result, it does not account for the fact that some characteristics are more highly associated with school failure than others. This strategy also does not account for the fact that some characteristics are associated with school success and may offset the risk associated with other characteristics. As a result, it fails to account for the possibility that, *given the same number of risk factors, different combinations of characteristics may indicate different degrees of risk*. In other words, because some characteristics are more strongly associated with academic outcomes than others, students with the same number of characteristics may actually be substantially more likely (or less likely) than one another to drop out of high school. A related limitation of this strategy is that it is based on categorical variables and is therefore unable to take advantage of the more subtle distinctions among students that are captured by continuous variables. For example, a

---

[5]See, for example, Natriello, 1987; National Center for Education Statistics (NCES), 1992; and Roderick, 1993.

“high” absentee rate might be defined as being absent 20 or more days during a school year. Students who were absent 40 days, as well as those absent 20 would be regarded as being equally “at-risk using risk criterion. Finally, the risk factor accumulation strategy may yield very small subgroups if relatively few sample members possess all or most of the highly specified characteristics.

Because it does not allow for a more complex set of relationships between risk factors and student outcomes, the simple risk-factor accumulation strategy may fail to produce subgroups with distinctly different academic trajectories. Therefore, in order to distinguish more effectively among subgroups of students who, in the absence of the program, would have experienced distinctly different outcomes, the Career Academies Evaluation employed an imputation strategy for identifying subgroups. This is referred to throughout this paper as a *regression-based subgroup strategy.*

IV. Regression-Based Subgroup Strategy

A. Overview of the Approach

The regression-based strategy described below involved four steps. The first three steps involved defining a key risk-related behavior (in this, case, dropping out of high school) that the Academies aim to prevent and that is likely to vary among non-Academy group members; identifying background characteristics that are likely to be empirically related to this risk-related behavior; and using multiple regression to generate estimates of the relationship between the background characteristics and the risk measure. The basic idea behind these steps in the regression-based subgroup strategy is to build on the opportunity created by the random assignment experimental design in order to identify the relationships between background characteristics and student failure *in the absence of the Academy intervention*. Based on these relationships, one then identifies the characteristics of the students who, in the absence of the program, are most likely to drop out of high school. Thus, the analyses in steps 2 and 3 were conducted with the non-Academy group in the research sample.[6]

The fourth step in the regression-based strategy used here involved using the estimates generated in the third step to create an index for those in both the Academy and on=-Academy groups that indicated their propensity toward dropping out of high school. The index was then used to divide the sample into subgroups with high, medium, and low risks of dropping out. Following is a brief description of each step.

Step 1: Defining a risk-related behavior. Students’ dropout behavior is a critical outcome for the evaluation. A key goal of Career Academies is to prevent student from dropping out of high school and keep them on target toward graduation. Dropping out is also a critical factor in students’ attainment of other outcomes both during and after their high school years. Thus,

[6]As discussed later in the paper, using regression parameter estimates from the non-Academy group has the potential to overstate positive impacts on the program for those at high-risk of dropping out of high school and understating negative effects for those at low-risk of dropping out. Although this distortion is actually quite small in this case, a preferred strategy for identifying background characteristics and generating regression parameter estimates is to rely on an external sample that is not used in the impact analysis. This strategy was used, for example, in Bloom, Kemple, Morris, Scrivener, Verma, and Hendra, 2000.

dropout status at the end of students' 12^{th}-grade year was identified in the key risk-related behavior for the regression-based subgroup strategy. Students were identified as having dropped out of high school if they were not found to be enrolled in any high school in the participating school district and if either the school districts' administrative records or students' self-reports indicated that the student had not graduated and was not enrolled in school elsewhere. Students who could not be confirmed as having dropped out were assigned a value of missing for this variable.

Step 2: Identifying relevant background characteristics. The second step in the regression-based strategy is to identify background characteristics that are highly correlated with the dropout indicator. For the Career Academies Evaluation subgroup analysis, these characteristics were chosen by considering both conceptual and empirical linkages with the dropout indicator. The list of characteristics is consistent with variables identified in prior research and were found to be empirically related to dropping out among students in the Career Academy sample. The background characteristics included in the Career Academies Evaluation subgroup analysis include:

- average daily attendance in the year the student applied for an Academy;
- grade point average for the year the student applied for an Academy;
- the number of credits earned toward graduation in the year the student applied for an Academy;
- whether the student was overage for grade when entering the Academy;
- whether the student had a sibling who dropped out of high school; and
- whether the student had transferred schools two or more times beyond the typical school transitions.

Step 3: Estimating the empirical relationship between background characteristics and the dropout indicator. This step involved using to use multiple regression to estimate the relationship between several background characteristics measured at the time students applied to the Academy and the probability that they would drop out of high school before the end of the 12^{th} grade.

