MEMORANDUM

SUBJECT: Identification of Schools within 200 meters of U.S. Primary and Secondary Roads

FROM: Meredith Pedde, Office of Transportation and Air Quality/Assessment and Standards Division
Chad Bailey, Office of Transportation and Air Quality/Assessment and Standards Division

TO: Docket EPA-HQ-OAR-2011-0135

Introduction

The purpose of this memo is to describe the approach used in EPA's analysis of students and schools in close proximity to primary and secondary roads in the United States, which is presented in section II.B.7 of the preamble to the Notice of Proposed Rulemaking (NPRM). The data used to create Tables 1-4 in this memo are provided as an attachment to this memo.

Methodology

EPA obtained public school data from the U.S. Department of Education's Institute of Education Sciences National Center for Education Statistics (NCES).[1] At the time this analysis was done, the most recent public school data available was for the 2009 – 2010 school year. These are the data used in our analysis. For each U.S. school (K-12), the data include the latitude and longitude, a unique NCES assigned ID, total number of students, number of students by race, and the number of students eligible for the free and reduced school lunch programs.

EPA obtained and used the U.S. Census Bureau's 2010 TIGER/Line® Shapefiles for primary and secondary roads. According to the technical documentation for the 2010 Census TIGER/Line® Shapefiles, "Primary roads are generally divided, limited-access highways within the Federal interstate highway system or under state management. These highways are distinguished by the presence of interchanges and are accessible by ramps, and may include some toll highways." [2] The technical documentation defines secondary roads as "main arteries, usually in the U.S. Highway, State Highway, or County Highway system. These roads have one or more lanes of traffic in each direction, may or may not be divided, and usually have at-grade intersections with many other roads and driveways. They often have both a local name and a route number."[3]

EPA imported the school data into ArcMap 10.0 as a set of geocoded points using the latitude and longitude provided in the schools dataset. EPA also imported the primary and secondary road

[1] http://nces.ed.gov/ccd/bat/
[2] http://www.census.gov/geo/www/tiger/tgrshp2010/TGRSHP10SF1.pdf
[3] http://www.census.gov/geo/www/tiger/tgrshp2010/TGRSHP10SF1.pdf

shapefiles into ArcMap 10.0 and then created 200 meter buffers around those roads. The school point data were intersected with the 200 meter road buffers to determine which schools are within 200 meters of U.S. primary and secondary roads.[4] Of the 99,804 U.S. schools in the dataset, 23,978 (24%) were identified to be located within the 200 meter buffers of U.S. primary and secondary roads.[5] Of the 23,978 schools within 200 meters of a primary or secondary road, 1,649 are near primary roads and 22,329 are near secondary roads, representing 1.7% and 22.4% of total schools, respectively.

This method of determining which schools are within 200 meters of primary or secondary roads underestimates the total number of schools from the edges of roads since the road shapefiles are line data – not polygon data. The line data represents the center line of a road (i.e., the center line between both directions of traffic), whereas polygon data would represent the area the roads encompass (i.e., the edges of a road). In situations where there are six lanes of traffic in each direction, for example, approximately 40 meters of the 200 meter buffers could be the road itself, resulting in the intersection analysis only identifying schools that are within approximately 160 meters of the edge of the road.

Analysis of Racial Demographics

Using SAS 9.1, EPA evaluated the racial composition and Hispanic status of the 49,136,240 students at the 99,804 U.S. schools, by number and percent of total students. This data is presented in Table 1.

Table 1: Number and Percent of Students by Race and Hispanic Status at all U.S. Schools[6]

	American Indian/Alaskan Native	Asian/Pacific Islander	Black	Hispanic	White	Two or More Races	No Race Reported	Total
Number of Students	584,756	2,461,820	8,166,410	10,775,975	26,311,473	334,459	501,347	49,136,240

[4] This approach does not consider the physical footprints of the schools, but represents each school as a discrete point in space. As such, the intersection of school points with roadway buffers represents an indicator of proximity to roadways with a degree of measurement error. Two primary sources of measurement error can be present. First, it is unknown where on school property each latitude/longitude coordinate represents (e.g., front office, school center). Second, schools have area in space not represented by a point, which may lead to some underestimation of the number of schools near major roads (i.e., where one wing of a school is nearby, but the coordinates are not)."

