High School Dropouts and Their Classmates

In the education literature, scholars and policy makers have demonstrated that the school a student attends plays a non-negligible role in shaping that student's academic outcomes. A considerable amount of effort has been put into showing that (and how) school-based factors such as the curriculum, teachers' characteristics, school culture or climate, etc. influence students' academic aspirations and achievement. In this paper, I focus on yet another component of the "school effects" literature – the social composition of a school – to answer the following question: Do high school dropouts influence the academic outcomes of their schoolmates? In short, I find that they do. More specifically, using data from the Educational Longitudinal Study of 2002, I show that attending a high school with a higher proportion of dropouts decreases the odds of on-time (high school) graduation and college enrollment for students who earn their diplomas. This finding suggests that there are social reasons we ought to remain concerned with dropout rates.

In order to contextualize my research question and findings, I've organized this working paper as follows. I begin by reviewing the literature and presenting the theoretical framework that is guiding this research project. Then, I summarize the dataset and the statistical methods I use to empirically answer this research question. I conclude by analyzing my main findings and presenting an overview of the next steps I will take to improve this paper.

Literature Review:

There are two strands of literature that have informed this paper: the body of work that links school composition to student achievement and the vast amount of research on understanding high school dropouts.

The research on school composition has, thus far, focused on understanding the relationship between the racial, sex, and socioeconomic composition of schools and student achievement. This body of work posits that – net of individual characteristics and other school resources – the demographic characteristics of a student's classmates influence his/her academic aspirations or achievement. The literature on the effect of racial composition on student achievement, much of which is related to the work on racial segregation, generally indicates that attending a school with high(er) levels of minority students decreases student achievement although the mechanisms are still unclear (Logan et al., 2012). Some research further indicates that minority concentration might not influence all students equally, with high-ability black students at greatest risk of suffering academically (Hanushek et al., 2002).

The work on the relationship between gender and academic achievement has focused on the effects of attending single-sex schools. Researchers have found that attending a single-sex school improves academic achievement for both, boys and girls (Lee and Bryk, 1986; Park et al., 2013) but it's unclear if one gender has an advantage over the other. Finally, and perhaps unsurprisingly, the literature on the socioeconomic composition of schools demonstrates that attending a school where fewer students qualify for free- or reduced-price meals has a positive influence on

students' achievement growth and literacy during their high school years (Rumberger and Palardy, 2005; Willms, 2010). Unlike racial composition, socioeconomic composition seems to influence students of all races equally (Rumberger and Palardy, 2005).

Another strand of the literature on school composition seeks to understand if, net of demographic characteristics, the abilities of peers and classmates influence a student's academic outcomes. The work in this field suggests that there is, indeed, a positive relationship between the two: As the "quality" of peers – measured by test scores – increases, so does student achievement (Ding and Lehrer, 2007; Hanushek et al., 2003). That being said, there is still some disagreement among scholars about the relationship between the variance of peers' scores and individual student achievement as well as whether all students benefit equally from their peers' abilities and achievements (Ding and Lehrer, 2007; Hanushek et al., 2003). In addition, some scholars have found that peers' attitudes and wishes are related to academic decisions – particularly college enrollment and attendance – for low-income, urban, and minority youth (Sokatch, 2006).

The second strand of literature that has informed this research project has been on high school dropouts. So far, this body of work has primarily focused on better understanding the causes and consequences of dropping out for those who drop out. While the literature on high school dropouts is too vast to effectively summarize here, the most relevant component of this literature to the development of this paper has been the work on dropouts and their peers. More specifically, researchers have shown that the number, characteristics, and closeness of friends

that a student has are important predictors of whether they drop out of high school. For instance, students with a larger number of close friends (Carbonaro and Workman, 2013) as well as students who are more engaged with other students who believe in the legitimacy of school are less likely to drop out (Ream and Rumberger, 2008).

