The Effect of School Inputs on Dropout Rates in Minnesota School Districts

David M. Vock

Abstract

Because of secondary school dropouts confront higher unemployment, lower wages, and a plethora of other societal problems, policy makers have a vested interest to find efficient ways to prevent dropouts. This study examined the effect of school inputs and expenditures on the likelihood of students dropping out of Minnesota school districts in 2003-2004. Hierarchal logistic modeling was used to account for the effect of an individual student's race and gender and the demographic factors of the school district to assess the impact of school inputs on the likelihood of students dropping out. The analysis indicated that students attending schools districts with larger percentages of non-white students and students enrolled in free and reduced lunch programs are more likely to dropout even after accounting for the race of the student. A diverse staff appears to lower dropout rates as increases in the percentages of non-white faculty were associated with decreasing dropout rates. But a more educated and well-paid faculty was related to an increased likelihood of dropping out. The results suggest that Minnesota should consider programs to even out the racial and economic composition of students in the school districts. Furthermore, rather than pay teachers more to obtain advanced degrees, Minnesota school districts should invest in recruiting a more diverse staff.

Introduction

Secondary school dropouts confront a bevy of potential problems of concern to public policy makers. According to data from the National Education Longitudinal Study of 1988, students who entered high school in 1988 and who dropped out and never received an equivalent high school degree had an unemployment rate of 11.4 percentage points higher and average annual earnings $3,100 less than traditional four-year high school graduates. The drawbacks of dropping out are particularly stark for women dropouts who had an unemployment rate of 26.4% (Klenzl and Kena 2006). Data further suggest that policy makers should focus on preventing secondary school dropouts because obtaining an equivalent high school degree does not appear to confer the same benefits as a traditional one. While half of all eighth graders in 1988 who dropped out of high school received an equivalent diploma by 2000, those students still faced increased unemployment rates and decreased earnings potential compared to traditional graduates. In fact, students receiving their high school degree more than six years after entering high school had higher unemployment rates and lower wages than those who never completed a high school degree (Klenzl and Kena 2006). Other research has noted increasing post-secondary schooling options presents the only advantage of obtaining an equivalent high school degree after dropping out (Cameron 1993).

In addition to societal costs incurred from higher unemployment and lower wages, dropouts commit crime at higher rates, have more health problems, and rely on government assistance programs more frequently (Rumberger 1995). High school dropouts' poor performance in government training programs designed to reduce unemployment and improve wages, such as the Comprehensive Employment and Training Act and Job Training Partnership Act, likely exacerbate the unemployment and wage gap (Goldschmidt 1999). Without any action to limit the number of high school dropouts, the societal problems stemming from dropping out are likely to get worse for two reasons. First, demographic shifts in the Untied States toward groups that historically have larger percentages of dropouts could increase the number of dropouts (Rumberger 1995). Secondly, the wage gap between people with and without high school degrees is increasing (Goldschmidt 1999).

Besides the economic problems associated with secondary school dropouts, lawmakers face legal pressures to provide quality education or adequate funding to prevent dropouts. In Minnesota a federal district court declared in a 1971 case, *Van Dusartz v. Hatfield* that Minnesota's school finance system was unconstitutional because of inequality between school districts. In *McDuffy v. Secretary of Education* (1993), the Supreme Judicial Court of Massachusetts ruled that poor communities were not receiving their constitutionally assured adequate education (Faville 2006). In addition to legal pressures, policy officials must address increasing demands for improvement from federal legislation such the No Child Left Behind Act.

As demands to improve public education increase, educational researchers have tried to elucidate which controllable factors are associated with improving dropout rates. The effect of educational funding and the allocation of resources on dropouts particularly concerns policy analysis. Economists confront large obstacles in identifying these factors because of limited longitudinal and student-specific data and difficulties in performing experiment that actually alter school inputs (a study by Kruger (1999) was able to test the effect of reduced class size through a controlled experiment, though). Establishing the effect of funding on the likelihood of dropping out is made more difficult in Minnesota because per pupil expenditures are fairly evenly distributed across the state and actually increase for students more at risk of dropping out. The Minnesota state legislature allocates each year a minimum per pupil allowance. However, the state recognizes that providing an adequate education cost more for certain at risk students and provides additional money to districts as the number of pupils with limited English proficiency or enrolled in free or reduced lunch programs increases. School districts can raise local levies through referenda which does create some variability in the amount of funding between school districts not associated with differences in the numbers of at risk students.

Previous Work

In spite of these difficulties, researchers have tried to determine how educational funding or educational inputs affect dropout rates using several different approaches including estimating a cost function for a given level of adequacy, deriving a production function where the dropout rate is a function of both school inputs and characteristics, and using hierarchal logistic modeling which separates individual student effects from the between-school effects.

- **Cost Function Adequacy**

To assess the effect of school resources, specifically expenditure, on educational achievement, some researchers have proposed developing a cost function for a given level of educational attainment desired while accounting for demographic factors of the district, the cost structure the district faces, and efficiency of a distinct. More precisely, Faville (2006) notes that school or district expenditures (E_j) can be thought of as a function of the number of inputs (I_j), the cost of the inputs (C_j), and the unobserved characteristics of a district including efficiency (u_j)

(1) $$E_j = f\,(I_j, C_j, u_j)$$

Similarly educational attainment or output (O_j) for school j can be thought of as a function of various school inputs (I_j), school demographics (D_j), and community or family demographics (F_j).

(2) $$O_j = g\,(I_j, D_j, F_j)$$

Equation (2) can be rearranged so that inputs are a function of educational attainment yielding equation (3) below. This can be substituted into equation (1) to relate educational output to expenditures (4).

(3) $I_j = h\ (O_h, D_j, F_j)$

(4) $E_j = j\ (O_h, D_j, F_j, C_j, u_j)$

The educational output of school/district and the school/district expenditures may be simultaneously determined which necessitates using two-stage least squares regression (2SLS). However, Faville (2006) notes the difficulty in finding instrumental variables (amongst demographic data) readily available to predict the educational output for a given district.

Although not particularly intuitive predictors of school achievement, Faville used measures of the community's wealth, educational attainment, and willingness to fund education as instruments for the prediction of the composite score on the Texas Assessment of Knowledge and Skill (TAKS) in the first stage regression. In the second stage, Faville found that after accounting for the school demographics, the teacher wage index, and an efficiency measure, the composite TAKS score was a significant predictor at the α=.10 level of the expenditure per pupil. An increase in the mean composite TAKS score for a district of 4.9 points (one quarter of standard deviation amongst school districts) would require the school district to increase expenditures by a factor of 3.53.

Rather than use demographic data from the district as instruments to account for the simultaneity bias, Duncombe et. al. (2003) used instruments derived from student characteristics of adjacent school districts. More specifically, the log of pupil density, average LEP students, and maximum and minimum performance on grade 8 exams in adjacent district were used as instruments to predict the performance index, a composite measure of the New York state exam and Regents exam, for each district in the first stage regression. After accounting for the demographic composition of the district, teacher salary, and a cost index, the performance index was found to be a statistically significant predictor of per pupil expenditure.

Generating a cost function provide useful "bottom line" information for policymakers to assess the additional cost of lowering the dropout rate or to determine how funds should be better allocated amongst different school districts. The method unfortunately does not specify a direct relationship between dropout rate and school expenditure and never relates school inputs like student-pupil ratio to dropout rate except possibly through simultaneous equations.

