



School Attendance Boundary Information Systems (SABINS) Database Design and Codebook

8/16/2011

Prepared by

Jeff Han, Senior Spatial Database Engineer

Version 1.0

1 Introduction

Although geography is a basic organizing feature of K to 12 education, the U.S. lacks an integrated database that describes student populations living in school attendance boundaries. The lack of a spatial database of school attendance boundaries has hindered the ability of researchers to undertake large-scale studies that explore phenomena such as the impact of school quality on housing prices, school socioeconomic composition on educational outcomes, or levels of racial and economic segregation across school catchment areas. To improve the quality and accessibility of geographic and demographic data to the scholarly community, this data infrastructure project will assemble and distribute a new spatial database called the School Attendance Boundary Information System (SABINS).

Each school attendance boundary in the SABINS database will have been integrated with data from the 2010 redistricting data, the U.S. Department of Education's Common Core of Data (CCD) and the Private School Universe Survey. SABINS will also link school attendance boundary geography with the American Community Survey and the Summary File 1 files when these datasets have been fully released. Linking demographic data with school catchment areas allows researchers to determine the characteristics of student-aged populations living in thousands of school catchment areas. This enables researchers, policy-analysts and educational administrators to work with customized datasets consisting of geography and corresponding statistical information of their choosing.

2 Definitions of School Attendance Boundaries and other Entities

2.1 Definitions

School districts are legal entities that provide educational services to all students within a region. Many school districts encompass a relatively large number of students spread over a sizable geographic area. For example, in many southern states such as Virginia, North Carolina and Florida, school districts are comprised of entire counties. In other areas of the country, school districts can be quite large—for example, the New York City School District covers roughly 300 square miles, contains over 1,000 school attendance boundaries and serves over 700,000 students. The U.S. Census Bureau distributes digital GIS files of school district boundaries for



the U.S. and there are roughly 13,700 school districts represented in the 2009 TIGER files available from the U.S. Census Bureau. Note, the Census Bureau distributes three separate school district level shapefiles for unified, elementary and secondary school districts.

School attendance boundaries are delineated by local, public school districts and serve as catchment areas that determine the (public) school a student likely attend (unless they opt to enroll in private school or other public school options such as a magnet school, a charter school, a vocational school, or a special education school). School attendance boundaries are labeled in many different ways by local school districts and are called many different things by the public. Examples include: "school boundaries," "attendance zones," "school zones," "school attendance areas," "school areas," "catchment areas," and, occasionally, "school districts." These various designations often cause confusion, particularly when distinguishing between the boundaries of school districts and the numerous school attendance boundaries within them.

Regular service schools are traditional, neighborhood public schools that serve children in a school attendance boundary. In most cases, one neighborhood school serves one school attendance boundary; however, schools and school attendance boundaries do not follow a strict "one-to-one" relationship. In some cases a neighborhood school can serve more than one school attendance boundary. In other cases two more schools can provide services to the same school attendance boundary.

Open enrollment zone is a type of school attendance boundary that allows children who live within a catchment area to attend any of two or more schools within the zone. For example, the school district of Minneapolis, MN has several open enrollment zones that contain two or more elementary schools within it. Students can apply to enroll in any of the elementary schools within the zone. (If a school receives more applications than its capacity, applications are subjected to a lottery.) In some school districts such as Milwaukee, all schools draw children from open enrollment zones. In Boston, the entire school district allows children to enroll in any of its schools and thus the entire district is one large open enrollment zone.

School attendance boundaries served by schools with sequential grades are school attendance boundaries that are served by two or more schools—but each school that serves children within the boundary enrolls children in different grades. For example, children in school "A" are enrolled in grades 00, 01, and 02 while children in school "B" are enrolled in grades 03, 04, and 05. Still, school A and B serve the same school attendance boundary.

Multipart boundary is a single geographic object in GIS that has two or more non-contiguous parts. A multipart feature can only share vertices but not edges. SABINS follows the Open GIS Consortium's (OGC) definition of "multipolygon," (<http://www.opengeospatial.org/>). Even though a multi-part boundary consists of multiple, non-contiguous areas, it still has one set of attributes/variables. For example, in a layer of states, the entire state of Hawaii is a multipart feature--all of its islands are described with the same information such as state name, number of residents, and so on. While most school attendance boundaries consist of a single feature, many



districts create school attendance boundaries with multiple parts. The SABINS data project will include information in its files that counts the number of parts associated with each school attendance boundary it distributes.

Common Core of Data (CCD) contains information for virtually every public school in the U.S. and is updated and distributed annually by the National Center for Educational Statistics (NCES). The data are edited by the U.S. Census Bureau and maintained in machine-readable datasets by NCES. The purpose of the CCD school-level survey is to provide a listing of all schools providing free, public elementary and secondary education, along with basic descriptive statistical information on each school. The CCD data are downloadable without cost to the public (from: <http://nces.ed.gov/ccd/pubschuniv.asp>) and are also available on the SABINS web page. The CCD provides variables describing the name and address of each school along with (approximate) geographic locations of school facilities. The CCD provides counts of the number of students (by race and gender) enrolled in each grade offered by the school. Additional variables include the schools operational status (i.e., whether it is active, it closed the previous year, or it will open during the following school year) and whether a school is a magnet, charter, or special educational school. These attributes allow us to determine whether a particular school *or schools* serve the children in a particular school catchment area (i.e., is a regular service school).