[5] 182 of these 23,978 schools were identified as 'temporarily closed' and 97 were identified 'to be operational within two years.' Neither of these types of schools had any student enrollment information (i.e., these 279 schools had total students of 0), so there are no students from these types of schools included in the student racial or SES analysis.

[6] In the dataset used here, race designations refer to non-Hispanic students. Hispanics are enumerated separately.

Percent of Total	1.19%	5.01%	16.62%	21.93%	53.55%	0.68%	1.02%	100.00%

EPA also evaluated the racial composition and Hispanic status of students at only the 23,978 schools within the 200 meter buffers of U.S. primary and secondary roads. Table 2 shows the number of students, by race, at these schools, as well as the percent of students by race.

Table 2: Number and Percent of Students by Race and Hispanic Status at U.S. Schools within 200 meters of Primary and Secondary Roads, by Road Type

		American Indian/Alaskan Native	Asian/Pacific Islander	Black	Hispanic	White	Two or More Races	No Race Reported	Total
Primary Road	**Number of Students**	6,242	56,105	175,128	244,618	313,169	9,216	7,409	811,887
	Percent of Total	0.77%	6.91%	21.57%	30.13%	38.57%	1.14%	0.91%	100.00%
Secondary Road	**Number of Students**	121,373	278,209	1,535,625	1,213,743	5,943,786	33,899	92,338	9,218,973
	Percent of Total	1.32%	3.02%	16.66%	13.17%	64.47%	0.37%	1.00%	100.00%
Primary and Secondary Roads	**Number of Students**	127,615	334,314	1,710,753	1,458,361	6,256,955	43,115	99,747	10,030,860
	Percent of Total	1.27%	3.33%	17.05%	14.54%	62.38%	0.43%	0.99%	100.00%

The data in Tables 1 and 2 shows that minorities are overrepresented at schools located within 200 meters of primary and secondary roads, with the overrepresentation being largest among students at schools located within 200 meters of primary roads. The highlighted cells in Table 2 indicate that a race is overrepresented in the student population for either the set of schools within the 200 meter buffers of primary roads, secondary roads, or primary and secondary roads combined, as compared to

the racial distribution of students attending all of the 99,804 schools in the U.S.[7] For example, Black students represent 21.57% of students at schools located within 200 meters of a primary road, whereas Black students represent 16.62% of students in all U.S. schools. Hispanic students represent 30.13% of students at schools located within 200 meters of a primary road, whereas Hispanic students represent 21.93% of students in all U.S. schools.

Figures 1-3 presents cumulative distributions of the demographic compositions of student bodies schools near primary and secondary roads compared with all U.S. schools. Figure 1 depicts the cumulative (percentile) distributions of the percentage of nonwhite and Hispanic students for schools near primary roads, secondary roads, and all schools. Consistent with the tabular data presented in Tables 1 and 2, across the entire distribution of schools, those near primary roads have higher nonwhite student fractions than other schools.

Figure 1 – Cumulative Distribution of Student Nonwhite Percentages for Schools Near and Far from Major Roads

[7] The highlighted race proportions in Table 2 were tested against the corresponding race's proportion in Table 1 to determine if the difference in the race's proportion in the observed populations (i.e., the race distributions for the schools within 200 meters of primary, secondary, and primary and secondary roads combined) and the whole U.S. population of schools is statistically significant. Using a chi-square goodness of fit test, the differences in the highlighted percentages in Table 2 and the corresponding percentages in Table 1 were all statistically significant ($p < 0.0001$).