- **Production Function**

The production function method eliminates some of the shortcomings of the above method by having school inputs, school demographics, and community or family demographics predict educational output (see equation 2). As with the cost function approach, any form of educational output including dropout rate may be used. Many studies just consider the effects of school inputs such as pupil-teacher ratio on educational output and never specificnally relate educational inputs back to educational expenditure (equation 2). One can rearrange equation (1) so that school inputs is a function of educational expenditure and then substitute this into equation (2) to obtain:

(5) $O_j = h\,(E_j, C_j, u_j, D_j, F_j)$

Regardless of whether researchers use equation (2) or (5) to estimate the production function, educational output is still likely to be simultaneously determined with school expenditures/inputs. As in the cost function, researchers have used 2SLS to account for the simultaneity bias. Using the production function approach Faville (2006) notes that median household income, tax price, and other district level demographic data can be used as instruments for school expenditure which is more intuitive than using the same demographic data as instruments for educational output in the cost function approach.

Other approaches have been used to deal with the simultaneity bias. Initial work using the production function to predict dropout rates focused on school, family, and community demographics and often omitted school expenditures/inputs as predictors and with it any need to consider the simultaneity bias. Using data from the National Longitudinal Survey of Youth Labor Mark Experience, Russel Rumberger (1983) regressed the probability of an individual dropping out on several student-level demographic factors using a probit model. Different models for each gender/race group (Black, Hispanic, and White – for a total of six models) were developed. Rumberger found that a cultural index which accounts for the prevalence of reading material and newspapers in the home was the most uniform predictor dropout rate. The mother or father's education and/or earnings also predicted whether a student would dropout in several of the models. Aside from not including school inputs, the potential applicability of the method is limited because the large amount of information needed on each individual student.

Some previous research using school or district-level data has used ordinary least squares regression and simply ignored the simultaneity bias. In a study of 293 public secondary schools in New Jersey all the school inputs considered, including the percentage of teachers with a BA, pupil-teacher ratio, and average teacher salary, were not found to significantly predict the percentage of students retained by the district - the inverse of dropout rate (Fowler and Walberg 1991). Because Fowler and Walberg did not consider any simultaneity bias between the retention rate and the school inputs, the magnitudes of the coefficients for the school input variables could be biased and erroneously lead the authors to conclude that the school input factors are not significant predictors.

Rather than use a full set of simultaneous equations to account for the possible endogeny between the school input/expenditure variables, several researchers have used versions of path analysis. This method allows researchers to specify the indirect effects of school inputs/expenditures or other resources on dropout rates by relating them to other factors found to relate significantly to dropout rate. For example, school size was shown to correlate positively with a composite variable for program/facility diversity and negatively with school social climate, another composite variable. The program/facility and social climate variables were in turn correlated positively and negatively, respectively, with the school dropout rate (Pittman and Haughwout 1987). This path analysis suggests that larger school size is related to higher dropout rates. Similarly research using ordinary least squares regression found that educational expenditure per student average daily attendance day (ADA) did not significantly predict dropout rates after accounting for the demographic background of the school and school inputs such as support staff-teacher ratio and pupil teacher ratio (Fitzpatrick and Yoels 1992). However, educational expenditure ADA did significantly predict both the staff-teacher ratio and pupil teacher ratio after accounting for the background demographics of the district. The results suggest that the effects of increasing per student expenditures on dropout rates are mediated through school structure variables. Unfortunately, path analysis does not allow for a quantitative assessment of the effect of school size or per student educational expenditure on dropout rates. Furthermore, because a full system of simultaneous equations was not specified, the correlations could be biased.

More recent research has constructed a full set of simultaneous equations to characterize the effect of school expenditure on achievement – although no studies found used this method with school dropout rate as the measure of educational output. However, the structural equations used when other measures of achievement are considered can be applied to dropout rate studies. For example, some researchers have treated measures of school expenditure as exogenous and then related expenditures to school environment and teacher-student ratio. Student-teacher ratios were allowed to affect school environment and both student-teacher ratios and school environment were used to predict achievement (Wenglinsky 1997). With this model, per student instructional expenditures were found to significantly predict achievement on the National Assessment of Educational Progress; an increase in $4,000 dollars on instructional expenditure was associated with an increase in 3.2 points on the exam where 12 points represents a grade level.

Rather than treat expenditures as exogenous, other researchers have postulated that educational achievement and expenditure are simultaneously determined. To account for this, researchers have used 2SLS with school demographic data and measures of community wealth, education, and willingness to pay for education as instrumental variables to model expenditure per pupil. The modeled estimate of per pupil expenditure is then used along with the demographic data to predict school achievement (Faville 2006). School expenditure was found not to predict significantly the school district result on the Texas standardized tests after accounting for demographic factors. Faville, however, did not include any other measure of school input such as student-teacher ratio nor did he divide expenditure per student into different types (instructional, administrative, etc.) as Wenglinsky had done.

The prevalence of the production function method for various forms of school achievement has led to work on combining the results of the various studies to obtain a better overall understanding of school inputs on achievement. Hanushek (1989) identified 187 studies which attempted to quantify the effect of inputs on educational outcomes (some of which used dropout rate) and, using vote-counting methodology, concluded that there was little evidence that school inputs affected outcomes. However, a re-analysis of the work using combined significance tests and combined estimation methods of meta-analysis suggested that there is a positive relationship between resource inputs and school outcomes (Hedges, Lane, and Greenwald 1994). While not all of the 187 studies used in the analysis concerned dropouts, the work highlights the ambiguous relationship between school inputs and achievement

Despite the prevalence of the model, researchers have noted two drawbacks to applying production function models to school dropout rates. First, dropout rates are not perfect measures for school achievement and may not be appropriate for a production function model (Wenglinsky 1997). Secondly, researchers often place individual predictors of dropping out (socio-economic status, race, etc. of the student) with school-wide predictors (socio-economic composition of the entire school, pupil-teacher ratio, etc.) This approach fails to consider the inherent structure of the data and can lead to statistical bias in the coefficients and misinterpretation (Goldschmidt and Wang 1999).

- **Hierarchal Generalized Linear Model**

To account for the second shortfall above, many researchers have turned to hierarchal generalized linear models (Rumberger 1995, McNeal 1997, Goldschmidt and Wang 1999, Rumberger and Thomas 2000). Using a logistic model, individual characteristics of student *i* in school *j*, including demographics (D_{ij}), such as race and gender; family characteristics (F_{ij}), including socioeconomic status, parent education, single parent household, and parental involvement; and student characteristics (S_{ij}), including previous academic achievement, engagement in class, and behavior, are used to predict the probability of the student dropping out.

(6) $$\log(p_{ij}/(1-p_{ij})) = \beta_{0j} + \beta_{1j}(D_{ij}) + \beta_{2j}(F_{ij}) + \beta_{3j}(S_{ij}) + \varepsilon_{ij}$$

In the second level of the model, student-level coefficients were allowed to vary between different school districts. School or district-wide data, such as overall student composition (SC_j), including the racial composition and average socio-economic status; school resources (R_j), such as pupil-teacher ratio, per student expenditures, and teacher experience; the structure of the school (ST_j), including school size, location, and control of the school (public or private); and school process/climate (P_j) which considers factors like absenteeism, were used to predict the student-level coefficients. Random effects were included in each of the level two equations to allow for variation between schools.

(6) $$\beta_{pj} = \gamma_{p0} + \gamma_{p1}(SC_j) + \gamma_{p2}(R_j) + \gamma_{p3}(ST_j) + \gamma_{p4}(P_j) + \mu_{pj}$$

Using this approach has lead to some consistent findings on the effects of school composition and inputs on dropout rates. While Goldman and Wang (1999) did not consider differences in educational inputs between schools when analyzing data from the National Education Longitudinal Survey (NELS), they did find that several school composition and community factors significantly predicted dropout rates after accounting for individual student factors. For example, increases in the percentage of parents with low levels of education, community members with low socio-economic status, and students held back or misbehaving were found to increase the odds of dropping out. However, students enrolled private religious and secular school had lower odds of dropping out. Without considering any school expenditure/input data, the study found that school-level factors accounted for approximately two-thirds of the difference between dropout rates in different schools.