In this paper, I hope to build on the two bodies of literature I've summarized above and to better understand our knowledge of the effects of school composition as well as the consequences of dropping out. In particular, I'd like to explore if there is a relationship between the proportion of dropouts in a school and the academic achievements of the students who persist in high school.

Theoretical Framework:

Although "school climate" and the tone set by teachers and administrators are important in shaping student achievement, in this paper I rely on a theoretical framework that emphasizes the importance of peer interactions and exchanges in schools. Put differently, I argue that the social composition of a school influences individual student outcomes because students convey to each other both, the *norms* for appropriate behavior as well as the requisite *information* to succeed in school.

As previous researchers have suggested, in a school setting, students look to one another to better understand how to behave. They implicitly establish norms – through interpersonal exchanges – for a wide range of academic and social behaviors: the appropriate amount of effort to put into learning, guidelines for classroom and hallway behavior, how to interact with the adults in the building, etc. Past researchers have identified this behavior and defined this process in different

ways – Gamoran (1992) views it as a social psychological process and Ryan (2000) refers to it as "peer pressure" and argues that it can be both, subtle or direct. In addition to setting norms collectively, students share information and skills with one another. This process can involve a direct "information exchange" in a conversation or can occur over time through observation and imitation of each other's behaviors (Ryan, 2000). As Ream and Rumberger (2008) suggest, the exchange of norms and information can be positive or negative – for a school as a whole or for individual students.

The question that remains to be answered, then, is whether high school dropouts might influence the norms and information at a high school. I argue that dropouts are able to influence the norms and information at their high schools through a combination of two processes. First, the research indicates that that dropouts are different from non-dropouts. In particular, they tend to have less engagement with school: they are more likely to have a history of truancy or absenteeism, are more likely to have been retained in a grade, and have weaker levels of school performance and engagement (Rumberger and Thomas, 2000; Ream and Rumberger, 2008). Second, dropping out is almost always related to student mobility. In other words, students who drop out either have a history with absenteeism and truancy as noted above or have made numerous residential or school-based moves (Rumberger, 2011). The research demonstrates that such student mobility has a negative influence on the students who move, particularly if the move was unscheduled (Stearns and Glennie, 2006), as well as the students who remain in school (Fiel et al., 2013).

To synthesize the above, I am arguing that the social composition of a school influences a student's academic outcomes because students are, as Ryan (2000) suggests, *socialized* by their peers to have a certain level of engagement, motivation, and achievement in school. As such, in a school with a higher proportion of dropouts, there is less value on behaviors that will lead to academic success. Similarly, in a school with a higher proportion of dropouts, not only will there be fewer academically inclined students who can share information about successful learning practices or behaviors, it is also likely that the information exchanges occurring between students will be focused on non-school related topics such as employment opportunities or family responsibilities.

Although some researchers might argue that selection into friendship networks and friendship homophily will lead to students being exposed to varying kinds of norms and information, empirical evidence suggests otherwise. A recent paper by Carbonaro and Workman (2013) finds, for instance, that "distant" friends – those individuals with whom students don't interact frequently – are more important in shaping their understanding of norms. In addition, Fletcher and Tienda (2009) show that entering college with a larger group of friends – who, presumably share the different norms and/or types of information – perform better in college. Membership in a cohort, school or broad social network, then, seems to confer on students some sense of how to behavior and information about shared values and goals. Perhaps this is the reason why scholars find evidence of "peer effects" even without examining the friendship groups to which students belong (e.g.: Ding and Lehrer, 2007).

Hypothesis:

Based on the empirical evidence from the literature on social composition as well as the theoretical framework I use, I hypothesize that students who attend schools with higher proportions of dropouts will have worse academic outcomes than otherwise equivalent students who attend schools with fewer proportions of dropouts. While some might contend that the process of selection makes my hypothesis untenable (as the "worse" students drop out, those who remain are better able to pursue their academic ambitions), I believe that the students who drop out have already shaped the norms and resources that persisting students encounter for the reasons I've presented in the previous section.