For the purposes of this study, a key goal of this analysis was to capitalize on the experimental design and estimate the relationships between background characteristics and dropping out of high school in the absence of access to an Academy. The random assignment research design ensures that the non-Academy group provides the best counterfactual for what would have occurred to students in the absence of access to an Academy. Thus, the non-Academy group was used as the basis for this regression. Table 1 presents the results of this regression analysis. The first column of parameter estimates reflects the relationship between the dropout rate and a unit change in the background characteristics. Numbers in the second column are standardized to reflect the relationship between the dropout rate and a *standard deviation* change in the background characteristics. As the table suggests, all the characteristics included in this regression model are statistically significant and are related to the probability that students would drop out of high school before the end of the 12th grade.[7]

Table 1

Career Academies Evaluation

Relationship Between Baseline Characteristics and the Probability of Dropping Out of High School Among Non-Academy Students

	Coefficients	
Baseline Characteristic	Unstandardized	Standardized
Sibling dropped out	0.08 ***	0.03 ***
	(0.03)	(0.01)
Overage for grade	0.06 **	0.02 **
	(0.03)	(0.01)
Transferred schools 2 or more times	0.07 ***	0.03 ***
	(0.03)	(0.01)
Attendance rate in year of random assignment	-0.01 ***	-0.04 ***
	(0.00)	(0.01)
Credits earned in year of random assignment	-0.05 ***	-0.05 ***
	(0.01)	(0.01)
Grade point average in year of random assignment	-0.03 *	-0.02 *
	(0.02)	(0.01)
Intercept	0.94 ***	0.12 ***
	(0.14)	-(0.14)
R-squared	0.10	0.10
Sample size	763	763

SOURCES: MDRC calculations from Student Baseline Questionnaire Database and Student School Records Database.

NOTES: Estimates are regression-adjusted using ordinary least squares, controlling for background characteristics of sample members. Rounding may cause slight discrepancies in calculating differences.
A two-tailed t-test was applied to differences between the Academy and non-Academy groups. In both cases, statistical significance levels are indicated as *** = 1 percent; ** = 5 percent; * = 10 percent.

Step 4: Applying the regression coefficient estimates to create a risk index and dividing the sample into discrete subgroups. This final step has two stages. The first is to combine the coefficients from the regression estimates for the non-Academy sample with the back-

[7]Other specifications of this model were tried. However, through an informal process of model specification, this six-variable model was found to be the most sensible and effective. The estimates (below) of the potential distortion caused by the regression-based approach do not take into account any effects of the model specification process on the impact estimates.

ground characteristics of each individual in both the Academy and the non-Academy groups. In other words, the coefficient estimates from the regression are used as weights multiplied by the relevant measured background characteristics of each individual. The weighted sum of these characteristics yields an index indicating the probability of dropping out of high school. This is referred to as the *risk index,* and it provides a basis for ranking sample members according to the predicted probability that they would drop out of high school.

For example, the parameter estimate associated with having a sibling who dropped out of school is .08 (that is, controlling for other background characteristics, students in the evaluation who had a sibling who already dropped out of high school were predicted to be 8 percentage points more likely to drop out of high school). Therefore, students with siblings who dropped out had .08 added to the index measuring their own risk of dropping out. By the same token, the regression estimates indicate that some characteristics are negatively correlated with dropping out. The weights assigned to these characteristics were multiplied by individual attributes and subtracted from the risk index.

In the second stage, the Academy and non-Academy students were divided into three subgroups based on the risk index. Following is a brief definition of each of the three risk subgroups.

- **The high-risk subgroup:** the students in the Academy and non-Academy groups with the combination of characteristics yielding scores at or above the 75th percentile of scores on the risk index (that is, those with the highest likelihood of dropping out)
- **The low-risk subgroup:** the students in the Academy and non-Academy groups with the combination of characteristics yielding scores at or below the 25th percentile of scores on the risk index (that is, those with the lowest likelihood of dropping out)
- **The medium-risk subgroup:** the remaining students in the Academy and non-Academy groups (approximately 50 percent of the study sample) with a mix of characteristics yielding scores between the 25th and 75th percentile on the risk index (that is, indicating they were not particularly likely to drop out but not necessarily highly engaged in school)[8]

B. Strengths of the Regression-Based Strategy

There are several important advantages to the regression-based strategy for defining subgroups. First, it incorporates factors which are both conceptually and empirically related to students' risk of dropping out of high school. At the same time, because these characteristics were measured prior to students' random assignment to the Academy and non-Academy groups, they are *exogenous* to the Academy treatment. In other words, while the background characteristics

[8]The 25th and 75th percentile cutoffs were based on the distribution of the risk index among the non-Academy students.

used to create the subgroups were correlated with the likelihood of dropping out, these characteristics did not influence the selection of students into the Academy group.

An important question for such an impact analysis is whether, within each subgroup, the random assignment research design is preserved. In other words, are there systematic differences between the background characteristics of the Academy and non-Academy students *within* each subgroup? To test this, a set of background characteristics (including those used to create the risk index) was regressed against a dummy variable indicating whether the student was assigned to the Academy group.[9] The results of this analysis revealed that, while there were differences between the background characteristics of Academy and non-Academy students on one or two individual characteristics within each subgroup, f-tests failed to reject the hypothesis that there are no overall systematic differences between the background characteristics of the Academy and non-Academy students. This suggests that the random assignment research design was preserved within each subgroup. In other words, the existing differences are not greater than those which would be expected to occur by chance.