Rumberger’s (1995) research on eighth grade dropouts also using the NELS reached many of the same conclusions on the influence of school composition on dropout rates.

Specifically, Rumberger found that after considering individual effects, mean socioeconomic status of the students, the percentage of minority students, and the percent perceiving the discipline system to be fair all significantly predicted mean dropout rates. When a model was constructed with only students attending lower socio-economic schools, school inputs such as student-teacher ratio in addition to the percentage held back and size of the school were associated with dropping out. More specifically, decreasing the student-teacher ratio by one student decreased the odds of dropping out by approximately 4% all else constant. Subsequent work using NELS data on the dropout rates of urban and suburban high schools confirmed that school inputs such as student-teacher ratio and percentage of excellent teachers are associated with dropout rates after accounting for student-level factors and school composition and structural characteristics (Rumberger and Thomas 2000). Similar to the previous studies mentioned, Rumberger and Thomas also found that mean socio-economic status of the school, school size, and the control of the school, public or private, significantly predicted dropout rates holding all else constant. Using data from the High School and Beyond database, McNeal (1997) also reached similar conclusions about the effect of student-teacher ratios on dropouts using a hierarchal linear model.

Hierarchal linear models traditionally require substantial data on both students who did and did not dropout. The data used in the research described above all came from longitudinal government surveys that asked detailed questions on the attitudes and academic habits of students. While hierarchal generalized linear models allow specification of different levels of correlation, they currently do not allow one to account for any simultaneity biases in the data. The other shortcoming of hierarchal generalized linear models is that researchers can only model dropping out versus not dropping out. Multinomial logistic regression, which Rumberger and

Larson (1998) used to model the outcome of the several possible choices students face such as dropping out, transferring, or staying in the same school, has not been extended to hierarchal modeling.

Research Questions

While most of the initial work on hierarchal models used data collected from longitudinal surveys, little work has been done to apply this method to school districts in one state. Because school districts in one state are more homogenous than ones from across the country, the coefficients for the school input terms in the regression would be less likely to capture other factors that influence dropout rate if all the school districts were from one state. An analysis of all the districts in one particular state would provide far more salient information for state-level policy makers. Therefore, the goal of this study was to extend the use of hierarchal generalized linear models to readily available census data for students in Minnesota public schools in order to assess the effect of school inputs and expenditures on the likelihood of students in grades 7 though 12 dropping out in 2003-2004.

Methods

- **Data and Variables**

Data on the number of dropouts, their race, and gender, by school district for grades 7-12 were obtained from the Minnesota Department of Education (DOE) online database for the 2003-2004 school year. Information concerning average teacher's salary, experience and age and the percentage of teachers by race, gender, and level of education were also acquired from the Minnesota DOE online database for the 2004-2005 school year, the oldest school year available. All other district-level data concerning demographics, teacher and staff resources, funding, and community demography were acquired through the National Center for Education Statistics

Common Core of Data (CCD) online database. The community data was also obtained from the CCD and contains information from the 2000 census (**Table 1**). While there is some slight discrepancy between the years of data the differences should not be great enough to significantly influence the results. Most of the data is only separated by one school year and the community-level data, which is three to four years older, likely has not changed dramatically over a few years.

All Minnesota school districts that were classified as local school districts and had enrolled students at both the junior high/middle school and high school levels were included in the study. Regional education agencies such as charter schools, specialty schools, and the state juvenile prison schools were excluded. Three school districts that did not have data on the surrounding community were not included in the analysis. In all 329 Minnesota school districts were examined.

While Minnesota tracks the number of dropouts, the number of "non-dropouts" in a given school year is not precisely determined. Instead both the Minnesota DOE and the CCD publish the number of students enrolled on October 1. For the purposes of this study, it will be assumed that the total enrollment on October 1 is the enrollment at the start of the school year. The number of non-dropouts will be calculated by subtracting the dropouts from this total. This process resulted in a negative number of non-dropouts in a few race classes in a couple of school districts. These negative values were changed to zero. The dropout rate is conventionally determined by taking the number of dropouts and dividing by the October 1 enrollment.

Of the school districts included in the study, the 2003-2004 dropout rates ranged from 0.0% to 15.9% with a mean of 1.29%. All but six school districts had dropout rates less than 5.0%. These six school districts could pose difficulty in modeling dropout rates as they might exert a particularly large influence on the estimated parameters for the district-level effects. However, for the hierarchal logistic models (see statistical methods) the variance about the regression increases as the probability of dropping out (up until the probability is equal to .5), so to some extent the method has some resistance to outliers.

The range of values, mean, and standard deviation for the possible explanatory variables are presented in **Table 1**. These covariates under consideration are all either fairly normally distributed or slightly right-skewed and should not pose a problem for the analysis. Perhaps the two exceptions to this claim above are the percentage of non-white students and the percentage of limited English proficiency students (**Figure 1, 2)**. There is considerable variability in these percentages between school districts with several outliers. However, regression analyses do not generally have any distributional assumptions for the explanatory variables. Of course there is some concern that the school districts with these extreme values could have a significant effect on the estimated parameters for these variables. Subsequent analyses will have to be done to assess the magnitude of those school districts influence on the model.

Of particular concern to this analysis is the high correlation between the different potential explanatory variables (**Figure 3- 5**). If the purpose of the study were to predict dropout rates in different Minnesota school districts, then the exact subset of explanatory variables selected for the model would not particularly matter and variables could be selected that are not substantially correlated. However, this study is concerned with the effect of certain predictor variables, such as school funding and resources, and dropout rates. Correlation between the explanatory variables is undesirable on the one hand because the inclusion or exclusion of one may substantially affect the interpretation of another variable. But inclusion of several correlated variable allows one to determine which factors directly affect the dropout rate whereas which factors are mediated though other ones.

- **Statistical Methods**

To assess the effect of various inputs on dropout rates, a hierarchal logistic model was employed. The advantage of this approach as compared to the cost or production function methods is that it allows for the consideration of individual level data which should lead to better estimates for the district-level effects. While, as noted above, there would be difficulties obtaining student level data, each state including Minnesota tracks the race and gender of each dropouts. Other student-level data, such as measures of student engagement, may be determined significantly by school inputs or climate and may capture the effects of related district-level factors of concern to policy makers. Thus, an analysis without the survey data may provide better evidence of district-level factors on the likelihood of students dropping out.

Given the data available for Minnesota school districts, the race and gender of student *i* from school district *j* were used to predict the probability (p_{ij}) of that student dropping out (1).

(1) $$\log[(p_{ij})/(1- p_{ij})] = \beta_{0j} + \beta_{1j}(AMI_{ij}) + \beta_{2j}(API_{ij}) + \beta_{3j}(HIS_{ij}) + \beta_{4j}(BLK_{ij}) + \beta_{5j}(MALE_{ij}) + \varepsilon_{ij}$$

where $\varepsilon_{ij} \sim N(0,\sigma^2)$

In the second level of the model, the baseline probability of a student dropping out was allowed to vary according to district level demographic data, teacher and staff input, teacher quality, funding, and community level demographics. That is the coefficient of the intercept term (β_{0j}) was allowed to vary between each school district according to the district and community level data described above. A random effect term (ξ_j) was included for each school district (2) (see **Table 1** for abbreviations of variable names).

(2) $\beta_{0j} = \gamma_{00} + \gamma_{01}$(% non-white) + γ_{02}(% migrant) + γ_{03}(% free and reduced) + γ_{04}(% LEP) + γ_{05}(% IEP) + γ_{06}(pupil-teacher ratio) + γ_{07}(instructional aide) + γ_{08}(school admin) + γ_{09}(guidance counselor) + γ_{010}(teacher salary) + γ_{011}(% new teacher) + γ_{012}(% non-white teachers) + γ_{013} (avg years experience) + γ_{014}(% teacher adv degree) + γ_{015}(% at least associate) + γ_{016}(median family income) + γ_{017}(instructional expenditure per student) + ξ_j

where $\xi_j \sim N(0, \tau^2)$

This second level of the model accounts for the correlation between students attending the same districts. Attempts to account for variation in the effect of race or gender between the different school districts as well as adding an additional level to the model, such as economic region, proved too computationally difficult for the statistical software package used. This results in the following overall composite model (3).