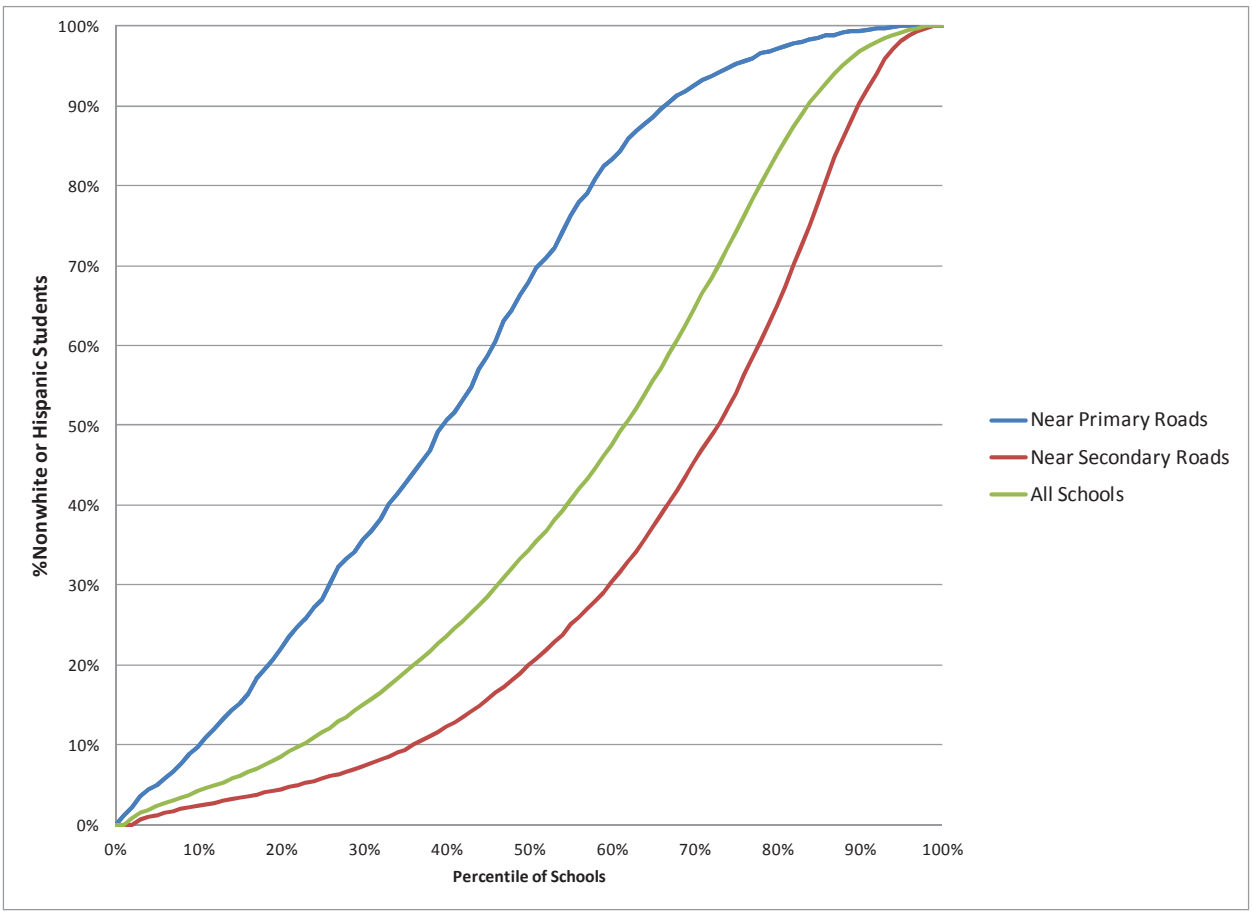
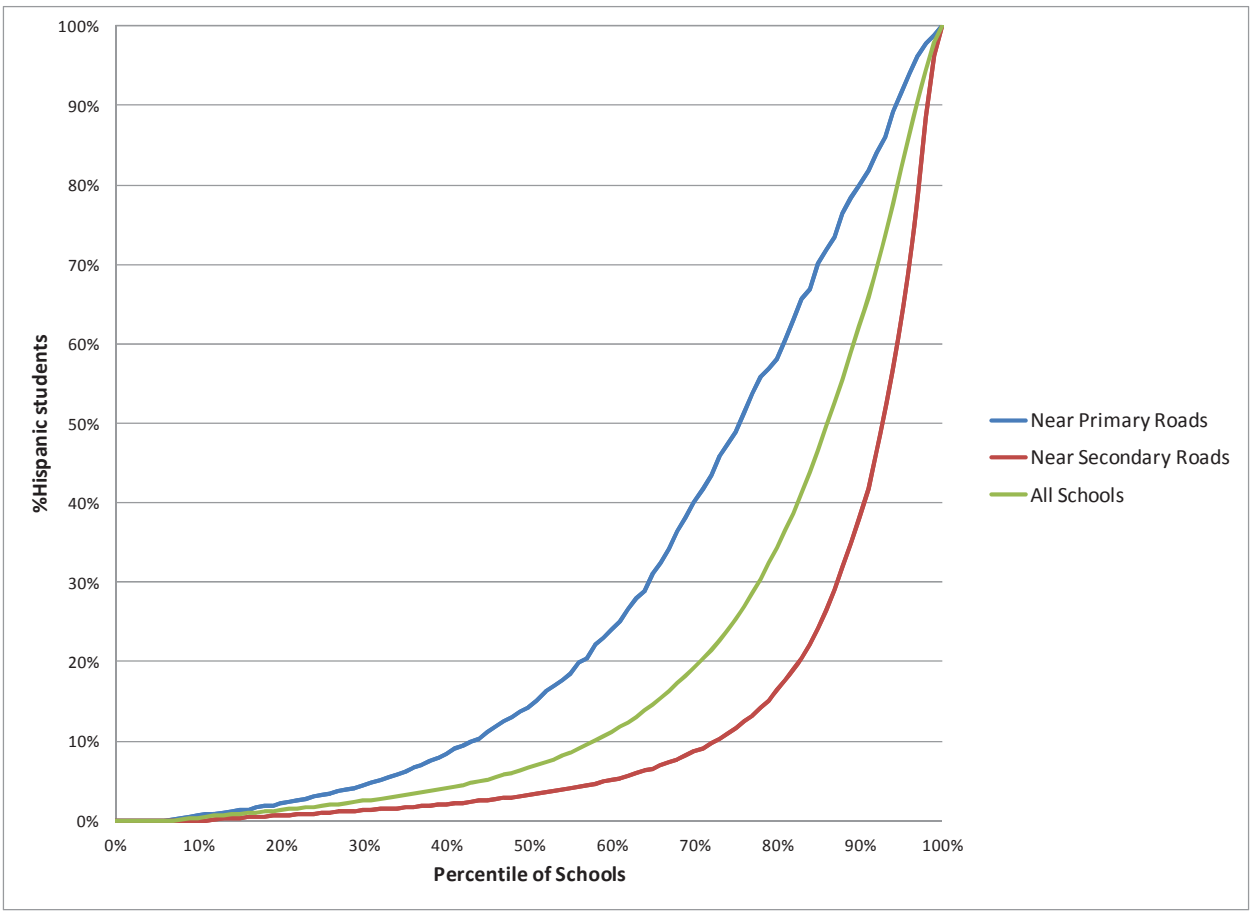