A second strength of this approach is that it incorporates the fact that the relationships between "risk factors" and student outcomes vary, depending on the background characteristic. For example, the coefficient estimates suggest that the effect of the number of credits earned in the year prior to random assignment and the effect of baseline attendance on the dropout rate are each at least twice as large as the effect of a student's baseline grade point average or whether a student was overage for grade.[10] Basing the subgroup definitions on these relationships allows these differences to be factored into the classification of students into the three risk subgroups. For example, these regression estimates suggest that an average student who had a sibling who had dropped out and who was overage for grade would have approximately a 24 percent chance of dropping out of high school before the end of the 12th grade. However, if that same student also had 98 percent attendance and was about a standard deviation above the average in terms of credits earned, he or she would have only a 16 percent chance of dropping out.[11]

Moreover, the regression-based strategy is capable of incorporating variation across students along continuous variables such as attendance and grade point average. Less flexible strategies that fail to incorporate these factors would not be as effective at distinguishing among students at different levels of academic risk. For example, an otherwise average student with perfect attendance (that is, 100 percent) has a 9 percent chance of dropping out; a similar student with an attendance rate of 95 percent has a 12 percent chance of dropping out; a student with a 90 percent attendance rate has a 15 percent chance of dropping out; and one with 85 percent attendance has a 17 percent chance of dropping out. In other words, there appears to be meaningful variation in the probability of dropping out that would not be captured by a simple categorical

---

[9]For a more detailed review of this analysis, see Appendix A in Kemple and Snipes, 2000.

[10]Note that these coefficients have been standardized to reflect the effect of a standard deviation change in the independent variable on the dropout rate, thus making the coefficient estimates directly comparable with one another.

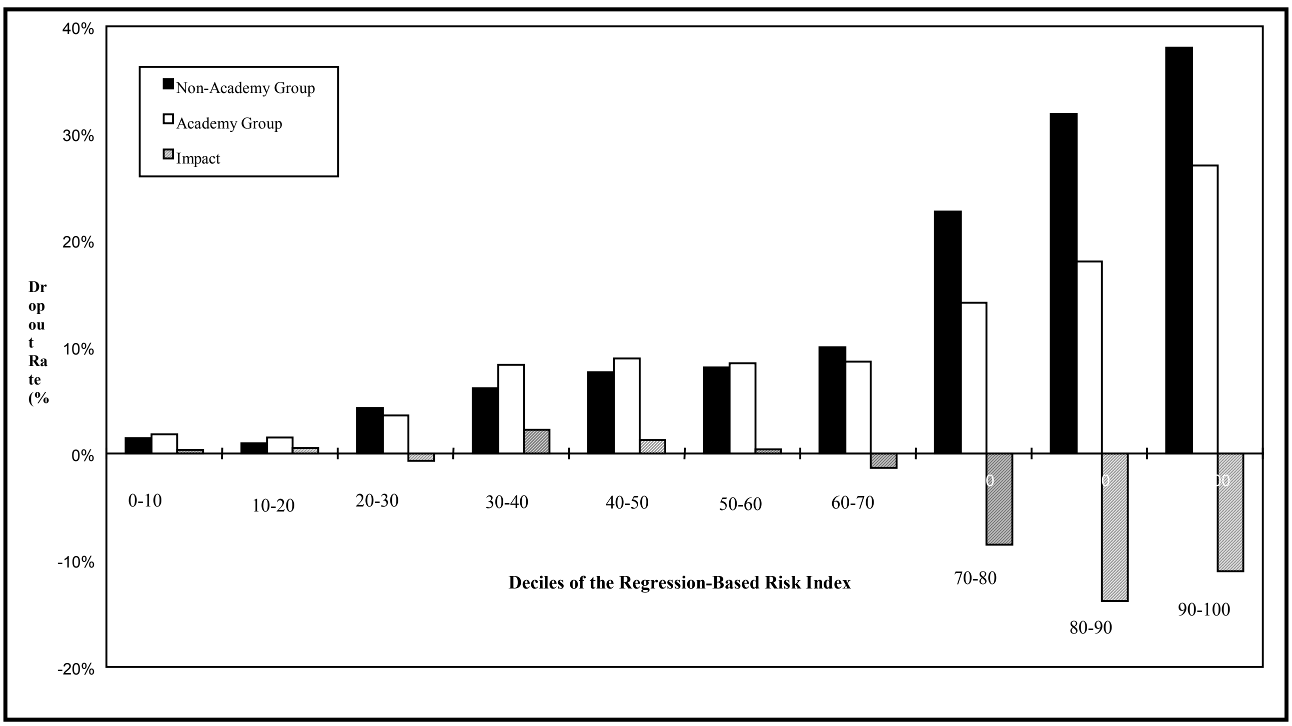
[11]The predicted probability of dropping out for the average student was estimated by multiplying the mean values of the independent variables among the students in the study sample by the coefficients in Table 1. The estimated probabilities for students with the hypothesized characteristics were estimated by substituting the hypothesized values for the mean values where appropriate.

measure of attendance. The regression-based subgroup strategy captures such variation and incorporates it into the assessment of each student's risk of school failure.

The third and perhaps most important strength of the regression-based strategy is that it effectively identifies students with distinct academic trajectories. Figure 1 presents the dropout rates for Academy and non-Academy students, as well as the difference between their dropout rates, at 10 percentile intervals on the regression-based risk index. The black bars represent the percentage of non-Academy students who dropped out of high school, and the white bars represent the percentage of Academy students who did so. The striped bars represent the difference between these two groups, that is, the impact of the Academy treatment on dropout rates. The pattern in this figure suggests that the risk index very effectively differentiates among students with different academic trajectories, and that the relationship between risk and the impact of Academies on dropout rates is not isolated to a small segment of the student population.

Figure 1

Impact of Career Academies on Dropout Rate, by Deciles Regression-Based Risk



SOURCES: MDRC calculations from Career Academies Evaluation Student School Records Database and 12th Grade Survey Database.

