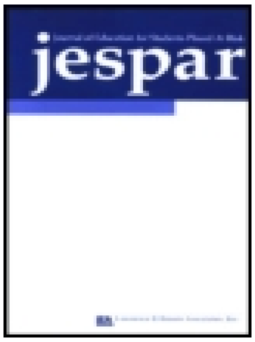

Journal of Education for Students Placed at Risk (JESPAR)

ISSN: 1082-4669 (Print) 1532-7671 (Online) Journal homepage: http://www.tandfonline.com/loi/hjsp20

Do We Know Whom to Serve? Issues in Using Risk Factors to Identify Dropouts

Philip Gleason & Mark Dynarski

JOURNAL OF EDUCATION FOR STUDENTS PLACED AT RISK, *7*(1), 25–41

Do We Know Whom to Serve? Issues in Using Risk Factors to Identify Dropouts

Philip Gleason and Mark Dynarski
Mathematica Policy Research, Inc.
Princeton, New Jersey

This article analyzes the effectiveness of widely used risk factors for identifying students who will drop out of school. The findings indicate that nearly all risk factors are not effective predictors of dropping out. The findings suggest that dropout prevention programs often serve students who would not have dropped out, and do not serve students who would have dropped out, which has implications for program effectiveness.

Dropout prevention programs commonly provide intensive services to those students deemed most likely to drop out within a school or district. To identify these students, program operators typically use "risk" factors—that is, student characteristics or measures of past school performance thought to be associated with future dropping out. Implicit in the use of these factors is the assumption that risk factors help to identify those students who will drop out if they do not receive program services. Even the highest quality dropout prevention programs will have little influence on the dropout problem if risk factors identify the wrong students (i.e., those who would not otherwise have dropped out).

Although researchers have examined student characteristics and past performance measures correlated with dropping out, many of these studies have not assessed the predictive validity of these factors. If a dropout prevention program uses a particular risk factor to identify participants, what proportion of these participants would have dropped out if they did not receive services? Alternatively, what proportion of the future dropouts of a school or community would be served by a program if that program used a particular risk factor to identify participants? In addition, if a particular risk factor used by a dropout prevention program did not pre-

Requests for reprints should be sent to Mark Dynarski, P.O. Box 2393, Princeton, NJ 08543–2393. E-mail: mdynarski@mathematica-mpr.com

dict the rate of dropping out well, how would that affect the impact of the program on a school's dropout rate? Using data from an evaluation of dropout prevention programs, we attempt to address these questions.

The main result of our analysis is that risk factors commonly used by dropout prevention programs are weak predictors of dropping out.[1] At both the middle and high school levels, students with risk factors were not likely to be dropouts 2 to 3 years later. A program designed to be large enough to serve all future dropouts in a school would end up serving well under half the dropouts if it were to use common risk factors to identify its participants. An important implication of this is that programs can have highly effective dropout prevention services, yet not significantly reduce a school's dropout rate.

The methodology we used to measure the effectiveness of risk factors for dropouts is relevant for any program that needs to target its services to a specific set of individuals. For example, states currently use risk factors to identify those unemployment insurance (UI) applicants most likely to exhaust their UI benefits for a program that provides special redeployment services. Other public programs use risk factors to identify recently employed welfare recipients who are most likely to have difficulty in keeping their jobs, and who may need employment retention services. The WIC program (Special Supplemental Nutrition Program for Women, Infants, and Children) uses "evidence of nutritional risk" as a criterion for program eligibility. Any program of this type must pay close attention to the risk factors it uses and whether the factors identify individuals who are most in need of services.

RISK FACTORS IDENTIFIED IN PREVIOUS RESEARCH

There is a well-established literature in education, sociology, and economics on factors associated with dropping out. Researchers have examined the relationships between dropping out and five types of variables:

1. Demographic characteristics and family background.
2. Past school performance.
3. Personal/psychological characteristics.
4. Adult responsibilities.
5. School or neighborhood characteristics.

[1]Throughout this article, we define students according to their dropout status; that is, whether or not they are a dropout at a particular point in time. An alternative way to conceptualize dropping out would be to define students according to whether they have ever dropped out of school.

Researchers have been in agreement on the factors related to dropping out even though their studies employed different data sources, covered different time periods, and differed in the extent to which they controlled for other factors in measuring these relationships.