(3) $\log[(p_{ij})/(1- p_{ij})] = \gamma_{00} + \gamma_{01}$(% non-white) + γ_{02}(% migrant) + γ_{03}(% free and reduced) + γ_{04}(% LEP) + γ_{05}(% IEP) + γ_{06}(pupil-teacher ratio) + γ_{07}(instructional aide) + γ_{08}(school admin) + γ_{09}(guidance counselor) + γ_{010}(teacher salary) + γ_{011}(% new teacher) + γ_{012}(% non-white teachers) + γ_{013}(avg years experience) + γ_{014}(% teacher adv degree) + γ_{015}(% at least associate) + γ_{016}(median family income) + γ_{017}(instructional expenditure per student) + $\beta_{1j}(AMI_{ij}) + \beta_{2j}(API_{ij}) + \beta_{3j}(HIS_{ij}) + \beta_{4j}(BLK_{ij}) + \beta_{5j}(MALE_{ij}) + \xi_j + \varepsilon_{ij}$

where $\varepsilon_{ij} \sim N(0,\sigma^2)$ and $\xi_j \sim N(0, \tau^2)$

Empirical Results and Discussion

Results from a number of models run with different interaction terms considered are included in **Tables 2** and **3**. For the purposes of the following discussion, the results from model 5 will be used since it is rich enough to include important information about factors that affect dropout rate without including insignificant interaction terms **Table 4**. While this model will be the focus of the analysis, any terms in this model that differ from the patterns observed in the other models will be noted.

- **Individual – level effects**

Both the race and the gender of the student were found to be significant as predictors of an individual student dropping out. Holding all other factors constant, the odds of a male student dropping out were 51.3% higher (95% confidence interval 42.0% to 61.2%). Students that were American Indian, Asian Pacific Islander, Hispanic, and Black all had statistically higher probabilities of dropping out as compared to White students (**Table 4**). The findings on the effect of race of the individual confirm those of McNeal (1997). These results differ from Goldschmidt and Wang's (1999) work which found that *ceteris paribus* African-Americans, Asians, and females were less likely to drop out. However, their analysis included far more student level characteristics, such information on family economic background, parents' education, and behavior, that are likely correlated with race and gender. Earlier work by Rumberger (1995) noted that demographic factors such as race and gender became insignificant after accounting for other factors including family background.

- **District – level demographic data**

Even after considering the race of the individual student, the percentage of non-white students in the school district was found to significantly predict the probability of a student dropping out from school in that district. The odds of a student dropping out from a school district with 10 percentage points more non-white students is 28.7% greater (95% confidence interval 20.1% to 36.0%) than a comparable student attending an otherwise similar district. The magnitude of this effect did not generally differ for students of different races (models 3, 4, and 6). However, there is some evidence that the effect of the percentage of non-white students is greater for Asian and Pacific Islander students (model 4). The results confirm the research of Rumberger (1995) and McNeal (1997) and suggest that large numbers of students of students at risk for dropping out put additional strains on the district that lead to higher probabilities of dropping out. There is also some evidence that increasing the percentage of students eligible for free and reduced lunch by 10% in the school district leads to an increase in the odds of all students dropping out of 8.0% (95% confidence interval -.679% to 17.5%). Individual-level data on the participation in the free or reduced lunch plan was not included, so it is difficult to determine if the increased probability is due to the fact that there are more individual students at risk for dropping out (an individual-level effect) or the fact that larger number of students at risk create a climate for dropping out that affects all students. In so far as the percentage of students in free and reduced lunch is a measure of the school socio-economic status, the results match those of Rumbereger (1995) and Rumberger and Thomas (2000) who found that mean socio-economic status of the school influenced dropout rates. Surprisingly, an increase in the number limited English proficiency students is associated with a decrease in the probability of dropping out. The percentage of migrant students and the percentage of students with an individualized education plan in the district did not significantly impact the probability of students in that district dropping out.

- **Human Resource Inputs**

In all the models examined, lower pupil-teacher ratios were significantly associated with a reduction in the likelihood of dropping out (z-stat 3.82, p-value <0.001). Reducing the ratio by one student results in a 7.3% (95% confidence interval 3.6% to 10.9%) decrease in the odds of dropping out. However, there is some evidence that the student-teacher ratio has far smaller effect for non-white students (model 6). This finding on the significance on pupil-teacher ratios on dropout rates furthers the work of Rumberger (1995), Rumberger and Thomas (2000), and McNeal (1997).

In model 5, increasing the number of instructional aides per student was found to be significantly associated with an increased probability of dropping out (z-stat = 2.09, p-value = 0.037). Of all the models entertained, the positive association between dropping out and instructional aides per student was only found in model 5. The number of school administrators per student and guidance counselors per student did not significantly predict the probability of a student dropping out.

- **Teacher Qualities**

Not surprisingly, a greater percentage of new teachers in a school district is significantly associated with an increase in the probability of a student dropping out (z-stat = 4.84, p-value < 0.001). An increase in the number of new teachers by one percentage one resulted in a 5.1% (95% confidence interval 3.1% to 7.1%) increase in the odds of a student dropping out. Presumably large numbers of new teachers indicate poorer teacher quality which could reasonably increase the likelihood of dropping out. While no study reviewed tied the likelihood of dropping out to the number of new teachers, Rumberger (1995) did find that teacher quality does impact dropout rates. The apparent detrimental effect of younger teachers does not seem to correspond with an increased benefit from having a well-experienced faculty. The average number of years of experience does not significantly predict the probability of a student dropping out (z-stat = .01, p-value=.993). Furthermore, a more educated, measured by the percentage that hold advanced degrees, and well-paid staff seems to be related to an increase in the likelihood of a student dropping out, all other factors equal. An increase in the number of teachers with advanced degrees increases the cost of labor in most districts without necessarily improving the quality of teaching. The increased cost of a more educated teacher, which in many cases the district cannot avoid due to tenure considerations, may prevent the district from hiring additional new teachers. Similarly the increased cost of employing a highly experienced teacher could also explain why the average number of years of experience for teachers does not significantly predict the probability of dropping out. The effect, or lack thereof, of average years of experience and teacher educations could explain the positive association between higher teacher salaries and increased likelihood of dropping out. In so far as higher average salaries are associated with larger amounts of experience and education which may not offer a substantial improvement in teaching, higher salaries will prevent money from being used to hire more teachers.

An increase in the percentage of non-white staff members was significantly associated with a decrease in the probability of dropping out (z-stat = -5.07, p-value < 0.001). The effects of increasing the number of minority faculty by one percentage point are fairly substantial: the odds of dropping decrease by 9.5% (95% confidence interval 5.9% to 12.9%). The results suggest that a diverse faculty does create a tangible benefit in terms of educational output. The benefits of a diverse faculty are realized by all racial groups but there is evidence that the benefit is greatest for Asian and Pacific Islander students (model 4).

- **Community-level effects**

The average income for families in the school district does have a significant effect on the probability of a student from that district dropping out (z-stat = -6.98, p-value < 0.001). An increase in the median family income of $10,000 is associated with a 32.4% decrease (95% confidence interval 24.6% to 39.4%) in the odds of dropping out. However, the percentage of adults in the school district with at least an associate's degree was positively associated with the probability of dropping out (z-stat = 2.93, p-value = .003).