Data and Methods:

To understand the relationship between high school dropouts and the students who go on to graduate from high school, I use the Educational Longitudinal Study of 2002 (ELS). This longitudinal survey, administered by the National Center for Educational Statistics, contains information on the academic and social lives of high schoolers and follows them for eight years after their expected high school graduation. I use information from the first two waves of this longitudinal study, drawing information from student, parent, and/or administrator interviews during the spring of the students' sophomore and senior years of high school.

There are two distinct advantages to using ELS to understand the relationship between dropouts. First, the survey has relatively low attrition (Domina et al., 2010). Attrition is always of concern to researchers who study dropouts because, as I mention above, they are a more mobile population. Second,

ELS gathers information about a student's dropout status from the student as well as the school they last attended. We can be confident, then, that dropout status isn't mis-reported as both, schools and family members of the dropout are liable to do (Swanson, 2004; Domina et al., 2010). Moreover, ELS offers collects a rich array of social and contextual variables that enable me to account for each respondent's home and school life.

Empirical Strategy:

As I have operationalized both of my dependent variables as binary outcomes (explained below), I employ logistic regressions to understand the relationship between the proportion of dropouts in a school and the academic trajectories of students who do not dropout. Following Carbonaro and Workman (2013), I account for the sampling strategy of ELS (where a random sample of students was drawn from a nationally representative sample of schools) by employing the "cluster" command in STATA, ensuring accurate estimates for the standard errors.[1]

Variables:

To calculate the key independent variable, the proportion of dropouts at each of the 750 schools in this dataset, I follow Ream and Rumberger (2008): Any enrolled tenth-grade student who is not enrolled in school in the twelfth-grade is considered a dropout.[2] This group of students includes those who are currently

[1] I explain other empirical strategies that I plan to pursue in the conclusion of this working paper.
[2] For the analysis presented here, I only include respondents who are not missing data on their schooling status in the baseline survey (sophomore year) and the first follow-up (senior year). I find that students who are missing data are more likely to share the same characteristics as those who identify as dropouts (in the literature on dropouts and the known dropouts in this dataset). As a result, I argue that the estimates I present in this working paper are conservative.

enrolled in or who have completed a GED program, those who identify as dropouts, as well as students whose status is unknown (but have not re-enrolled as regular students at another school). I have chosen to categorize students who earn GEDs as "dropouts" because No Child Left Behind regulations stipulate that only those who earn regular high school diplomas can be considered high school graduates (Swanson, 2004). Moreover, students who enroll in GED programs are generally not in the same physical space as students who persist in high school and, consequently, cannot shape the norms and information available to other students.

An important advantage of this operationalization of the independent variable is that it avoids the "reflection problem" (Ding and Lehrer, 2007) that occurs when the independent and dependent variables are measured at the same time. The independent variable, in this case, reflects the school context students experienced between their sophomore and senior years of high school. The dependent variables, as I explain further below, are measures from the end of high school. One limitation of this operationalization – and any others that draw on ELS data – is that I miss any students who drop out prior to their tenth-grade year. This might be problematic as some cross-sectional analyses indicate that ninth-grade boys have the highest dropout rate of all high school students (Stearns and Glennie, 2006).

In this working paper, I've chosen to examine the relationship between the proportion of dropouts at a high school and two key outcome variables: on-time high school graduation and students' plans to enroll in college immediately after high school completion. The first variable, on-time graduation, indicates the extent

to which the timing of an important educational milestone is related to the social composition of the school a student attended. In my current operationalization, students who graduated with regular high school diplomas in the early summer (May, June, or July) of their senior year – two years after their sophomore year – are considered "on-time graduates".

The second dependent variable – college enrollment - captures students' abilities to fulfill their aspirations and plans for post-secondary education. In my current operationalization, students are considered college enrollees if they planned to *and* were able to successfully enroll in a post-secondary institution by the fall after their high school graduation. This variable indicates whether students' abilities to fulfill their ambitions are related to the social composition of their high schools. In both cases, I expect that the norms and information to which students are exposed influence their ability to complete school on time and their desires to pursue a post-secondary education.