Regarding demographic characteristics, many researchers have found that Black and Hispanic youths are more likely than White youths to drop out of high school (e.g., Eckstrom, Goertz, Pollack, & Rock, 1987; Natriello, McDill, & Pallas,1990; Rumberger, 1995). Researchers examining family background have found that family income, socioeconomic status, and parents' educational attainment are related to dropping out (Barro & Kolstad, 1987; Mare, 1980; National Center for Education Statistics [NCES], 1990, 1992; Rumberger, 1983, 1995). In addition, researchers have found that students with limited English proficiency are more likely to drop out (Natriello et al., 1990; NCES, 1990; Rumberger, 1995), as are students whose families receive welfare (Haveman, Wolfe, & Spaulding, 1991; Mare), and those students with a sibling who has already dropped out (NCES, 1990).

Researchers have consistently found that past school performance is related to dropping out. Pallas (1987) stated that, "Poor academic performance is the best predictor of who drops out of school" (p. 4). Important indicators of poor academic performance that are associated with dropping out have included low grades, poor test scores, and placement on a nonacademic track (Barro & Kolstad, 1987; Eckstrom et al., 1987; NCES, 1992; Rumberger 1995; U.S. Department of Education [USDOE], 1983). Other school performance measures related to dropping out have included being overage for grade level, having disciplinary problems in school, truancy, and spending little time on homework (Barro & Kolstad, 1987; Eckstrom et al. 1987; NCES, 1992; Pallas, 1984).

Researchers have found that students with adult responsibilities, such as being employed or having to take care of a child, are more likely to drop out than are their counterparts without these responsibilities. Barro and Kolstad (1987), Pallas (1987), and Rumberger (1983) have found that having a child is positively associated with a student's likelihood of dropping out. D'Amico (1984) and Pallas (1984) have identified a positive association between working more than 20 hr a week and dropping out.

Personal and psychological characteristics related to dropping out include factors that reflect students' commitment to schooling and ability to follow through on this commitment. Research by Eckstrom et al. (1987) and Rumberger (1983) has shown that students with low self-esteem and an external locus of control (i.e., a feeling that they have little control over their own destinies) are more likely to drop out. Rumberger (1983) found that students with low educational expectations or plans are more likely to drop out.

Dropping out may be influenced not only by student and family characteristics but also by characteristics of schools and neighborhoods. Researchers have shown that dropout rates are higher among students in urban schools than in rural or suburban schools (e.g., Pallas, 1987). Clark (1992), Crane (1991), and Vartanian and Gleason (1999) have shown that students living in neighborhoods whose residents are poor are more likely to drop out than those in neighborhoods with wealthier residents, even if their family incomes are not high.

A few studies have examined the predictive ability of risk factors. Lloyd (1978) looked at how well a set of student characteristics of third graders predicted whether these students would drop out after entering high school. The study found that risk factors identified a group of students who were more likely than other third graders to drop out. However, the risk factors also identified a substantial number of students who did not drop out, and did not identify a substantial number of students who ultimately did drop out. Weber (1989) reviewed several studies of dropout prediction scales and concluded that the typical dropout prediction scale identifies only about one third of actual dropouts.

More recently, McKee, Melvin, Ditoro, and McKee (1998) examined two dropout prediction scales, one based on information available from student records, and another based on information from a student survey. The administrative records scale predicted dropping out well among a sample of 10th-grade students. However, the analysis of this scale was based on a small sample (N = 49) and predicted dropping out over a short period. The dropout prediction scale was measured in October and November of a school year, and students' dropout status was measured at the end of the school year (June). It was not clear whether the administrative records scale would predict dropping out as well over a longer period.

Studies that have examined the effectiveness of risk factors in contexts other than dropout prevention have found that risk factors were effective. For example, Rangarajan, Schochet, and Chu (1998) found that risk factors effectively predicted which newly hired welfare recipients would have difficulty sustaining employment. Devaney (1996) found that some risk factors routinely used by the WIC program to identify individuals who showed evidence of nutritional risk were effective in identifying those who were truly at risk, though other risk factors were less effective.

LONGITUDINAL DATA ALLOWED RISK FACTORS TO BE ASSESSED

From the perspective of a dropout prevention program, an efficient risk factor is one that identifies as many students as possible who are likely to drop out. If the program serves most of these students, referred to here as future dropouts, the program will have the greatest likelihood of reducing the dropout rate in a school or community.