- **Funding**

The total instructional expenditure per student was not found to significantly predict dropout rates (z-stat = .48, p-value = .632). Model 6, though, does suggest that increasing funding may be associated with a significant decrease in the likelihood of dropping out for non-white students.

Policy Implications and Discussion

Although perhaps not politically popular, the results suggest that the state should consider mandating some level of integration between school districts in Minnesota. In addition to ensuring adequate racial balance between the schools, Minnesota should consider ensuring that the economic background of the students, in particular those that qualify for free or reduced school lunches, should be evenly distributed across districts. In addition, the findings indicate that school districts have done an adequate job ensuring that large numbers of students in certain risk groups (migrant students, limited English proficiency, and individual education plan) do not create a climate where students overall are more likely to dropout. From a policy perspective, this could indicate that the Minnesota state funding formula which allocates additional funding to schools with larger amounts of students with limited English proficiency and individual education plans have served its purpose.

The results suggest that school districts should focus on improving pupil-teacher ratios and perhaps should divert some money from paying school administrators or instructional aides to hiring more teachers. Minnesota should create programs that not only encourage districts to decrease their pupil-teacher ratio but that create a greater pool of qualified teachers in the state. In particular, Minnesota should focus on developing initiatives that encourage non-whites to pursue careers in teaching. The models suggest that there are tangible benefits to having a diverse faculty and that these gains are realized by students of all races. Typically several have argued for an increased presence of minority faculty so that minority students could more closely identify with some teachers. The underpinnings of this argument suggest that minority students would benefit more from a diverse staff than white students. But the data suggest that all races benefit fairly equally and significantly from a diverse staff.

This finding on the diversity of the faculty and others suggest that school districts should re-evaluate the factors that they consider when hiring or paying teachers. School districts should reconsider paying for teachers to obtain advanced degrees because the model suggests that there is little benefit to having high percentage of faculty with advanced degrees. Because the average number of years of experience is not associated with the likelihood of a student dropping out, this seems to suggest some plateau in teaching effectiveness. Districts must also carefully weigh the appropriate pay increases a teacher should receive for increasing experience. While experience does not seem confer any benefit to reducing the dropout rate, poor wages for experienced teachers could cause them to leave the district. This would require the district to hire new teachers which have been shown to have a detrimental effect on dropout rates. A school district with flexibility in its labor force would ideally like to hire teachers with a few years experience and then structure its wages so that pay levels off after a certain number of years of experience. In short, school districts should spend less on developing a more-educated and heavily experienced staff and instead focus recruiting a diverse faculty. Because of licensure requirements, the supply of minority teachers is relatively fixed. Therefore, increases in the efforts of districts to recruit minority teachers must be offset by efforts by the state or other agencies to get more minorities into the teaching profession.

This overall finding on the influence of funding on dropping out does not necessarily suggest that school funding is not relevant to student performance. Pupil-teacher ratios as noted earlier are associated significantly with lower dropout rates and are significantly correlated instructional spending per student. Pupil-teacher ratio may account for all the variability in funding within this model. That is the effect of funding may be mediated through pupil-teacher ratios and other factors considered in this model. This result perhaps suggests that instructional money spent beyond teacher salaries or other factors included in the model is not significantly related to dropout rate.

The results also suggest that school districts need to focus on hiring more teachers rather than other school personnel. The data indicate that schools do not necessarily improve the dropout rate by increasing the number of administrators overseeing the school. The result that increases in the number of instructional aides are associated with increases in the probability of dropping out seem to conflict with the findings on the impact of student-teacher ratios. While the finding is not particularly intuitive, increases in the percentage of students with limited English proficiency and individualized education plans are associated with increases in the number of instructional aides per students. Perhaps the increased probability of those students dropping out is actually reflected in the increased probability of dropping out associated with increases in the instructional aides per student (**Figure 6, 7**).

This study concerned the effects of district inputs on the likelihood of a student dropping out, but dropout rate is only one measure of a school district performance. The policy implications above only apply to the findings on dropouts. Some of the factors above that do not appear to play a role in dropping out may in fact play crucial roles in other measures of school achievement. More research should be done that utilizes hierarchal modeling to determine the district-level effects on other measures of school achievement.

In spite of the limitations, the results do provide a starting point for districts to consider which inputs actually are associated in educational achievement. The study contradicts traditional thinking that a more educated and experienced faculty will always lead to improved educational outcomes. Instead these data indicate that a diverse faculty leads to improved educational outcomes for all races of students. Rather than pay for teachers to obtain advanced degrees, districts should spend money recruiting a diverse staff.

Figures and Tables

Table 1. Numerical summaries of the covariates considered in the model.

Covariate Name	Covariate short name	Year	Mean	Standard Deviation	Minimum	Maximum
Percentage of non-white students	% non-white	2003-2004	8.96	12.16	0	100
Percentage of Students on Free or Reduced Lunch	% free and reduced	2003-2004	31.55	14.01	3.17	84.17
Percentage of Students with Limited English Proficiency	% LEP	2003-2004	2.56	5.02	0	33.54
Percentage of Students with an Individual Education Plan	% IEP	2003-2004	13.70	3.06	6.6	25.75
Percentage of Migrant Students	% migrant	2003-2004	0.21	1.67	0	17.16
Pupil-Teacher Ratio	pupil-teacher ratio	2003-2004	14.97	2.55	7.1	22.8
Instructional Aides per Student	instructional aide	2003-2004	0.018	0.0076	0	0.046
School Administrators per Student	school admin	2003-2004	0.0027	0.0010	0	0.0066
Guidance Counselors per Student	guidence counselor	2003-2004	0.012	0.00077	0	0.0040
Average Teacher Salary	teacher salary	2004-2005	$42,983	$4,623	$31,767	$55,773

Percentage of non-white Teachers	% non-white teachers	2004-2005	0.857	1.847	0	16.0
Percentage of New Teachers	% new teachers	2004-2005	3.31	2.90	0	16.0
Average Years of Experience for Teachers	avg years experience	2004-2005	16.35	2.62	9.44	24.81
Percentage of Teachers with Advanced Degrees	% teacher adv degree	2004-2005	34.06	18.39	0	81.00
Educational Expenditures per Student	instructional expenditure per student	2003-2004	$7,916	$1,166	$6,019	$14,953
Percentage of Adults with an Associates Degree	% at least associate	2000	32.97	12.29	9.04	82.44
Median Family Income	median family income	2000	$49,138.00	$11,497.00	$19,969.00	$96,855.00

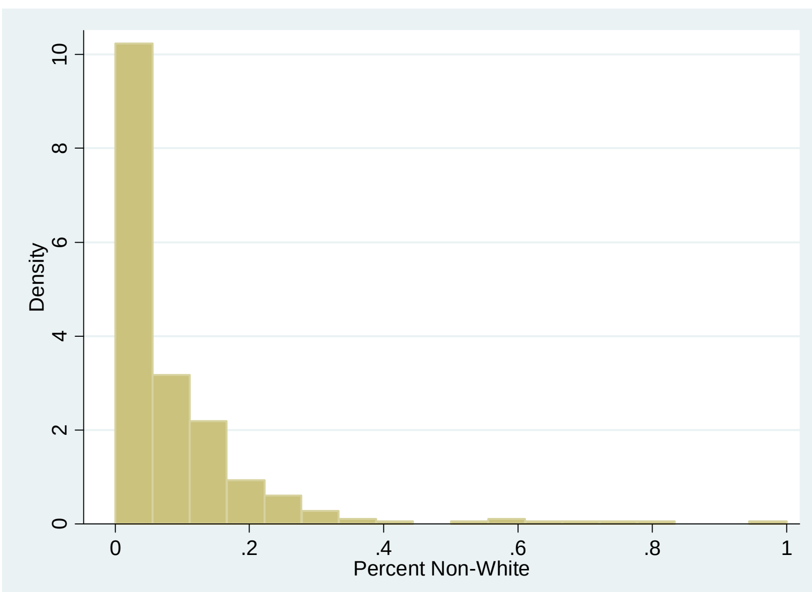

Figure 1. The distribution of the percentage of non-white students in Minnesota school districts in 2003-2004.