The figure indicates that both the risk of dropping out and the impact of the program on this outcome generally increase with the percentiles of the risk index. In particular, the dropout rate among the non-Academy group appears to increase steadily with the percentiles of the risk index, and it grows sharply after the 70th percentile. The impact on the dropout rate follows essentially the same pattern. From the 30th percentile through the 90th, the difference between the Academy and non-Academy groups becomes increasingly negative. The magnitude of this reduction in dropout rates appears to increase dramatically after the 70^{th} percentile, and then it shrinks slightly among students above the 90^{th} percentile of risk. This pattern suggests that, for the individuals with low to moderate risk of dropping out, the impact of the program on dropout rates appears to be rather negligible. However, as the risk of academic failure becomes more serious, the impact of the Academy approach appears to grow. Finally, for those at greatest risk, the impact on dropout rates is substantial, but it is not as great as for those who are slightly less at risk.

In short, this graph illustrates that the regression-based strategy is quite effective at differentiating among students with different dropout rates, and that the impact of the Academies on the dropout rate is strongly related to this definition of risk.

Table 2 illustrates that the regression-based strategy is effective at differentiating among students with different trajectories across a variety of school outcomes, and that it is more effective than the risk-factor accumulation strategy for making these distinctions. The table presents several key measures of student performance during high school for the non-Academy students within each risk subgroup. The first panel of the table presents non-Academy outcome levels and estimated impacts based on the risk-factor accumulation approach, and the second panel presents these estimates based on the regression-based approach to defining subgroups. As the table illustrates, the regression-based strategy does a better job of making distinctions among students with different levels of academic risk.

According to the estimates generated by the regression-based approach, while 32 percent of the non-Academy students in the high-risk subgroup dropped out of high school before the end of the 12^{th} grade, 8 percent in the medium-risk subgroup dropped out, and less than 3 percent in the low-risk subgroup did so. Moreover, while only 27 percent of the non-Academy students in the high-risk subgroup earned enough credits to graduate from high school, 65 percent in the medium-risk subgroup and 75 percent in the low-risk subgroup did so. Similar patterns were found for most other measures as well. This indicates that, without access to a Career Academy, the students in the different risk subgroups would have had substantially different outcomes.

Table 2 also provides outcome levels and estimated impacts for subgroups based on the risk-factor accumulation approach. Not surprisingly, these estimates are not as distinct from one another as the estimates generated by the regression-based approach. For example, 22 percent of students in the "high-risk" subgroup dropped out of high school before the end of the 12^{th} grade, compared with 7 percent in the "medium-risk" subgroup and 5 percent in the "low-risk" subgroup. Moreover, 44 percent in the "high-risk" subgroup earned enough credits to graduate, compared with 66 and 78 percent in the "medium-risk" and "low-risk" subgroups, respectively.

These patterns in outcome levels among students who weren't exposed to the Academy treatment suggest that the regression-based strategy is the more effective means for defining subgroups of Career Academy students with substantially different academic trajectories. Interestingly, the impact estimates suggest that while the estimates generated by the regression-based approach tend to be somewhat larger, their pattern is similar to the pattern of estimates based on the risk-factor accumulation model. For example, among students in the high-risk subgroup, both

Table 2

Career Academies Evaluation

Selected Outcomes Among Non-Academy Students, by Risk Subgroups Defined Using Risk-Factor Accumulation and Regression-Based Index

	Accumulation Approach		Regression-Based Approach	
Outcome	Non-Academy Outcomes (%)	Impact	Non-Academy Outcomes (%)	Impact
High-risk subgroup				
Dropped out of high school	27.4	-5.6 *	32.3	-11.4 ***
Earned credits to graduate	34.1	10.0 **	27.0	12.8 ***
Completed basic academic core	9.9	9.1 *	5.6	8.0 *
Reported any negative risk-taking	36.9	-5.9	38.9	-3.8
Reported positive youth development	56.8	8.7 *	54.9	8.0
Medium-risk subgroup				
Dropped out of high school	9.3	-2.0	7.9	0.9
Earned credits to graduate	66.2	-0.6	64.8	0.8
Completed basic academic core	31.0	-2.7	30.3	-1.2
Reported any negative risk-taking	23.1	0.5	25.7	-2.2
Reported positive youth development	67.5	4.1	69.7	1.6
Low-risk subgroup				
Dropped out of high school	4.2	0.5	2.8	-1.2
Earned credits to graduate	69.5	8.4 **	74.8	12.9 **
Completed basic academic core	33.9	0.9	36.6	4.5
Reported any negative risk-taking	22.0	-2.1	15.8	-1.0
Reported positive youth development	75.4	0.8	75.5	6.3

SOURCES: MDRC calculations from Career Academies Evaluation Student School Records Database and 12th Grade Survey Database.

NOTES: Estimates are regression-adjusted using ordinary least squares, controlling for background characteristics of sample members. Rounding may cause slight discrepancies in calculating

A two-tailed t-test was applied to differences between the Academy and non-Academy groups. In both cases, statistical significance levels are indicated as *** = 1 percent; ** = 5 percent; * = 10 percent.

the regression-based approach and the risk-factor accumulation approach found that Academies significantly reduced dropout rates, increased credits earned toward graduation, and increased the percentage of students who completed a core academic curriculum. So while the regression-based approach was more effective at identifying students who, in the absence of the Academy treatment, would have had substantially different outcomes, it did not distort the basic pattern of impacts generated by the experiment.