The approach we used to measure efficiency was to select a group of students at a point in time, measure their risk factors, and determine whether they dropped out of school 2 or 3 years later.[2] The efficiency of various risk factors was assessed by calculating the dropout rate for students with different risk factors—the higher the dropout rate, the more efficient the risk factor. In particular, the higher the dropout rate for a given risk factor, the larger the number of future dropouts that a dropout prevention program using this risk factor would be able to serve (assuming it had a fixed number of program slots).[3]

Data Were From Middle Schools and High Schools in Four Cities

The data were from the evaluation of the School Dropout Demonstration Assistance Program (SDDAP), sponsored by the USDOE. The SDDAP provided federal assistance to 85 local dropout prevention programs, operating in sites throughout the United States, between September 1991 and May 1995. Our analysis made use of extensive data from program schools and comparison schools in four sites: (a) Dallas, Texas; (b) Grand Rapids, Michigan; (c) Phoenix, Arizona; and (d) Santa Ana, California.

The data provided ample information about students who were disadvantaged, diverse with respect to race and ethnicity, and at high risk of dropping out.[4] The evaluation collected information on samples of middle and high school students at a subset of schools in these sites that were undergoing school wide reform designed to reduce dropping out and improve other outcomes, as well as from similar schools that were not undergoing reform. The middle school sample included data from three of the four sites (all but Phoenix, which was a high school district) in which students were sampled and interviewed at the beginning of their seventh-grade year. The baseline interview asked questions about students' characteristics and attitudes, as well as their experiences in school during the previous year, when most were sixth graders. The high school sample included data from all four sites. In three of the sites, the baseline interview took place at the beginning of the students' 10th-grade year; in Phoenix, this interview took place at the beginning of

---

[2]We analyzed students' dropout status at a point in time 2 to 3 years after their risk characteristics were initially measured. This "point-in-time" dropout status may differ from the students' ultimate dropout status, as students who drop out in one year may return to school in another, and students who remain in school one year may drop out in the next. However, the point-in-time dropout status and ultimate dropout status are likely to be highly correlated. This was the assumption made in the analysis.

[3]One might argue that, even if a risk factor does not identify future dropouts, it is likely to identify students who are struggling in school and who will benefit from the type of services provided by dropout prevention programs. The extent to which this argument is true will be discussed later.

[4]Gleason and Dynarski (1994) provide more information on the characteristics of students in these sites.

their ninth-grade year. All students in both the middle and high school samples were enrolled at the time of the baseline interview.

In all four sites, baseline data were collected from one cohort in Fall 1992 and from another cohort in Fall 1993. Students were interviewed again in Spring 1995, nearly 3 years after baseline for Cohort 1, and 2 years after baseline for Cohort 2. Baseline and follow-up data were obtained for 2,672 middle school students from three sites, and from 2,808 high school students in four sites.

The data from the four cities had many features that enabled us to assess the efficiency of risk factors for middle and high school students. However, the data were not nationally representative, so the relations we found between risk factors and dropping out were correctly interpreted as applying only to the four cities. Additional analyses would be needed to assess whether the relationships we found were also evident at the national level.

Defining Risk Factors

We defined risk factors using data from the baseline survey. Dropout status was measured 2 to 3 years later, in what should have been the end of students' 8th- or 9th-grade year among middle school students, and 11th- or 12th-grade year among high school students (10th or 11th grade in Phoenix). Dropping out was defined as not being enrolled in school, not on summer vacation, and not having earned a high school degree or General Education Development certificate, as reported by students. After the risk factors and dropout status were defined, two measures were calculated to assess the efficiency of a risk factor: (a) the proportion of all students with the risk factor, and (b) the proportion of students with the risk factor who were dropouts at the time of the follow-up survey.

The set of risk factors captured characteristics associated with dropping out derived from the research literature. Many of the risk factors were, in fact, used by programs that were in the evaluation, as noted by researchers who visited the programs.[5] Each risk factor could take on one of two values—a value of 1, if the student had the characteristic, and a value of 0, if the student did not. Some characteristics on which risk factors were based were continuous, so cutoffs were chosen to create a binary factor.[6] The set of risk factors included the following:

Family background

1. Single-parent family.

[5]Gleason and Dynarski (1994) provide more details on the definition of variables used as risk factors.

[6]The choice of this cutoff point was somewhat arbitrary, but we conducted robustness analysis whenever possible to determine whether the selection of a different cutoff point would have changed the results.

2. Family receives public assistance.
3. Primary language of family is not English.
4. Has sibling who has dropped out of school.
5. Mother did not graduate from high school.

Previous school experiences

1. Absent 20 or more times during the school year.
2. Overage for grade level by at least 1 year.
3. History of having dropped out of school (high school sample only).
4. Low grades (Cs and Ds or below).
5. Disciplinary problems during school year.
6. Has attended five or more schools during lifetime (six or more schools for the high school sample).
7. Spends less than 1 hr per week on homework.
8. Spends no time each week reading for fun.