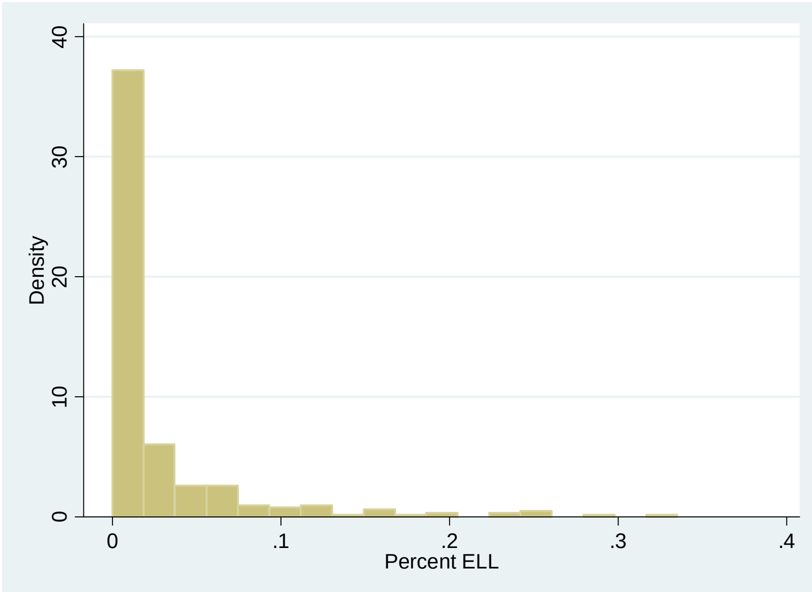

Figure 2. The distribution of the percentage of limited English students in Minnesota school districts in 2003-2004.

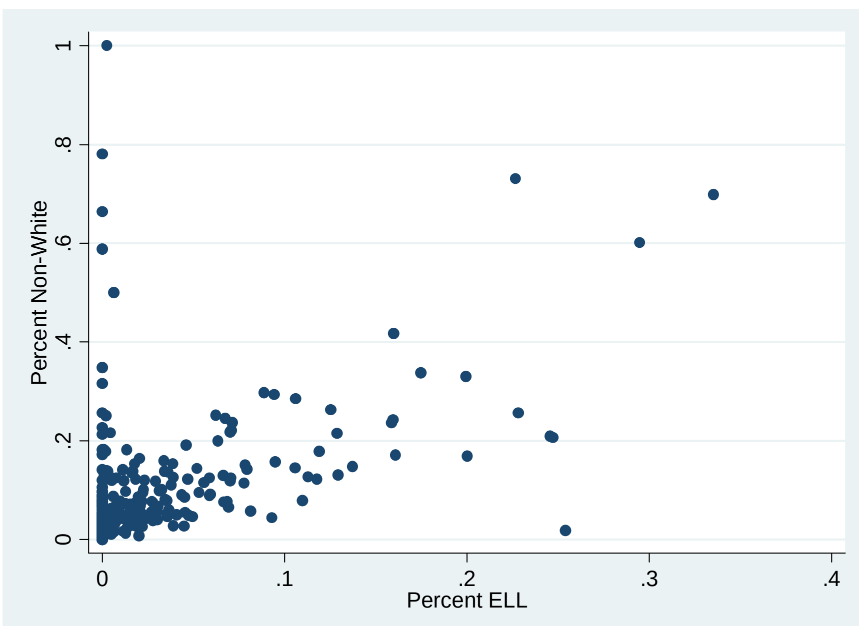

Figure 3. Relationship between the percentage of limited English students and the percentage of non-white students in Minnesota school districts.

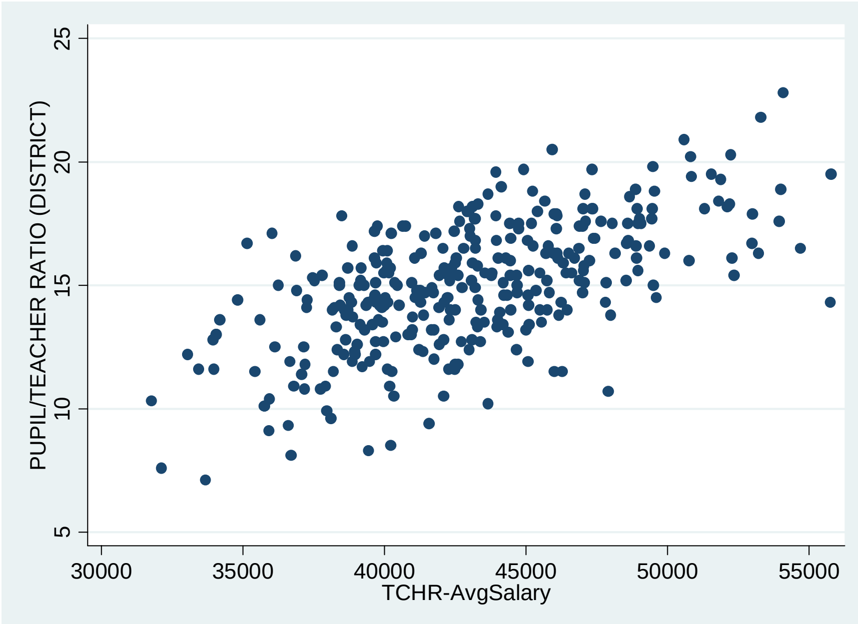

Figure 4. Relationship between the average teacher salary and the pupil-teacher ratio in Minnesota school districts.

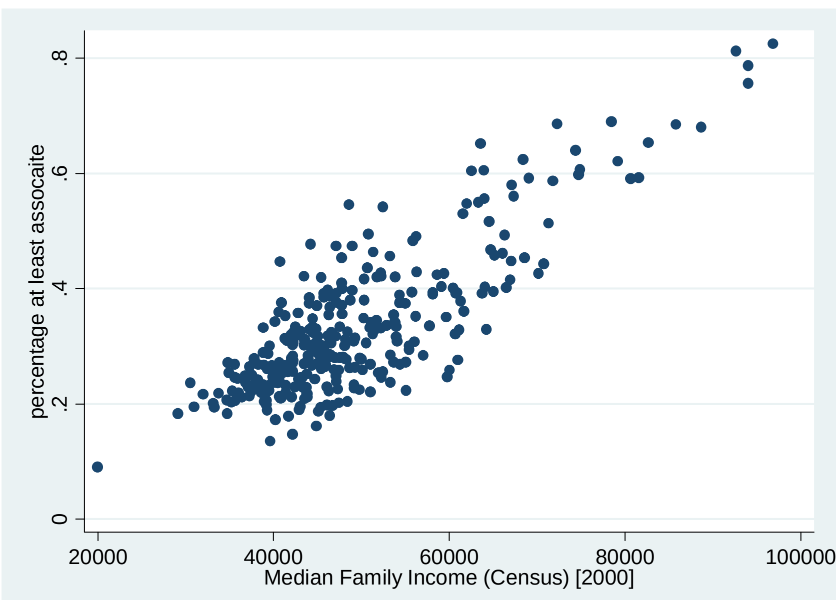

Figure 5. Relationship between the percentage with an associates degree and the median family income in Minnesota school districts.

Table 2. Coefficients for the covariates in the various models. The estimated coefficients with p-value < 0.01 are shaded a deep orange, those with p-values between 0.01 and 0.05 are shaded a lighter orange, and those with p-values between 0.05 and 0.10 are colored tan.