C. Potential Limitations of the Regression-Based Approach

While the regression-based strategy is more effective than the risk-factor accumulation strategy at identifying students who were likely to experience different academic trajectories in the absence of the Academy, it has some potentially important limitations. First, although it is more systematic, it is also less straightforward than the risk-accumulation strategy in terms of the manner by which subgroups of students might be identified by school administrators. In particular, to the extent that these subgroup findings might be used to target program resources toward particular individuals, the subgroups defined using the regression-based strategy might be more difficult to identify than subgroups based on a simple accumulation approach. While it is unclear that the implications of the findings from this particular study suggest that targeting would be advantageous, such thinking may be a factor when applying this strategy to the study of programs in which the implications of targeting are less ambiguous.

Although it is not discussed in this paper, the regression-based approach can be applied in a practical way and may, in fact, be a more systematic way of targeting resources toward students most likely to benefit from them. For example, this type of approach has been used in research designed to develop approaches for the targeting of benefits and associated employment services to workers eligible for unemployment insurance as well as for targeting employment resources to individuals in welfare-to-work programs. In particular, several of these programs have used historical data to estimate the relationship between background characteristics and policy-relevant outcomes, and then to combine these estimates with individual characteristics in order to predict outcomes and target services. This has been done in welfare-to-work programs in Michigan as well as in unemployment programs in Michigan, New Jersey, and Washington.[12]

A more important potential limitation of the regression-based subgroup strategy is related to the manner in which the strategy generates weights relating background characteristics to risk. In short, theoretically, the strategy has the potential to overstate any positive impacts of the program on the high-risk subgroup and to overstate the magnitude of any negative impacts on the low- and medium-risk subgroups.

The problem has its genesis in the fact that the regression parameter estimates that are used as weights to translate student characteristics into academic risk are the result of estimates that are specific to the non-Academy group. In a sample from any population, estimated regression coefficients reflect both the relationships that exist in the population and a random element that is specific to that sample. In other words, on average, each coefficient from such a random

[12]See O'Leary, Decker, and Wandner, 1998; and Eberts, 1997.

sample is unbiased. However, it is highly unlikely that, in any given sample, the estimated regression coefficient will exactly equal the true regression coefficient from the entire population from which that sample is drawn. Therefore, the regression estimates from the non-Academy group include some random error that is particular to the non-Academy group and that is correlated with the outcome in question — in this case, whether or not a student dropped out before the end of the 12th grade.

For example, Equation 1 is a simple regression predicting dropout from a set of background characteristics for a sample of students drawn from the population of students who applied to a Career Academy:

$$Y_i = \hat{\alpha} + \hat{\beta} X_i + e_i \quad (1)$$

where:

Y_i = 1 if student *i* dropped out; 0 otherwise;

X_i = 1 if student i had ever been held back; 0 otherwise (this could be any important background characteristic);

$\hat{\alpha}$ = the intercept term, that is, the average outcome (Y_i) among those where $X = 0$; and

$\hat{\beta}$ = the estimated relationship between X_i and Y_i, that is, the estimated effect of X_i on the probability that a student drops out of high school.

In this case, it would also be true that:

$$\hat{\beta} = \beta + \beta_s \quad (2)$$

where:

β = the true relationship between *X* and *Y* in the population from which our sample was drawn; and

β_s = the difference (or error) between the relationship between X and Y in the population from which the sample was drawn and the relationship between *X* and *Y* in the sample, that is, the element of the estimate which is idiosyncratic to the particular sample.

While β is a characteristic of the population and does not change from sample to sample, β_s is particular to the sample upon which the regression is estimated, and it *will* vary from sample to sample. As a result, while β never changes, $\hat{\beta}$ will vary from sample to sample. Furthermore, it is also highly unlikely that the random error in a coefficient estimated from one sample drawn from a population will be exactly the same as the random element in any other sample drawn from the same population.

The students in the Career Academies evaluation sample were assigned to the Academy or non-Academy groups at random; therefore, one can have a high degree of confidence that there are no *systematic* differences between these two groups in terms of observable or unobservable characteristics. They can be thought of as two random samples drawn from the same population of students at these sites who applied to and were eligible for the Career Academies. While the program may have changed the relationships between background characteristics and the probability of dropping out, the underlying relationship between background characteristics and the likelihood of dropping out in the absence of the Academy intervention (β) is the same for these two groups.

However, *even in the absence of the Academy program*, it is unlikely that the estimated coefficients relating the background characteristics to the dropout rate among the students who ended up in the program group would have been exactly the same as those in the non-Academy group. In other words, while the underlying relationship between background characteristics and the probability of dropping out (β) would not vary across these two samples, the idiosyncratic element (or error term) of the *estimated* relationship (β_s), and therefore the estimated relationship itself ($\hat{\beta}$), *would* vary.

Therefore, it is highly unlikely that the estimated relationship between background characteristics and dropout would have been *exactly* the same among the Academy group as it was among their non-Academy group counterparts. Because the regression weights were generated from the non-Academy group, the regression-based strategy might more accurately distinguish among students with different levels of academic risk for this group than it does for the Academy group. In other words, the risk index might distinguish different levels of risk more effectively among non-Academy students that it does among Academy group students.