Personal/psychological characteristics

1. External locus of control.
2. Low self-esteem.
3. Parents do not talk to student about things studied in school.
4. Student not "very sure" of graduating from high school.

Adult responsibilities

1. Has child (high school sample only).

Risk Factors Based on Multiple Characteristics

A program operator could view a student as being at risk if the student had a combination of characteristics. We addressed this possibility by constructing two risk factors, based on a combination of characteristics: (a) a composite risk factor, and (b) a regression risk factor. The composite risk factor defined a student as at risk if he or she had some minimum number of single risk factors. Here, the composite risk factor was based on eight single risk factors from the list.[7] We analyzed alternative ver-

[7]The eight contributing risk factors were selected because they were the most predictive individual risk factors. For the middle school sample they were (a) receipt of public assistance, (b) high absenteeism, (c) being overage for grade, (d) low grades, (e) disciplinary problems, (f) spending no time reading,

(*continued*)

sions of the composite risk factor based on whether students had at least two, three, or four of eight single risk factors.

The regression risk factor used a model to combine single factors. We estimated a logistic model that related 40 student characteristics and risk factors to dropping out, and used the estimated model to calculate, for each student, the predicted probability of dropping out.[8] We defined students to be at risk if their predicted probabilities exceeded a threshold, setting the threshold so that the proportion of students classified as at risk equaled the proportion of actual dropouts in the sample. Middle school students were coded as being at risk if their probability was above the 94th percentile of the middle school distribution (i.e., in the top 6%); high school students were coded as being at risk if their probability was above the 85th percentile of the high school distribution (i.e., in the top 15%).

The idea behind the regression risk factor was to use as much information as possible about a student to determine whether he or she was at risk. In this sense, the regression risk factor made the most use of the data. The performance of the regression risk factor relative to the single risk factors provided a sense of whether using additional data and more sophisticated statistical techniques helped create a better risk indicator.

RESULTS

An efficient risk factor would be highly predictive, and observable at a low cost. Ideally, students who had the risk factor would develop the "condition" and students who did not have the risk factor would not.

Our results showed that none of the single risk factors effectively identified future dropouts among middle school students, as highlighted in Table 1. The table lists

1. The single risk factors and the risk factors based on multiple characteristics.
2. The proportion of all students who were at risk, according to the factor.
3. The dropout rate among these students.

[7](*continued*) (g) having attended five or more schools, and (h) having parents who do not talk to the student about things studied in school. For the high school sample, the eight contributing risk factors were (a) high absenteeism, (b) being overage for grade, (c) low grades, (d) having a child, (e) having a sibling who has dropped out, (f) having previously dropped out, (g) being unsure of graduating from high school, and (h) spending less than 1 hr a week on homework.

[8]We estimated logistic regression models to generate the propensity score that both included and excluded a binary variable indicating whether a student was a treatment group student in a restructuring school or a control group student in a comparison school. This variable did not significantly affect dropping out, nor did it affect our results. Details of the regression models can be obtained from the corresponding author on request.

TABLE 1
Middle School Risk-Factor Efficiency

Risk Factor	*Percentage of All Students With Risk Factor*	*Dropout Rate Among Students With Risk Factor (%)*
Family characteristics		
Single parent	34	8
Public assistance	24	8
Mother was high school dropout	24	5
Sibling has dropped out	22	7
English not primary language	18	5
Previous school experiences		
Does little homework	42	6
Overage for grade	30	9
Disciplinary problems	30	9
Does not read for fun	18	9
Has attended 5 or more schools	17	10
Low grades	12	8
High absenteeism	6	15
Overage by 2 or more years	5	16
Personal and psychological characteristics		
External locus of control	46	7
Not sure of high school graduation	37	8
Low self-esteem	33	7
Parents do not talk to student about school	17	11
Multiple characteristics		
Composite risk factor (3 of 8)	15	14
High absenteeism or overage by 2 or more years	11	15
Composite risk factor (4 of 8)	6	19
Propensity score risk factor	6	23

Note. From Gleason and Dynarski (1998). The sample size was 2,568. Among this sample, 6% dropped out within 2 to 3 years.

A risk factor was deemed efficient if students who had the factor also had a high dropout rate. Overall, 6% of the middle school students in the sample had dropped out by the end of eighth or ninth grade.[9]

[9]Among students in both cohorts, the dropout rate after 2 years (by the end of the eighth grade) was 5.5%. Among students in Cohort 1 with 3 years of follow-up data, the dropout rate at 3 years (by the end of the ninth grade) was 7.5%.