In addition to the independent and dependent variables I've described here, all of my regression analyses control for key individual- and school-based covariates[3]. For a complete list of these variables and how they were constructed, please refer to Appendix A.

Results and Analyses:

Taken together, a descriptive and multivariate analysis of the data point to two key results. First, I find evidence to support my hypothesis. That is, as the

[3] In future versions of this paper, I plan to add more contextual or school-based variables that might be correlated with the outcomes. These include school racial composition, school size, school curriculum (e.g.: tracking) and school quality.

proportion of dropouts in a school increases, the odds of on-time high school graduation and college enrollment decrease. Second, the analyses suggest that the influence of school composition – as measured by the proportion of dropouts in a school – operates *in addition* to the other school effects described in the literature review, particularly socioeconomic status. Put differently, the social composition of a school has an independent relationship with the outcome variables I've measured here.

Descriptive Results:

The descriptive results, presented in Table 1, provide some background information on how students are sorted into schools. Based on my operationalization of the key independent variable – the proportion of dropouts at a school – I've categorized schools as being "low", "medium", or "high" dropout schools. As the literature on school choice and residential segregation suggests, there are greater percentages of minority and low-income students at high-dropout schools relative to low- and medium-dropout schools. In addition, students who attend high-dropout schools tend to be, on average, slightly older, and have lower grade point averages (at the end of their tenth-grade year).

In terms of school characteristics, high-dropout schools are more likely to be public (as opposed to private) and located in urban areas. Additionally, students who attend high-dropout schools are more likely than their peers at medium- and low-dropout schools to have classmates who qualify for free- or reduced-price meals. Perhaps unsurprisingly, the descriptive data also show a correlation between the independent and dependent variables of interest in this paper: high-dropout

schools have the lowest rates of on-time graduates and the percentage of college enrollees. I run and present results from multivariate regressions to demonstrate that these correlations are not merely an artifact of the selection processes that sort students into schools.

Multivariate Results:

The first set of multivariate results, as shown in Table 2, highlight the relationship between the proportion of dropouts in a school and on-time graduation (among the students who do persist). The full model – which controls for individual and school characteristics – is presented in the third column. The coefficient for school dropout indicates that a one percent increase in the number of dropouts changes the odds of on-time high school graduation 0.1265 times. In other words, a one percent increase in dropouts at a school decreases the odds of on-time graduation by roughly 87 percent.[4] This result supports my hypothesis that an increasing proportion of dropouts influence the academic outcomes of all students, perhaps because of the manner in which or the substance of the norms and information exchanged among classmates.

The results also indicate that other school composition variables are associated with on-time high school graduation. In particular, attending a public school decreases the odds of on-time graduation by about 45 percent. Additionally, the results from the multivariate analysis suggest that attending a poorer school – a

[4] This result – while in the right direction – seems extremely large (in magnitude). One reason that this might be happening is because my key independent variable is skewed (please see Appendix B for a histogram). In subsequent drafts of this paper, I'm hoping to construct a school dropout rate using alternative measures of high school completion and to re-create my independent variable using the Common Core of Data (CCD) dropout rates. I hypothesize that while the direction and significance of alternative operationalizations of my independent variable will remain the same, it will decrease in magnitude.

one-unit increase in school poverty – decreases the odds of on-time graduation by 6 percent. Together, these findings raise the empirical question of whether the information exchange or norm-setting processes at private schools and/or schools with richer student populations are somehow different than those at private schools.[5]

The second set of regression results, presented in Table 3, demonstrate the relationship between the proportion of dropouts at a school and college enrollment. The substantive results are similar to those for on-time graduation. The full model – which is, again, presented in the third column and includes individual- and school-based covariates, demonstrates that a one percent increase in the number of dropouts changes the odds of on-time high school graduation 0.2295 times. This means that the odds of enrolling in college decreases by approximately 77 percent for every one percent increase in dropouts at a school. While the results from Table 2 indicate that attending a school with a larger proportion of dropouts influences the timing of key educational outcomes, these results demonstrate that school composition also influences students' capacities to translate their ambitions into actions. As above, I speculate that this result is related to the norms established and information exchanged within schools.