This creates the possibility that, although their *observable* characteristics were the same, students in the "high-risk" non-Academy group were actually more at risk than students in the "high-risk" Academy group. It also creates the possibility that students in the "low-risk" non-Academy group were actually less at risk than students in the "low-risk" Academy group.

To the extent that this occurred, it would result in overstating positive impacts for the high-risk subgroups and overstating the magnitude of negative impacts for the "low-risk" subgroups. However, as the next section will reveal, the magnitude of this potential distortion can be estimated. Furthermore, the magnitude of the distortion appears to be minimal, and it is not large enough to have a meaningful effect on the overall pattern of impact estimates.

V. Magnitude of Potential Distortion in the Regressions-Based Approach

In order to understand whether this potential limitation outweighs the analytic advantages of the regression-based approach discussed earlier, it is important to estimate the magnitude of the potential distortion.

Theoretically, in order to estimate the magnitude of this distortion, one would like to compare the outcomes of the students within each risk category in the non-Academy group with what would have been observed among the Academy students in the same risk subgroup *in the absence of the treatment*. However, because the Academy group received the treatment and the treatment may have actually affected these outcomes, this comparison cannot be made. The ideal basis for such a comparison would be a second non-Academy group that was neither used in order to estimate the dropout regression nor exposed to the program. In the absence of any distortion, one would expect that, within each risk subgroup, the outcomes for the students in this sample would be identical to the outcomes for these in the original non-Academy group. Therefore, any differences between outcomes for these students and outcomes for the original non-Academy group could be confidently attributed to the distortion created by the regression-based strategy.

Although a second non-Academy group for this study is not available, a strategy for estimating the potential distortion in the original estimates is to use bootstrap sampling in order to simulate a second sample. Bootstrap sampling is commonly used to generate estimates of standard errors and other population characteristics from relatively small samples (Stine, 1990). It rests on the assumption that the sample from which the observations are drawn is representative of the population as a whole. In this case, to the extent that the initial non-Academy group can be thought of as representative of the population of students from whom the evaluation sample was drawn, bootstrap sampling procedures can be used to simulate new samples of non-Academy group students. Within each subgroup, these samples can be used in order to compare the outcomes for the students on whom the dropout regression was based with the outcomes for a sample of students who were not included in this regression. These differences would constitute a reliable estimate of the distortion created by the regression-based subgroup strategy.

The mechanics of this process are as follows:

1. Use a random number generator to draw a bootstrap sample of students the size of the original non-Academy group, sampling *with replacement* the observations from the original non-Academy group sample.
    a. Use a random number generator to select an observation from the original non-Academy group.
    b. Copy that observation to a new data set.
    c. Replace that observation into the sampling frame from which it was drawn (the original non-Academy group sample).
    d. Repeat steps a through c until the new sample equals the size of the original non-Academy group (n=805 times). This sample will be referred to as the *model group*.

This creates a sample which is the same size as the original non-Academy sample and which, theoretically, is drawn from the same population.[13] However, this sample is not the same as the non-Academy group, because steps a through d typically create a sample which omits several observations from the original sample and creates multiple copies of other observations.

2. Use this bootstrap sample to estimate the relationship between the six background characteristics used to define academic risk and the probability that a student will drop out of high school prior to the end of the 12th grade.

3. Repeat steps 1a through 1d to draw (with replacement) a second bootstrap sample, the size of the Academy group, from the original non-Academy sample. This sample will be referred to as the *non-model group.*

4. Repeat steps 1a through 1d once more, this time drawing from the *Academy* sample, to produce a bootstrap sample of students from each risk subgroup who received the Academy treatment. This sample will be referred to as the *simulated Academy group*.

5. Apply the coefficients from the regression model to the background characteristics of the individuals in all three bootstrap samples in order to create the risk index.

6. Use the 25th and 75th percentiles of the risk index in the first bootstrap sample (the model group) in order to divide the samples into high-, medium-, and low-risk subgroups.

7. Compare the average outcomes from the model group with those from the second bootstrap sample (the non-model group). The difference between the two groups represents the distortion created by the regression-based strategy.

8. Repeat steps 1 through 7 another 200 times. The average difference across these iterations between the subgroup outcomes for the model group and the non-model group provides a bootstrap estimate of the potential distortion created by the regression-based subgroup strategy. The average levels across these iterations among the simulated Academy group represents a bootstrap estimate of the outcome levels among the Academy students.

Table 3 presents the results of this estimation process for five key outcomes. The numbers in this table represent the average outcomes of 200 iterations of the bootstrap process describe above. As such, they are intended to simulate what one would expect to observe if one repeated the experimental analysis contained 200 times, with 200 different samples from the same population. The first column of the table presents the average outcomes among students from the bootstrap samples upon which the dropout regression was estimated (the model group). The numbers in this column represent the outcome levels one would expect to observe as a result of the regression-based approach among the sample of non-Academy students on whom the regression was fit.

[13]In particular, this replaces the unknown theoretical distribution of the population from which the non-Academy group is drawn with the empirical distribution of the non-Academy sample itself.