The factors associated with the highest dropout rates were high absenteeism (students with that factor had a dropout rate of 15%) and overage by 2 or more years (students with that factor had a dropout rate of 16%). These two risk factors were more efficient than other risk factors that all had dropout rates below 11%.

High absenteeism and being overage by 2 or more years were more efficient than other single risk factors, but they were not efficient in the absolute sense. In fact, dropout prevention programs using these risk factors would not reach most of their target populations. Only 15% of students with these two factors were dropouts at the time of the follow-up survey. Programs using these risk factors to identify participants would end up providing dropout prevention services mostly to students who would not have been dropouts.

Composite risk factors were more efficient than the most efficient single risk factors, as expected. The dropout rate for students with the composite factor, defined by whether students had four of eight single factors, was 18%, slightly higher than the dropout rates for students with the high absenteeism and overage risk factors (see Table 1). The regression risk factor predicted dropping out most efficiently. The dropout rate among students defined as at risk by the regression risk factor was 24%, almost four times as high as the overall dropout rate. However, even the regression risk factor did not identify most dropouts. More than three fourths of students who were identified as at risk of dropping out according to the regression risk factor were not dropouts 2 to 3 years later.[10]

A numerical example might help to underscore the implications of using inefficient risk factors. Suppose that a large middle school had 1,000 students beginning seventh grade. Suppose further that, on the basis of past experience, the school expected 60 of these students to drop out by the end of what would be their ninth-grade year, so it established a dropout prevention program to serve 60 students. If the program used the regression risk factor to identify 60 participants, the program would end up serving only 14 students who would have been dropouts (about 23% of 60). The other 46 students served by the program would not have been dropouts. The flip side is that the regression risk factor would incorrectly code 46 students as not at risk; these students would not be served by the program, even though they would eventually drop out.

Using inefficient risk factors means that even if a program provided highly effective dropout prevention services, it would not reduce the dropout rate much. For example, if the program described in the previous example prevented half the dropouts it served from dropping out, it would reduce the number of dropouts in

[10]The regression risk factor does better if we consider a broader definition of educational failure. In particular, if we consider an individual to have failed if he or she drops out, is absent daily or almost daily, skips school more than once a week, receives mainly Ds and FS in school, or has been arrested, slightly less than half the students with the regression risk factor will have failed (overall, 22% of the middle school sample falls into this category).

the school from 60 to 53 (from 6.0% to 5.3%), an overall impact on the dropout rate of only 12%. Although the program might serve 60 students, it would prevent only 7 students who would have dropped out from doing so. The cost of using inefficient risk factors is that the program's maximum feasible impact of 50% would be reduced to only 12%.

The inefficiency of risk factors might be mitigated in two ways. First, a program that served less students overall might more easily identify those students most at risk for dropping out. In the previous example, suppose that the program was able to serve only 20 students. Under this assumption, a program that selected students in the top 2% of the regression distribution (the top 20 out of 1000 students) would serve a larger proportion of dropouts. In our data, the dropout rate for these students was 27%, more than four times larger than the overall dropout rate. To be more efficient, then, programs may need to be smaller, presenting a tradeoff between coverage and efficiency that we will discuss more next.[11]

Second, measuring risk factor efficiency within a relatively short time frame may understate efficiency if students would have dropped out at a later point. In the previous example, of the 46 students that were served by the program but had not dropped out 2 years later, some may have dropped out 4 years later, increasing the ultimate efficiency of the risk factor somewhat.[12]

The high school analysis yielded results analogous to those of the middle school analysis, as shown in Table 2. The high school sample consisted of 2,615 students, most beginning the 10th grade. Within 2 to 3 years, 15% were dropouts.[13]

As in the middle school analysis, none of the single risk factors efficiently identified dropouts.[14] Of the single risk factors, previous school experiences were generally the best predictors. For example, 31% of high school students were overage for grade, and 21% of these were dropouts at the time of the follow-up. However,

[11]Students who are not served 1 year may be served later on. For example, a student who is not served because he or she exhibits no risk factors may later develop a risk factor. For example, a student who once attended school regularly may begin to miss school more often and be categorized as a frequent absentee. This might then qualify them for service by a program. However, as long as students do not exhibit risk factors, it is unlikely that they will be served at any point because there would be no information to suggest that serving them would be useful.