Figure 2 depicts the cumulative (percentile) distributions of the percentage of Hispanic students for schools near primary and secondary roads, and for all schools. Consistent with the tabular data presented in Tables 1 and 2, across the entire distribution of schools, those near primary roads have higher Hispanic student fractions than other schools.

Figure 2 – Cumulative Distribution of Student Hispanic Percentages for Schools Near and Far from Major Roads



Analysis of Free and Reduced-Price Lunch Eligibility

To evaluate the socioeconomic status of students who attend schools within 200 meters of primary and secondary roads, EPA used SAS 9.1 to evaluate the number of students who are eligible for the U.S. Department of Agriculture's free or reduced-price school lunch program. Children from families with incomes at or below 130 percent of the poverty level are eligible for free meals. Those with incomes between 130 percent and 185 percent of the poverty level are eligible for reduced-price meals.[8]

[8] United States Department of Agriculture: Food and Nutrition Service, National School Lunch Program Fact Sheet. Obtained from: http://www.fns.usda.gov/cnd/Lunch/AboutLunch/NSLPFactSheet.pdf, September, 2010. For the period July 1, 2010, through June 30, 2011, 130 percent of the poverty level is $28,665 for a family of four; 185 percent is $40,793.

Table 3 shows the number and percent of students that are eligible for free or reduced-price school lunches out of the 49,136,240 students at all 99,804 U.S. schools, by number and percent of total students.