	Model Number					
Variable	1	2	3	4	5	6
Percentage of non-white students	2.043	2.410	2.777	3.007	2.456	1.578
Percentage of Migrant Students	6.334	2.449	0.0597	-1.319	-0.450	2.461
Percentage of Students on Free or Reduced Lunch	0.365	0.440	1.159	1.161	0.772	0.350
Percentage of Students with Limited English Proficiency	-1.453	-0.671	-0.572	-0.783	-3.730	0.417
Percentage of Students with an Individual Education Plan	4.632	3.862	0.456	0.142	-0.0412	4.222
Pupil Teacher Ratio	0.0432	0.131	0.123	0.125	0.0762	0.0654
Instructional Aides per Student	-4.208	0.722	-2.253	-3.645	9.626	7.559
School Administrators per Student	15.393	46.984	94.268	61.809	66.398	54.137
Guidance Counselors per Student	-142.656	-84.911	-54.515	-61.368	-50.203	-88.918
Average Teacher Salary	0.0000405	0.0000294	0.0000162	0.0000192	0.0000356	0.0000501
Percentage of New Teachers	5.428	4.178	6.232	4.910	5.263	0.189
Percentage of non-white Teachers	-0.462	-7.565	-13.357	-14.429	-9.927	-7.079
Average Years of Experience for Teachers	0.0188	0.00334	0.0306	0.0190	0.000130	-0.0104
Percentage of Teachers with Advanced Degrees	1.624	0.745	1.355	1.164	1.338	1.296
Percentage of Adults with an Associates Degree	1.624	0.0309	1.107	1.405	1.057	0.307
Median Family Income	-0.0000391	-0.0000364	-0.0000451	-0.0000462	-0.0000392	-0.0000395
Educational Expenditures per Student		0.0001141	0.0001086	0.0000511	0.0000385	-0.000016
Male	0.414	0.335	0.416	0.414	0.414	0.416
American Indian	1.241			1.193	1.199	
Asian Pacific Islander	0.374			-0.121	0.391	
Hispanic	1.314			1.163	1.314	
Black	1.027			1.143	1.057	
Non-white		0.990	0.979			0.416
Metro Area			0.0624			
Non-white x % non-white			0.482			0.571
American Indian x % non-white				-0.115		
Asian Pacific Islander x % non-white				4.122		
Hispanic x % non-white				0.760		
Black x % non-white				0.247		
American Indian x % non-white teachers				2.328		
Asian Pacific Islander x % non-white teachers				-15.567		
Hispanic x % non-white teachers				2.332		
Black x % non-white teachers				-4.776		

Non-white x Pupil-teacher ratio						-0.0744
Non-white x Instructional aide						9.842
Non-white x % Non-white teachers						-0.690
Non-white x Avg Exp Teachers						-0.00203
Non-white x Educational Expenditure per student						-0.0002282

Table 3. Odds ratios for an increase in the covariate listed in column two or compared to the reference group. Because the odds ratio for an interaction term is not easily interpretable, the coefficient is given for these terms.

		Model Number					
Variable	Comparison	1	2	3	4	5	6
Percentage of non-white students	10.00%	1.227	1.273	1.320	1.351	1.278	1.171
Percentage of Migrant Students	10.00%	1.884	1.277	1.006	0.876	0.956	1.279
Percentage of Students on Free or Reduced Lunch	10.00%	1.037	1.045	1.123	1.123	1.080	1.036
Percentage of Students with Limited English Proficiency	10.00%	0.865	0.935	0.944	0.925	0.689	1.043
Percentage of Students with an Individual Education Plan	10.00%	1.589	1.471	1.047	1.014	0.996	1.525
Pupil Teacher Ratio	1.00	1.044	1.140	1.131	1.133	1.079	1.068
Instructional Aides per Student	0.01	0.959	1.007	0.978	0.964	1.101	1.079
School Administrators per Student	0.001	1.016	1.048	1.099	1.064	1.069	1.056
Guidance Counselors per Student	0.01	0.240	0.428	0.580	0.541	0.605	0.411
Average Teacher Salary	$10,000	1.499	1.342	1.176	1.212	1.428	1.650
Percentage of New Teachers	1.00%	1.056	1.043	1.064	1.050	1.054	1.002
Percentage of non-white Teachers	1.00%	0.995	0.927	0.875	0.866	0.905	0.932
Average Years of Experience for Teachers	1.00	1.019	1.003	1.031	1.019	1.000	0.990
Percentage of Teachers with Advanced Degrees	10.00%	1.176	1.077	1.145	1.123	1.143	1.138
Percentage of Adults with an Associates Degree	10.00%	1.176	1.003	1.117	1.151	1.111	1.031
Median Family Income	$10,000	0.676	0.695	0.637	0.630	0.676	0.674
Educational Expenditures per Student	$1,000		1.121	1.115	1.052	1.039	0.984
Male	Female	1.513	1.398	1.516	1.513	1.513	1.516
American Indian	White	3.459			3.297	3.317	
Asian Pacific Islander	White	1.454			0.886	1.478	

Hispanic	White	3.721			3.200	3.721	
Black	White	2.793			3.136	2.878	
Non-white	White		2.691	2.662			1.516
Metro Area	non-Metro area			1.064			
Non-white x % non-white				0.482			0.571
American Indian x % non-white					-0.115		
Asian Pacific Islander x % non-white					4.122		
Hispanic x % non-white					0.760		
Black x % non-white					0.247		
American Indian x % non-white teachers					2.328		
Asian Pacific Islander x % non-white teachers					-15.57		
Hispanic x % non-white teachers					2.332		
Black x % non-white teachers					-4.776		
Non-white x Pupil-teacher ratio							-0.0744
Non-white x Instructional aide							9.842
Non-white x % Non-white teachers							-0.690
Non-white x Avg Exp Teachers							-0.00203
Non-white x Educational Expenditure per student							-0.00023

Table 4. Odds ratios and corresponding 95% confidence intervals for an increase in the covariate listed in column two or compared to the reference group for model 5.

Covariate	Comparison	Coefficient	p-value	Odds Ratio	Lower Bound 95% Confidence Interval of Odds Ratio	Upper Bound 95% Confidence Interval of Odds Ratio
Percentage of non-white students	10.00%	2.456	<0.001	1.28	1.202	1.360
Percentage of Migrant Students	10.00%	-0.450	0.844	0.96	0.611	1.496
Percentage of Students on Free or Reduced Lunch	10.00%	0.772	0.072	1.08	0.993	1.175
Percentage of Students with Limited English Proficiency	10.00%	-3.730	<0.001	0.69	0.613	0.774
Percentage of Students with an Individual	10.00%	-0.0412	0.979	1.00	0.732	1.354

Education Plan						
Pupil Teacher Ratio	1.00	0.0762	<0.001	1.08	1.038	1.122
Instructional Aides per Student	0.01	9.626	0.037	1.10	1.006	1.205
School Administrators per Student	0.001	66.398	0.158	1.07	0.997	1.016
Guidance Counselors per Student	0.01	-50.203	0.189	0.61	0.286	1.280
Average Teacher Salary	$10,000	0.0000356	0.001	1.43	1.153	1.768
Percentage of New Teachers	1.00%	5.263	<0.001	1.05	1.032	1.077
Percentage of non-white Teachers	1.00%	-9.927	<0.001	0.91	0.871	0.941
Average Years of Experience for Teachers	1.00	0.000130	0.993	1.00	0.972	1.029
Percentage of Teachers with Advanced Degrees	10.00%	1.338	<0.001	1.14	1.101	1.186
Percentage of Adults with an Associates Degree	10.00%	1.057	0.003	1.11	1.036	1.193
Median Family Income	$10,000	-3.92E-05	<0.001	0.68	0.605	0.754
Educational Expenditures per Student	$1,000	0.0000385	0.632	1.04	0.888	1.217
Male	Female	0.414	<0.001	1.51	1.420	1.612
American Indian	White	1.199	<0.001	3.32	2.861	3.849
Asian Pacific Islander	White	0.391	<0.001	1.48	1.233	1.773
Hispanic	White	1.314	<0.001	3.72	3.318	4.176
Black	White	1.057	<0.001	2.88	2.522	3.282