[12]In the theoretical case in which no students drop out 2 years after risk factors are assessed, and all students drop out 4 years after risk factors are assessed, risk factors would have no predictive value within 2 years, and would be perfectly predictive within 4 years.

[13]Among students in both cohorts, the dropout rate after 2 years for the high school sample (by the end of the 11th grade) was 13.1%. Among students in Cohort 1 with 3 years of follow-up data, the dropout rate after 3 years (by the end of the 12th grade) was 17.0%.

[14]The high school risk factors included having ever dropped out or having a child, as well as the factors used for middle school students. Definitions of some risk factors also differed for the high school sample—a student had to have attended six or more schools (rather than five or more) to be considered at risk, and the two composite variables required individuals to have had two out of eight and three out of eight single risk factors to be considered at risk (rather than three and four out of eight).

TABLE 2
High School Risk-Factor Efficiency

Risk Factor	*Percentage of Students With Risk Factor*	*Dropout Rate Among Students With Risk Factor (%)*
Family characteristics		
Single parent	35	18
Public assistance	15	19
Mother was high school dropout	26	18
Sibling has dropped out	22	21
English not primary language	19	17
Previous school experiences		
High absenteeism	9	27
Overage for grade	31	21
Overage by 2 or more years	9	28
Low grades	14	27
Disciplinary problems	19	23
Ever dropped out	7	13
Does little homework	31	21
Does not read for fun	16	17
Has attended 6 or more schools	18	17
Personal and psychological characteristics		
External locus of control	38	18
Low self-esteem	23	19
Not sure of high school graduation	23	25
Parents do not talk to student about school	22	20
Watches much television	15	17
Adult responsibilities		
Has child	3	32
Multiple characteristics		
High absenteeism or overage by 2 or more years	17	27
Low grades or has child	18	28
High absenteeism or has child or ever dropped out	17	27
Composite risk factor (2 of 8)	30	25
Composite risk factor (3 of 8)	14	34
Propensity score risk factor	15	42

Note. From Gleason and Dynarski (1998). The sample size was 2615. Among this sample, 14.6% dropped out within 2 to 3 years.

even using the best single risk factors, students classified as being at risk were much more likely to be in school 2 to 3 years later than to be dropouts. The composite risk factor based on whether students had at least three of eight specific risk factors was associated with a dropout rate of 34% among students it identified, higher than the dropout rate of 28% for the top single risk factor.

The regression risk factor performed best, with a dropout rate among identified students of 42%. Nonetheless, the regression risk factor was wrong more often than it was right. In particular, 58% of the students identified as likely to drop out using the regression did not drop out.[15] Considering that the regression risk factor used 40 variables measuring student characteristics and past school performance, dropping out clearly was difficult to predict.

A numerical example illustrates the implications of using an inefficient risk factor to identify high school dropouts. Suppose that a high school expects 146 of its 1,000 entering 10th graders to drop out, and establishes a dropout prevention program to serve 146 students. If the school uses the regression risk factor to identify program participants, about 61 program participants would be dropouts in the future. Thus, the program would serve more students who do not need dropout prevention services than students who do need them. Of the146 future dropouts in the school, 85 would not be detected and would not be served by the program. Ultimately, even if the program provided services that prevented 31 of the 61 future dropouts it served from dropping out, it would reduce the overall school dropout rate from 14.6% to only 11.5%, a reduction of 21% compared to the assumed maximum feasible reduction of 50%.

As with middle school students, having less capacity to serve dropouts might increase efficiency. If a program had the capacity to serve only 50 students, it could identify them by using the top 5% of students in the regression distribution, for which the dropout rate is 53%—so that at least the majority of students served by the program would be dropouts. Here again, the increase in efficiency is at the expense of coverage because the program is set up to serve only one third of the actual dropouts in the school.

COULD OTHER FACTORS PREDICT DROPPING OUT BETTER?

The analysis here suggests that the task of predicting who drops out is more difficult than simply taking into account known risk factors. Even for the most efficient risk factor, many of those served may not need the services, and many of those in need may not be served. However effective the program's dropout prevention services may be, using inefficient risk factors means that the program will have a modest impact on the school's dropout rate.

[15]Using the broader definition of educational failure described previously, 33% of the high school sample and 65% of those with the regression risk factor would fail. Thus, more than one third of the students served by a dropout prevention program would not need these services, according to this measure.

One could argue that it is not surprising that characteristics measured when students are beginning the 7th grade do not predict accurately whether they will drop out. The factors leading to the dropout decision may become evident only later. However, risk factors measured in the 7th grade and risk factors measured in the 10th grade do not reliably predict dropping out, suggesting that the problem is not simply the point when risk factors are measured.