Table 3: Number and Percent of Students Eligible for Free or Reduced-Price Lunch at all U.S. Schools

	Free Lunch Eligible	Reduced-Price Lunch Eligible	Total Free or Reduced-Price Lunch Eligible
Number of Students	18,684,568	3,738,839	22,423,407
Percent of Total	38.03%	7.61%	45.64%

EPA also evaluated the number of students eligible for free or reduced-price school lunches at only the 23,978 schools within the 200 meter buffers of U.S. primary and secondary roads. Table 4 shows the number and percent of students, by lunch program eligibility, at these schools, by road type.

Table 4: Number and Percent of Students Eligible for Free or Reduced-Price Lunch at U.S. Schools within 200 meters of Primary and Secondary Roads, by Road Type

		Free Lunch Eligible	Reduced-Price Lunch Eligible	Total Free or Reduced-Price Lunch Eligible
Primary Road	**Number of Students**	323,223	55,792	379,015
	Percent of Total	39.81%	6.87%	46.68%
Secondary Road	**Number of Students**	3,587,928	732,779	4,320,707
	Percent of Total	38.92%	7.95%	46.87%

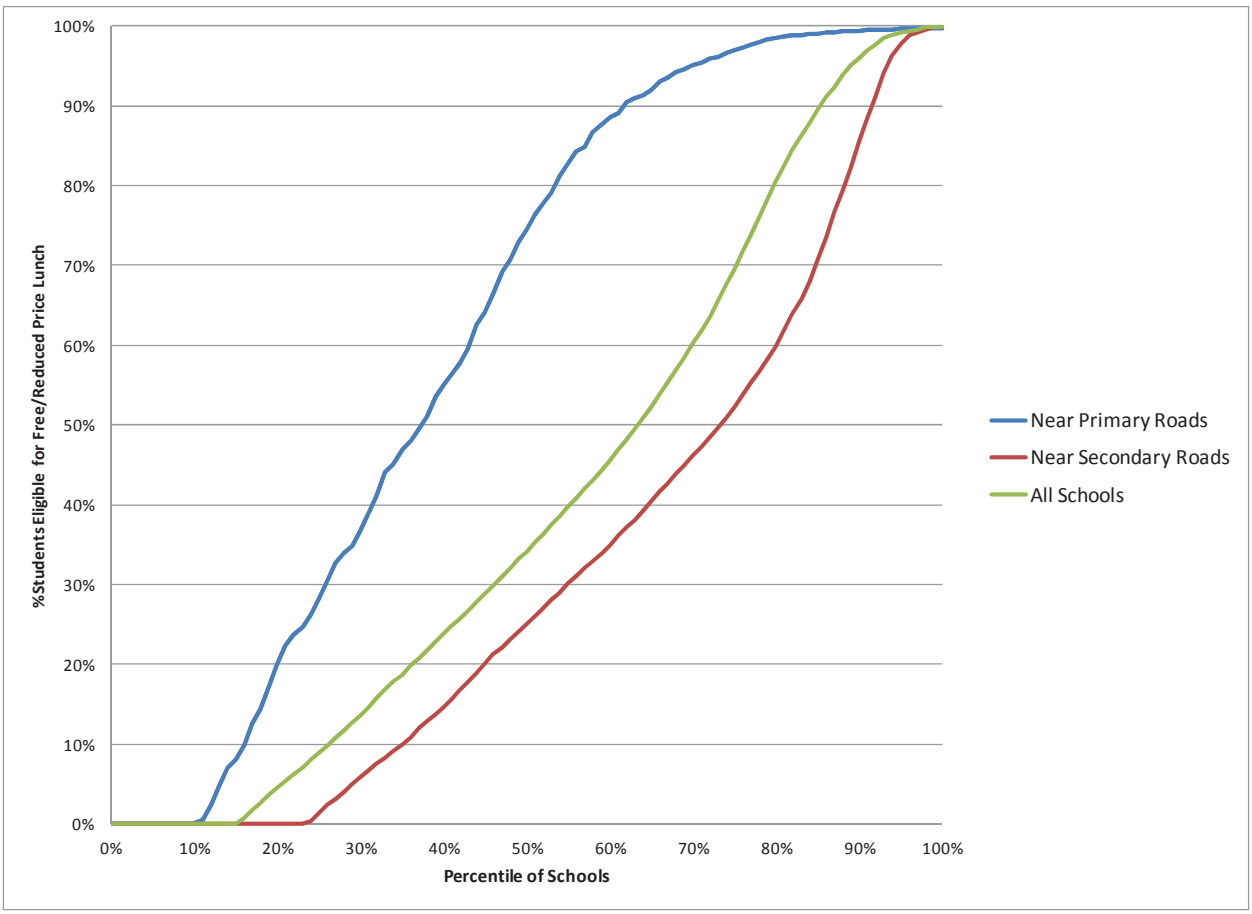
Primary and Secondary Roads	**Number of Students**	3,911,151	788,571	4,699,722
	Percent of Total	38.99%	7.86%	46.85%