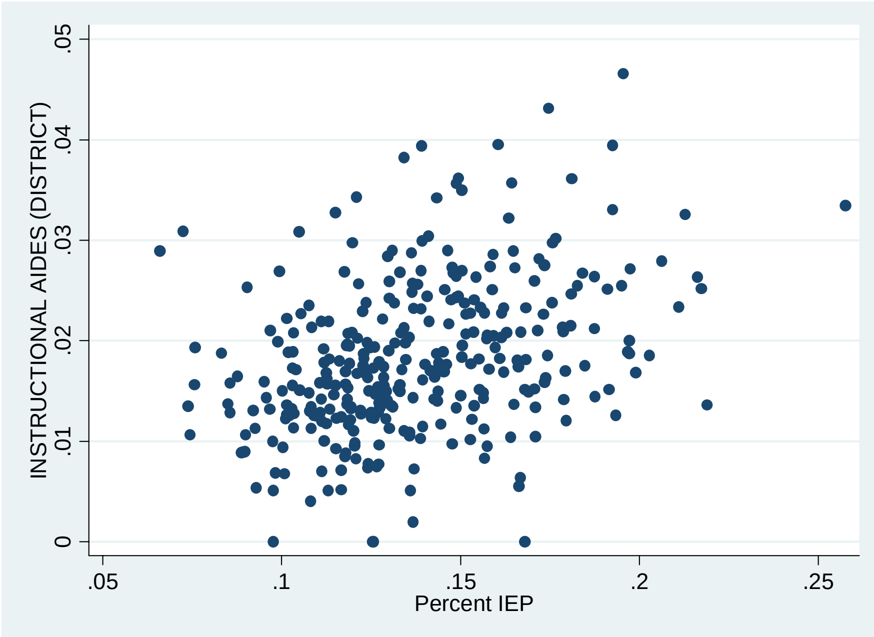

Figure 6. Relationship between percentage of individual education program students and the instructional aides per student ratio.

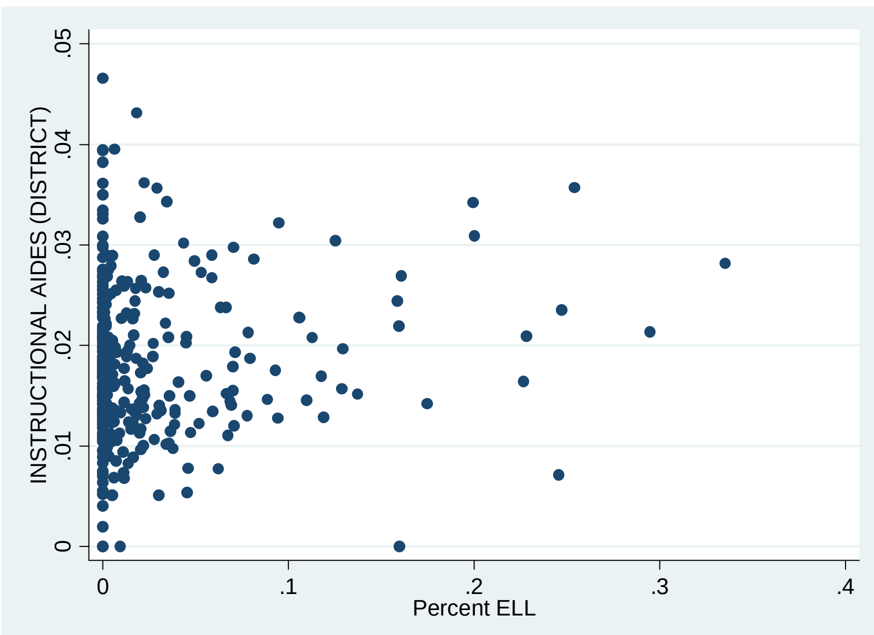

Figure 7. Relationship between percentage of limited English proficiency students and the instructional aides per student ratio.

References

Cameron, Stephen V. and Heckman, James J. "The Nonequivalence of High School Equivalents." *Journal of Labor Economics*, 1993, *11*(1), pp. 1-47.

Duncombe, William; Lukemeyer, Anna and Yinger, John. "Financing an Adequate Education: A Case Study of New York," *Discussion paper prepared for the National Center for Education Statistics.* 2003.

Faville, Brett. "Does More Money Man Better Test Scores? Bridging the Gap between Cost Function Adequacy Estimations and Educational Production Functions," *Department of Economics.* Stanford, CA: Stanford University, 2006, 59.

Fitzpatrick, Kevin M. and Yoels, William C. "Policy, School Structure, and Sociodemographic Effects on Statewide High School Dropout Rates." *Sociology of Education*, 1992, *65*(1), pp. 76-93.

Fowler, William J., Jr. and Walberg, Herbert J. "School Size, Characteristics, and Outcomes." *Educational Evaluation and Policy Analysis*, 1991, *13*(2), pp. 189-202.

Fuchs, Victor R. "Economics, Values, and Health Care Reform." *The American Economic Review*, 1996, *86*(1), pp. 1-24.

Goldschmidt, Pete and Wang, Jia. "When Can Schools Affect Dropout Behavior? A Longitudinal Multilevel Analysis." *American Educational Research Journal*, 1999, *36*(4), pp. 715-38.

Hanushek, Eric A. "The Impact of Differential Expenditures on School Performance." *Educational Researcher*, 1989, *18*(4), pp. 45-62.

Hedges, Larry V.; Laine, Richard D. and Greenwald, Rob. "An Exchange: Part I: Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes." *Educational Researcher*, 1994, *23*(3), pp. 5-14.

Klenzl, Gregory and Kena, Grace. "Economic Outcomes of High School Completers and Noncompleters 8 Years Later," National Center for Educational Statistics, 2006.

Krueger, Alan B. "Experimental Estimates of Education Production Functions." *The Quarterly Journal of Economics*, 1999, *114*(2), pp. 497-532.

Lee, Jong-Wha and Barro, Robert J. "Schooling Quality in a Cross-Section of Countries." *Economica*, 2001, *68*(272), pp. 465-88.

McNeal Jr., Ralph B. "High School Dropouts: A Closer Examination of School Effects." *Social Science Quarterly*, 1997, *78*(1), pp. 209-22.

Pittman, Robert B. and Haughwout, Perri. "Influence of High School Size on Dropout Rate." *Educational Evaluation and Policy Analysis*, 1987, *9*(4), pp. 337-43.

Rumberger, Russell W. "Dropping out of High School: The Influence of Race, Sex, and Family Background." *American Educational Research Journal*, 1983, *20*(2), pp. 199-220.

____. "Dropping out of Middle School: A Multilevel Analysis of Students and Schools." *American Educational Research Journal*, 1995, *32*(3), pp. 583-625.

Rumberger, Russell W. and Larson, Katherine A. "Student Mobility and the Increased Risk of High School Dropout." *American Journal of Education*, 1998, *107*(1), pp. 1-35.

Rumberger, Russell W. and Thomas, Scott L. "The Distribution of Dropout and Turnover Rates among Urban and Suburban High Schools." *Sociology of Education*, 2000, *73*(1), pp. 39-67.

Wenglinsky, Harold. "School District Expenditures, School Resources and Student Achievement: Modeling the Production Function," W. J. Fowler, Jr., *Developments in school finance : fiscal proceedings from the Annual NCES State Data Conference* Washington, D.C.: National Center for Educational Statistics, 1997.