Characteristics or factors not examined in this article might possibly have better identified future dropouts. Four possible sets of factors include the following:

1. Ecological characteristics.
2. Unobserved psychological factors.
3. Measures of the persistence of specific characteristics over time.
4. Transitory events occurring after the point at which baseline risk factors were measured.

Ecological factors that might have influenced dropping out, but that were not measured in this article, include neighborhood characteristics and peer group characteristics. Students growing up in very poor neighborhoods and surrounded by many dropouts have been shown to be more likely to drop out than students living in other areas, where staying in school is more the norm (Crane, 1991). Being able to carefully measure neighborhood or peer group characteristics may improve the ability of a regression risk factor to predict dropping out. However, our findings indicated that knowing whether a student had a sibling who dropped out did not predict well whether that student would drop out. This suggests that knowing peer or neighborhood characteristics may improve prediction only marginally.

Unobserved psychological factors may be important determinants of whether individuals drop out. Risk factors examined here include students' self-esteem, locus of control, and educational expectations. These risk factors may measure the underlying psychological constructs imprecisely, or other psychological factors may be more important in predicting dropping out. For example, the degree to which individuals persevere through difficult times or are optimistic about the future may better predict their future success in school. However, these constructs are difficult to measure.

The other two explanations involve viewing risk in a more dynamic way. The risk factors used in this article measured student characteristics at a single point in time and school performance in a single academic year. Characteristics or factors that lead to dropping out may have a more cumulative effect. Performing poorly in school one year may lead to temporary disappointment and an increased resolve to do better; performing poorly for several years in a row may lead students to become detached from school and believe that they are failures, eventually leading them to drop out (Finn, 1989). Risk factors that measure trends in student performance over several school years may be better predictors of dropping out than those that measure performance over a single year.[16] However, the longitudinal data available here do not support a test of this type of trend risk factor.

Some students may have dropped out not because of the cumulative effects of poor academic performance but because of an unexpected event. A student may have become a parent, been arrested, started using drugs, or had serious personal problems at home. Despite previous success in school, the event may have led the student to decide abruptly to leave. Again, the likelihood of these events occurring would be difficult to quantify.

An alternative approach for identifying students in need of dropout prevention services is self-selection. The person most likely to know whether a student will drop out may be the student himself or herself. Dropout prevention programs that serve students who apply on their own may be more likely to serve students who truly need services than programs that use risk factors to identify at-risk students. For this approach to be effective, however, students would not only have to expect to drop out, but would also have to be motivated to seek help. The one piece of evidence we have about this approach is not promising. We found that identifying students as being at risk if they reported that they were "not sure" that they would graduate from high school predicted future dropping out no better than many of the other risk factors we examined.

CONCLUSION

What is the "right" level of efficiency for risk factors? We have assumed up to this point that programs want to be as efficient as possible in identifying dropouts. This is a sensible perspective if dropouts are expensive to serve, or if the benefits of serving them are moderate, in which case programs want to ensure that their services reach those most in need. However, if serving dropouts is not expensive and the benefits of serving them are large—for example, if the savings to society are large in terms of more earnings and reduced public assistance (among students who do not drop out because of program intervention)—then programs can be effective from a benefit–cost perspective even if they do not serve many dropouts. Levin (1972) argued that the benefits of serving dropouts are on the order of $6 for every

[16]Some of our analysis casts doubt on the possibility that risk factors measured over more than 1 year would do better. We measured students' risk factors over 2 years and examined the dropout rate between the 2nd and 3rd years. Overall, 7% of students dropped out between the 2nd and 3rd years. Among students with high absenteeism (one of the best individual risk factors) in the first year, 15% had dropped out by Year 3. Adding Year 2 information improved the efficiency of this risk factor only slightly—among students with high absenteeism in both the first and 2nd years, 17% had dropped out by Year 3.

$1 spent serving them. With this high a return, programs can use inefficient risk factors yet be solid investments from a social perspective.

Regardless of whether programs have large benefit–cost ratios, however, they can be more beneficial if they target their services more efficiently. That is, society is better off if programs use more efficient risk factors, as long as the expense of using them is not so considerable as to outweigh the gains in terms of additional dropouts helped. The regression risk factor constructed here would improve on single risk factors that programs commonly employ and would be a useful starting point for considering how to better identify likely dropouts.

ACKNOWLEDGMENT

We would like to thank Alan Hershey and Audrey Pendleton for their support and helpful comments, and two referees for helpful comments.