Table 3

Career Academies Evaluation

Outcome Levels for Bootstrap Control Samples and Program Group, by Risk Subgroups

Outcome	Model Group (%)	Non-Model Group (%)	Model Minus Non-Model	Program Group (%)	Program Minus Model	Program Minus Non-Model
High-risk subgroup						
Dropped out of high school	31	30.3	0.7 **	20.7	-10.3	-9.6
Earned credits to graduate	28.5	28.6	-0.1	40.2	11.7	11.6
Completed basic academic core	7.3	7.2	0.1	15.3	8	8.1
Reported any negative risk-taking	39.5	39	0.5	32.9	-6.6	-6.1
Reported positive youth development	55.4	56.1	-0.7 *	64.5	9.1	8.4
Medium-risk subgroup						
Dropped out of high school	8.4	8.9	-0.5 ***	9	0.6	0.1
Earned credits to graduate	63.4	63.4	0	65.6	2.2	2.2
Completed basic academic core	28.9	29	-0.1	28.2	-0.7	-0.8
Reported any negative risk-taking	25.3	25.1	0.2	24.2	-1.1	-0.9
Reported positive youth development	68.6	68.8	-0.2	70.6	2	1.8
Low-risk subgroup						
Dropped out of high school	2.8	2.9	-0.1	2.2	-0.6	-0.7
Earned credits to graduate	75.8	76.1	-0.3	84.9	9.1	8.8
Completed basic academic core	36.7	36.6	0.1	39	2.3	2.4
Reported any negative risk-taking	16.8	16.7	0.1	15.7	-1.1	-1
Reported positive youth development	77.2	76.8	0.4	80.3	3.1	3.5

SOURCES: MDRC calculations from Career Academies Evaluation Student School Records Database and 12th Grade Survey Database.

NOTES: Estimates are regression-adjusted using ordinary least squares, controlling for background characteristics of sample members. Rounding may cause slight discrepancies in calculating differences.

A two-tailed t-test was applied to differences between the model and non-model groups. In both cases, statistical significance levels are indicated as *** = 1 percent; ** = 5 percent; * = 10 percent.

Column 2 of Table 3 presents the average outcomes among students from the bootstrap samples that were not used for this regression (the non-model group). The numbers in this column represent the pattern of outcomes one would expect to observe if one had a sample of non-Academy students who were not the basis for the regression model but for whom the coefficients from the regression-based strategy were combined with individual characteristics in order to estimate the risk of school failure.

The third column of Table 3 presents the differences between the two averages for the model and non-model groups. Because the second column of estimates is not affected by the potential distortion described above, these numbers represent the estimate of the potential distortion created by the regression-based strategy for each outcome.

The fourth column of Table 3 presents the average outcomes for the high-, medium-, and low-risk subgroups from the simulated Academy (program) group. The fifth column presents the average differences between the simulated Academy group and the model group from column 1. This represents a bootstrap estimate of the program impact. The sixth column presents the average differences between the simulated Academy group and the non-model group from column 2. This represents a bootstrap estimate of the program impact, *absent any distortion created by the regression-based subgroup strategy*.

The estimates in Table 3 suggest that the magnitude of the distortion created by the regression-based subgroup strategy is not large enough to have a meaningful effect on the pattern of impacts. In particular, for each of the outcomes in this table, the estimated distortion appears to be less than 1 percentage point. For example, the first row of the table presents the bootstrap estimates of the dropout rate for the high-risk subgroup. Inasmuch as whether or not a student dropped out of high school was the dependent variable in the regression used to define the subgroups, the potential magnitude of the distortion should be largest with respect to that outcome. However, the estimate in this row suggests that the potential distortion in the impact estimate is seven-tenths of 1 percentage point. In particular, across 200 replications, the average dropout rate or the high-risk sample from the model group is 31 percent, while the average for the sample that was *not* used to estimate the regression (from the non-model group) is 30.3 percent, a difference of .7 percentage points.[14]

Columns 4 and 5 of Table 3 indicate that subtracting the potential distortion does not result in a meaningfully different estimate of the program impact. In particular, the estimate of the impact and the estimate of the impact minus any potential distortion appear to be within rounding error of one another. Moreover, the other estimates in this table reveal a similar pattern. The estimated distortion is never larger than seven-tenths of a percentage point, and the pattern of ef-

[14]An alternate estimate of the distortion was generated by performing what might be called a randomization test. This entailed taking the entire evaluation sample, including Academy and non-Academy students, and randomly assigning them to two groups. The dropout regression was then estimated within one group, and the coefficients were used to generate an index and divide the sample into risk categories in both groups. The difference between the outcomes for these groups would represent an alternative estimate of the distortion. After performing this process 200 times, it was found that this alternative method yielded a pattern of estimated distortion similar to that produced by the initial method. In particular, the estimated distortion on the dropout variable was 1.3 percentage points, and the estimated distortion on all other variables was smaller than that.

fects in the impact estimates is not substantially different from the pattern of effects in the column estimates that account for the distortion. This suggests that, while the regression-based subgroup strategy has theoretical limitations, the limitations do not have any meaningful effect on the pattern of impacts.

The asterisks in the table indicate the results of statistical significance tests regarding the differences between the model and non-model groups. They suggest that, across the five outcomes and three subgroups considered, the estimated distortion created by the regression-based subgroup strategy was statistically significant in only three cases. In particular, for the high-risk subgroup, the estimated distortion created with respect to the dropout variable and the percentage of students who participated in positive youth development activities was significantly different from zero. For the medium-risk subgroup, the distortion created regarding the dropout rate was also statistically significant. The estimated distortion across all other outcomes was not significantly different from zero.