Finally, as with on-time graduation, associations between other measures of school composition and college enrollment remain significant. Attending a public school decreases the odds of college enrollment by roughly 45 percent. Moreover, a

[5] In subsequent drafts of this paper, I'd like to align my analysis more closely with the literature I reviewed on school composition. In particular, I'd like to see if there is an influence of racial and socioeconomic composition on academic outcomes in addition to the proportion of dropouts.

one-unit increase in school poverty, measured by the percentage of students eligible for free- or reduce-price meals changes the odds of college enrollment by 0.935. This translates into a 6.5 percent decrease in the odds of college enrollment. Taken together, these findings suggest that school composition, when measured by examining the number of dropouts, has a different and independent influence on students' outcomes than other school composition variables, such as the socioeconomic composition of a school. Why this might be the case remains an empirical question but the underlying processes driving these outcomes could, potentially be the same.

The results I've presented in this section indicate that the low academic performance of *some* students is negatively associated with academic outcomes for *all* students. The implication of these findings is that we ought to be concerned with the consequences of dropping out on more than the individual dropouts themselves. In fact, as Gamoran (1992) argues, the low productivity of some students creates the potential for greater inequality across schools. This happens because high school dropouts negatively influence even the students who are "on-track", holding back the higher achieving students at the schools with the highest levels of dropping out.

Next Steps:

In order to improve this working paper, I plan to complete the following larger-scale tasks (in addition to attending to smaller concerns like increasing the number of contextual variables, dealing with missing data, thinking about issues of selection that might plague this analysis, etc.):

1. I'd like to better develop the theoretical framework I've presented in this paper and link it more clearly to the dependent variables I choose. I'd also like to use the theoretical framework to *explain* my results.

2. Re-think my empirical strategy: There is some evidence, in the literature, that it might be more appropriate to incorporate a multi-level modeling strategy and/or use structural equation modeling.

3. I'd like to repeat the analysis – once I find the appropriate strategy – to understand if proportion of dropouts influences all students equally (I suspect not). For instance, are the highest-achieving students in a school more protected than those who are in the middle?

Table 1: Descriptive Statistics by School:

	Low Dropout Schools N= 385	Medium Dropout Schools N = 164	High Dropout Schools N = 202
Characteristics of Students:			
Percent Male	49.6	50.1	49.3
Percent by Race:			
Asian	9.9	9.6	7.1*
Black	8.6	14.4	21*
Hispanic	11.3	15.2	19.5*
White	65	55	45.4*
Other	5.2	5.8	7*
Age in Years	16.39	16.45	16.56*
Average Family SES Quartile[1]	2.86	2.43	2.2*
Average GPA	2.9	2.7	2.6*
Characteristics of Schools:			
Percent Dropout	0.78	7.3	16.94*
Percent On-Time Graduates	92.2	82.8	72.7*
Percent College Enrollees	67.3	49.7	40.3*
Public School	62.4	92.9	95.7*
Private School	37.6	7.1	4.3*
Urban	33.5	27	37.6*
Suburban	49.9	51.9	41.7
Rural	16.5	21.1	20.7
Percent of Free Lunch Students	6-10	11-20	21-30

Source: Educational Longitudinal Study of 2002 (Baseline Survey and First Follow-Up Survey)
[1]Families are placed in one of four quartiles based on their socioeconomic status (see Appendix A for more details). Lowest-SES families are given "1" and highest-SES families are given a "4"
* Significant ($p < 0.05$) results from t-tests comparing high-dropout schools to low- and medium-dropout schools (combined)

Table 2: Odds Ratios for On-Time High School Graduation:

	On-Time Graduation[1]	On-Time Graduation[1]	On-Time Graduation[1]
School Dropout	0.0063***		0.1265**
Individual Characteristics:			
Male [Female]		0.9466	0.937
Race [White]			
Black		1.069	1.08
Hispanic		0.7335**	0.731*
Asian		0.913	0.893
Other		0.6006**	0.607**
Family SES [Rich]			
Poor		0.664**	0.669**
Middle		0.810†	0.8215†
Upper middle		0.874	0.88
GPA		1.77***	1.77***
School Characteristics:			
Public [Private]		0.501**	0.553**
Urbanicity [Suburban]			
Urban		0.847	0.893
Rural		1.01	1.006
Percent of Free Lunch Students		0.922	0.94†
Constant	12.574***	2.889***	2.857***
Observations	13,073	11,004	11,004
Pseudo R^2	0.0238	0.1573	0.16

Source: Educational Longitudinal Study of 2002 (Baseline Survey and First Follow-Up Survey)
Note: *** p<0.001, ** p<0.01, *p<0.05, †p<0.10
[1]Results displayed in this column are from logistic regression with clustered standard errors (by school)

Table 3: Odds Ratios for College Enrollment:

	College Enrollment[1]	College Enrollment[1]	College Enrollment[1]
School Dropout	0.00683***		0.2295***
Individual Characteristics:			
Male [Female]		0.7096***	0.7071***
Race [White]			
Black		1.32**	1.339***
Hispanic		0.8793†	0.881†
Asian		1.62***	`.6***
Other		0.7358**	0.7434**
Family SES [Rich]			
Poor		0.3367***	0.339***
Middle		0.4672***	0.471***
Upper middle		0.6569***	0.661***
GPA		1.7***	1.7***
School Characteristics:			
Public [Private]		0.514***	0.548***
Urbanicity [Suburban]			
Urban		1.1155	1.145†
Rural		1.0316	1.028
Percent of Free Lunch Students		0.9235***	0.935***
Constant	2.006***	0.4453***	0.441***
Observations	13,073	11,004	11,004
Pseudo R^2	0.022	0.1941	0.1953

Source: Educational Longitudinal Study of 2002 (Baseline Survey and First Follow-Up Survey)
Note: *** p<0.001, ** p<0.01, *p<0.05, †p<0.10
[1]Results displayed in this column are from logistic regression with clustered standard errors (by school)

References

Carbonaro, W. and J. Workman. 2013. "Dropping Out of High School: Effects of Close and Distant Friendships." *Social Science Research* 42: 1254-1268.

Ding, W. and S. Lehrer. 2007. "Do Peers Affect Student Achievement in China's Secondary Schools?" *The Review of Economics and Statistics* 89(2): 300-312.

Domina, T., Ghosh-Dastidar, B. and M. Tienda. 2010. "Students Left Behind: Measuring 10th to 12th Grade Student Persistence Rates in Texas High Schools. *Educational Evaluation and Policy Analysis* 32(2): 324-346.

Fiel, J., Haskins, A., and R. Turley. 2013. "Reducing School Mobility: A Randomized Trial of a Relationship-Building Intervention." *American Educational Research Journal* 50: 1188-1218.

Fletcher, J. and M. Tienda. 2009. "High School Classmates and College Success." *Sociology of Education* 82: 287-314.

Gamoran, A. 1992. "Social Factors in Education." Pp. 1222-1229 in *Encyclopedia of Educational Research*, ed. Marvin Alkin. New York, NY: Macmillian.

Hanushek, E., Kain, J., Markman, J. and S. Rivkin. 2003. "Does Peer Ability Affect Student Achievement?" *Journal of Applied Econometrics* 18(5): 527-544.

Hanushek, E., Kain, J., and S. Rivkin. 2002. "New Evidence about Brown v. Board of Education: The Complex Effects of School Racial Composition on Achievement." NBER Working Paper No. 8741.

Lee, V. and A. Bryk. 1986. "Effects of Single-Sex Secondary Schools on Student Achievement and Attitudes." *Journal of Educational Psychology* 78(5): 381-395.