The findings in this article are based on research performed under contract LC91015001 with the U.S. Department of Education. The contents of the article do not necessarily reflect the views or policies of the U.S. Department of Education, nor does any mention of trade names, commercial products, or organizations imply endorsement of these by the U.S. Government.

REFERENCES

Barro, S. M., & Kolstad, A. (1987). *Who drops out of high school? Findings from high school and beyond* (Report no. CS 87–397c). Washington, DC: U.S. Department of Education.

Clark, R. L. (1992, December). *Neighborhood effects on dropping out of school among teenage boys* (Population Studies Center Discussion Paper Series, UI–DSC–UI–13). Washington, DC: The Urban Institute.

Crane, J. (1991). The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing. *American Journal of Sociology, 96*, 1226–1259.

D'Amico, R. (1984). Does employment during high school impair academic progress? *Sociology of Education, 57*, 152–164.

Devaney, B. (1996). *WIC Nutrition Criteria: A Scientific Assessment* (Report of the Committee on Scientific Evaluation of WIC Nutrition Risk Criteria, Food and Nutrition Board).Washington, DC: Institute of Medicine.

Eckstrom, R. B., Goertz, M., Pollack, J., & Rock, D. (1987). Who drops out of high school and why? Findings from a national study. In G. Natriello (Ed.), *School dropouts: Patterns and policies* (pp. 52–69). New York: Teachers College Press.

Finn, J. (1989). Withdrawing from school. *Review of Educational Research, 59*, 117–142.

Gleason, P. M., & Dynarski, M. (1994). *Falling behind: Characteristics of students in federally funded dropout prevention programs* (Report submitted to the U.S. Department of Education). Princeton, NJ: Mathematica Policy Research.

Gleason, P., & Dynarski, M. (1998). *Do we know whom to serve? Issues in using risk factors to identify dropouts* (Report submitted to the U.S. Department of Education). Princeton, NJ: Mathematica Policy Research.

Haveman, R., Wolfe, B., & Spaulding, J. (1991). Childhood events and circumstances influencing high school completion. *Demography*, *28*(1), 133–157.

Levin, H. (1972). *The costs to the nation of inadequate education*. Washington, DC: Government Printing Office.

Lloyd, D. N. (1978). Prediction of school failure from third-grade data. *Educational and Psychological Measurement, 38,* 1193–1200.

Maddalla, G. S. (1983). *Limited dependent and qualitative variables in econometrics*. Cambridge, England: Cambridge University Press.

Mare, R. D. (1980). Social background and school continuation decisions. *Journal of the American Statistical Association, 75*, 295–305.

McKee, J. M., Melvin, K. B., Ditoro, V., & McKee, S. P. (1998). SARIS: Student at-risk identification scale. *Journal of At-Risk Issues, 4*(2), 24–33.

National Center for Education Statistics. (1990). *A profile of the American eighth grader*. Washington, DC: U.S. Department of Education.

National Center for Education Statistics. (1992). *Characteristics of at-risk students in NELS:88* (Report No. 92–042). Washington, DC: U.S. Department of Education.

Natriello, G., McDill, E. L., & Pallas, A. M. (1990). *Schooling disadvantaged children: Racing against catastrophe.* New York: Teachers College Press.

Pallas, A. M. (1984). *The determinants of high school dropout.* Unpublished doctoral dissertation, Johns Hopkins University, Baltimore.

Pallas, A. M. (1987). *School dropouts in the United States*. Washington, DC: U.S. Department of Education.

Rangarajan, A., Schochet, P., & Chu, D. (1998). *Employment experiences of welfare recipients who find jobs: Is targeting possible?* Princeton, NJ: Mathematica Policy Research.

Rumberger, R. W. (1983). High school dropouts: A review of issues and evidence. *Review of Educational Research, 20,* 101–121.

Rumberger, R. W. (1995). Dropping out of high school: The influence of race, sex, and family background. *American Educational Research Journal, 20*(2), 199–220.

U.S. Department of Education. (1983). *High school dropouts: Descriptive information from high school and beyond* (Publication No. NCES 83–221b). Washington, DC: National Center for Education Statistics.

Vartanian, T. P., & Gleason, P. M. (1999). Do neighborhood conditions affect high school dropout and college graduation rates? *Journal of Socioeconomics, 28*(1)*,* 21–41.

Weber, J. M. (1989). *Identifying potential dropouts: A compilation and evaluation of selected procedures.* Columbus, OH: Ohio State University.