The data in Tables 3 and 4 shows that students eligible for the free and reduced-price lunch program are overrepresented at schools located within 200 meters of primary and secondary roads. The highlighted cells in Table 4 indicate that free and/or reduced-price lunch eligibility is overrepresented in the student population for the sets of schools within the 200 meter buffers of primary roads, secondary roads, and primary and secondary roads combined, as compared to the free and/or reduced-price lunch eligibility distribution of students attending all of the 99,804 schools in the U.S.[9] For example, at schools located within 200 meters of a primary road or secondary road, 46.85% of students are eligible for either free or reduced-price lunches, whereas nationally, 45.64% of students are eligible for either free or reduced-price lunches.

Figure 3 depicts the cumulative (percentile) distributions of the percentage of students eligible for free/reduced-price lunches for schools near primary and secondary roads, and for all schools. Consistent with the tabular data presented in Tables 1 and 2, schools near primary roads tend to have higher fractions of eligible students, though the association is inconsistent across the distribution. Notably, the majority of schools located near primary roads have student populations of which >50% are eligible for free or reduced price lunch.

Figure 3 – Cumulative Distribution of Percentage of Student Eligible for Free or Reduced-Price Lunches for Schools Near and Far from Major Roads

[9] The highlighted free and/or reduced-price lunch eligibility proportions in Table 4 were tested against the corresponding free and/or reduced-price lunch eligibility's proportion in Table 3 to determine if the difference in the eligibility's proportion in the observed populations (i.e., the free and/or reduced-price lunch eligibility distributions for the schools within 200 meters of primary, secondary, and primary and secondary roads combined) and the whole U.S. population of schools is statistically significant. Using a chi-square goodness of fit test, the differences in the highlighted percentages in Table 4 and the corresponding percentages in Table 3 were all statistically significant ($p < 0.0001$).



The data and results are presented in the attached Microsoft Excel workbook, "Schools Near Roads Analysis for the Tier 3 NPRM docket." The first worksheet, 'All U.S. Schools,' lists each of the 99,804 U.S. schools and its associated location and site information (school name, address, county, state, latitude, longitude, NCES assigned School ID) demographic data (the number of students who identify as American Indian/Alaskan Native; Asian/Pacific Islander; Black; Hispanic; White, non-Hispanic; Two or more Races), school lunch program eligibility data, and total enrollment data. The second worksheet, 'Schools in 200m Buffers,' lists each of the 23,978 schools within 200 meters of a primary or secondary road (1,649 are near primary roads and 22,329 are near secondary roads). The data identifies each school's site information (school name, address, county, state, latitude, longitude, NCES assigned School ID) demographic data (the number of students who identify as American Indian/Alaskan Native; Asian/Pacific Islander; Black; Hispanic; White, non-Hispanic; Two or more Races), school lunch program eligibility data, and total enrollment data. The data also identifies the name of the road the school is within 200 meters of, whether that road is identified as a primary or secondary road, and the U.S. Census Bureau's ID for the road segment.

Summary

On a population-wide basis, there are millions of students attending classes in locations likely to have poor air quality resulting from the emissions of vehicles on the nearby roads.

The analyses described above provide evidence that schools near major roads include student populations which have higher percentages of nonwhite racial minorities and of recipients of free or reduced-price lunches. These analyses suggest that children in these demographic groups experience worse air quality, and as such, have potential for elevated risks of adverse health outcomes.

Overall, this analysis indicates that nonwhite and low SES students are likely to experience significant benefit as a result of the proposed standards.