Logan, J., Minca, E., and S. Adar. 2012. "The Geography of Inequality: Why Separate Means Unequal in American Public Schools." *Sociology of Education*,85(3): 287-301.

Park, H., Behrman, J., and J. Choi. 2013. "Causal Effects of Single-Sex Schools on Entrance Exams and College Attendance: Random Assignment in Seoul High Schools." *Demography* 50: 447-469.

Ream, R. and R. Rumberger. 2008. "Student Engagement, Peer Social Capital, and School Dropout anong Mexican American and Non-Latino White Students." *Sociology of Education* 81(2): 109-139.

Rumberger, R. 2011. *Dropping Out: Why Students Drop Out of High School and What Can Be Done About It*. Cambridge, MA: Harvard University Press.

Rumberger, R. and G. Palardy. 2005. "Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School." *Teachers College Record* 107(9): 1999-2045.

Rumberger, R. and S. Thomas. 2000 "The Distribution of Dropout and Turnover Rates Among Urban and Suburban High Schools." *Sociology of Education* 73(1): 39-67.

Ryan, A. 2000. "Peer Groups as a Context for the Socialization of Adolescents' Motivation, Engagement, and Achievement in School." *Educational Psychologist* 35(2): 101-111.

Sokatch, A. 2006. "Peer Influences on College-Going Decisions of Low Socioeconomic Status and Urban Youth. *Education and Urban Society* 39(1): 128-146.

Stearns, E. and E. Glennie. 2006. "When and Why Dropouts Leave High School." *Youth and Society* 38(1): 29-57.

Swanson, C. 2004. "Who Graduates? Who Doesn't? A Statistical Portrait of Public High School Graduation, Class of 2001". Washington, DC: The Urban Institute Education Policy Center.

Willms, J. 2010. "School Composition and Contextual Effects on Student Outcomes." *Teachers College Record* 112(4): 1008-1037.

Appendix A
Variable List and Description

Variable Name in Tables	ELS ID for Variable	Description or Notes
Individual Characteristics:		
Male	bysex	Re-coded bysex such that male students were coded as "1" and females as "0"
Race – Other	BYRACE_R	Created a new variable, "Other" which was "1"if BYRACE_R was coded as 1, 6, or 7
Race – Asian	BYRACE_R	Created a new variable, "Asian" which was "1" if BYRACE_R was coded as 2
Race – Black	BYRACE_R	Created a new variable, "Black" which was "1" if BYRACE_R was coded as 3
Race – Hispanic	BYRACE_R	Created a new variable, "Hispanic" which was "1" if BYRACE_R was coded as 4 or 5
Race – White	BYRACE_R	Created a new variable, "White" which was "1" if BYRACE_R was coded as 8
Age	BYDOB_P	Created a new variable, "AgeG10" which subtracted the year of birth (taken from BYDOB_P) from 2002 to calculate age in 10th grade
Family SES1/2/3	BYSES2QU	Used as-is; a composite variable based on parents' education, occupation, and total family income. The variable is divided into four equally-sized quartiles.
Grades	F1RGP9 and F1RGP9P	Both were used as is; the first variable (used in the descriptive tables) just provides raw GPA (as calculated by the NCES staff); the second variable (used in regression analyses) divides students' GPAs into seven categories
School Characteristics:		
Public	bysctrl	Created a new variable, "Public" which was "1" if bysctrl was coded as 1
Urbanicty - Urban	byurban	Created a new variable, "Urban"

		which was “1” if byurban was coded as 1
Urbanicity – Rural	byurban	Created a new variable, “Rural” which was “1” if byurban was coded as 3
Urbanicity - Suburban	byurban	Created a new variable, “Suburban” which was “1” if byurban was coded as 2
Percent Free Lunch	BY10FLP	Renamed BY10FLP “FLCatG10”; the variable provides seven categories of the percent of free-lunch eligible students by school

Note: For all of the above variables, I just re-coded missing data so STATA sees a . (instead of the NCES coding)

Appendix B
Histogram of School Dropout Rates