This pattern, combined with the magnitude of the effects, suggests two conclusions. First, the estimated distortion created by the regression-based subgroup strategy appears to converge around some non-zero number, but that effect does not appear to be large enough to affect the basic pattern of impacts. Second, the distortion appears to be restricted mainly to the outcome variable that was the basis for defining the subgroups, and it was concentrated within the high-risk subgroup.[15]

IV. Conclusions

The evidence and discussion in this paper strongly support the idea that accounting for the heterogeneity of students in the Career Academies Evaluation is an important element of any strategy designed to assess the impact of the Academies on the diverse group of students they serve. Impact estimates which aggregate results across students with different academic trajectories conceal a substantial amount of variation across students in the effects of the Academies on key outcomes. Therefore, in order to assess the effects of Career Academies more sensitively, it is necessary to develop a strategy for differentiating among students who, in the absence of the Academy treatment, would experience different academic outcomes.

Traditional approaches toward defining subgroups go part of the way toward differentiating among students with different academic trajectories. However, the experimental design present in the Career Academies Evaluation provides a rare opportunity to improve on these strategies by estimating the relationship between student characteristics and the likelihood of school failure *in the absence of the Academy treatment*.

This regression-based approach offers a number of distinct advantages over its alternatives, and its potential limitations are highly unlikely to change the pattern of any of the find-

[15]It is important to note that the potential for introducing distortion or bias into the impact estimates can be avoided altogether by using an external sample to derive both the list of background characteristics and regression parameter estimates used to create the risk index. By external sample we mean a sample that is not also used in the impact analysis.

ings. The regression-based approach takes multiple factors into account, weighting them according to the strength of their effect on student failure. It also allows the use of all relevant variation in student characteristics in order to estimate risk, as opposed to classifying students on the basis of arbitrary cutoffs in otherwise continuous measures of risk. Most important, it is a highly effective strategy for identifying students who, *in the absence of the Academy intervention*, would have had substantially different outcomes. As a result, it reveals differences in the effects of Career Academies that would be masked by impacts which are averaged across the entire population of Academy students — and would be at least partly masked by traditional approaches to defining subgroups.

The major drawback of the regression-based strategy is that it has the potential to generate a distortion in the impact estimates that would overstate the impact of the Academies on students in the high-risk subgroup. However, the best estimates of the potential distortion in impact estimates suggest that its magnitude is negligible. In particular, the estimates suggest that the distortion, at it largest, is seven-tenths of a percentage point. Moreover, any distortion that exists appears to be concentrated within the high-risk subgroup and to be restricted primarily to one outcome. In other words, both the magnitude and pattern of distortion suggest that this phenomenon is neither large nor pervasive enough to affect the overall pattern of impacts.

References

Bloom, Dan, James J. Kemple, Pamela Morris, Susan Scrivener, Nandita Verma, and Richard Hendra. 2000. *The Family Transition Program: Final Report on Florida's Initial Time-Limited Welfare Program.* New York: Manpower Demonstration Research Corporation.

Eberts, Randall W. 1997. *The Use of Profiling to Target Services in State Welfare-to-Work Programs: An Example of Process and Implementation.* Staff Working Papers, No. 98-52. Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Kemple, James, and JoAnn Leah Rock. 1996. *Career Academies: Early Implementation Lessons from a 10-Site Evaluation.* New York: Manpower Demonstration Research Corporation.

Kemple, James J., and Jason C. Snipes. 2000. *Career Academies: Impacts on Students Engagement and Performance in High School.* New York: Manpower Demonstration Research Corporation.

National Center for Education Statistics (NCES). 1992. *National Education Longitudinal Study of 1988: Characteristics of At-Risk Students in NELS: 88*. Washington, DC: U.S. Government Printing Office.

Natriello, Gary, ed. 1987. *School Dropouts: Patterns and Policies*. New York: Teachers College Press.

O'Leary, Christopher J., Paul Decker, and Stephen A. Wandner. 1998. *Reemployment Bonuses and Profiling.* Staff Working Papers, No. 98-51, Kalamazoo, MI: W. E. Upjohn Institure for Employment Research.

Roderick, Melissa. 1993. *The Path to Dropping Out: Evidence for Intervention*. Westport, CT: Auburn House.

About MDRC

The Manpower Demonstration Research Corporation (MDRC) is a nonprofit, nonpartisan social policy research organization. We are dedicated to learning what works to improve the well-being of low-income people. Through our research and the active communication of our findings, we seek to enhance the effectiveness of social policies and programs. MDRC was founded in 1974 and is located in New York City and San Francisco.

MDRC's current projects focus on welfare and economic security, education, and employment and community initiatives. Complementing our evaluations of a wide range of welfare reforms are new studies of supports for the working poor and emerging analyses of how programs affect children's development and their families' well-being. In the field of education, we are testing reforms aimed at improving the performance of public schools, especially in urban areas. Finally, our community projects are using innovative approaches to increase employment in low-income neighborhoods.

Our projects are a mix of demonstrations — field tests of promising program models — and evaluations of government and community initiatives, and we employ a wide range of methods such as large-scale studies to determine a program's effects, surveys, case studies, and ethnographies of individuals and families. We share the findings and lessons from our work — including best practices for program operators — with a broad audience within the policy and practitioner community, as well as the general public and the media.

Over the past quarter century, MDRC has worked in almost every state, all of the nation's largest cities, and Canada. We conduct our projects in partnership with state and local governments, the federal government, public school systems, community organizations, and numerous private philanthropies.