De Facto school attendance boundaries are school districts that are coincident with a single school attendance boundary. In such districts, there is one and only one regular service school in the school district that enrolls one or more children in each grade 00 to 12. Dublin City, Georgia is an example of a *De Facto* school attendance boundary. As shown in Table 3.1.1, Dublin City has five regular service schools:

Table 3.1.1 Schools by enrollment within Dublin City, GA School District

<u>School Name</u>	<u>00</u>	<u>01</u>	<u>02</u>	<u>03</u>	<u>04</u>	<u>05</u>	<u>06</u>	<u>07</u>	<u>08</u>	<u>09</u>	<u>10</u>	<u>11</u>	<u>12</u>
DASHER	292	243											
SAXON H.			222	194									
MOORE ST.					192	187							
DUBLIN M.							202	211	178				
DUBLIN H.										246	198	211	162

The Common Core of Data is used to estimate the number of De Facto school attendance boundaries. The CCD contains enrollment figures for each grade. If a school district has one and only one school with enrollment (at least one student) for each grade 00 to 12, the entire district is the same as a school attendance boundary. Schools identified as magnet, charter or special education schools are excluded from the process of identifying school districts that are *De Facto*



school attendance boundaries. (See appendix for STATA code that is used to identify school districts that are *De Facto* school attendance boundaries.)

There are roughly 13,700 elementary, secondary and high school districts in the U.S. (This estimate is based on the number of polygons present in the school district feature classes in the Census Bureau's 2009 TIGER files; some polygons represent areas that are not school districts.) Of these districts, roughly 7,000 are De Facto school attendance boundaries. These districts contain about 14 percent of the public school children in the 2007-2008 Common Core of Data (this estimate will be updated when the 2008-2009 CCD become available).

Using the CCD data to classify school districts as *De Facto* school attendance boundaries is not foolproof. The estimate of 7,000 school districts that are De Facto school attendance boundaries is a slight underestimate. Some regular servicing schools have very low enrollment numbers in one or more grades relative to the other grades served by the school. To illustrate the challenges of using the CCD to estimate the number of districts that are coincident with their schools, Table 3.1.2 provides revised student enrollment figures by grade for schools in Dublin City School District:

Table 3.1.2 Schools by enrollment within Dublin City, GA School District

<u>School Name</u>	<u>00</u>	<u>01</u>	<u>02</u>	<u>03</u>	<u>04</u>	<u>05</u>	<u>06</u>	<u>07</u>	<u>08</u>	<u>09</u>	<u>10</u>	<u>11</u>	<u>12</u>
DASHER	292	243											
SAXON			222	194									
MOORE					192	187							
DUBLIN M.							202	211	178				
DUBLIN H.									2	246	198	211	162

The last row of Table 3.1.2 shows enrollment figures for Dublin High School. The highlighted cell in the table shows that there are two children enrolled in 8th grade (even though there are well over 100 students enrolled in grades 9 to 12). The two children classified as 8th graders are either data entry errors or students who were demoted or promoted during the year (but nevertheless are classified in the 8th grade). Although there is one and only one school in Dublin City that serves 8th grade students, in this scenario Dublin City School District would not be identified as a *De Facto* school attendance boundary.

2.2 Data structure, file names and field descriptions

2.2.1 Background Information: grade-specific school boundary polygons



SABINS will assign a unique identification number to every grade-specific school attendance boundary polygon collected over the duration of the project. The structure of the unique identification number (which we describe in detail below) overcomes three challenges. The first is assigning unique identification numbers to school attendance boundary polygons that vary widely in the grades each of those boundaries serve. SABINS finds it intractable to devise a user-friendly *spatial* data structure that represents what are commonly known as “elementary schools,” “middle schools” and “high schools.” Although schools generally fall into the categories of elementary, middle, and high, the classification is different in different school districts—and even within the same school district. For example, “elementary schools” within the same school district can have grades 00 to 04; or 00 to 05; or 03 to 06. A middle school within this same district can have grade ranges such as 04 to 08, or 06 to 08. Finally, some schools within a district can have grades 00 to 08, or 00 to 12. Attempting to impose a three tier classification system of elementary, middle, and highs onto school attendance boundary polygons would be incompatible with the disparate grade range definitions adopted by school districts across the country. Also, a three-tier classification system may pose difficulties in accurately linking the school demographic data with school catchment areas. To overcome these challenges, SABINS will distribute the school attendance boundaries by individual grades. Hence, the spatial data for SABINS will consist of 13 polygon layers, with one polygon layer for each grade 00 through 12.

Determining whether a given school attendance boundary polygon serves a particular grade will be derived from one of two sources. The primary source is the Common Core of Data (which records public school enrollment by grade for all schools in the U.S.). In cases where the CCD provides questionable data describing whether a school attendance boundary serves a particular grade, SABINS staff contacts schools and school districts to verify the grade range of a school’s catchment area.

The second challenge is assigning a unique ID to polygon features in such a manner that it is possible to determine that a polygon in one single-grade feature class is coincident with a polygon in other single-grade feature classes. For example, many school attendance boundary polygons in the grade-two feature class will be coincident with school attendance boundary polygons in the grade-three feature class. We have developed an approach in which a portion of the primary key will be the same for any polygon features that are geographically coincident across the 13 feature classes. Thus, in a so-called “elementary school catchment area” that serves students enrolled in grades 00 to 04—and in which the school attendance boundaries in grades 00 to 04 are coincident—a portion of the primary key in feature classes 00 to 04 will be identical. The primary key fields will allow users to aggregate information stored in each grade with school attendance boundaries that are coincident.

The third challenge is devising a primary key that has the potential to allow users to identify attendance boundaries that remain coincident over time. This can be accomplished by linking the primary key field across vintages of SABINS features classes. If two school attendance



boundaries are coincident the 2009-2010 and 2010-2011 school years, their primary key will be the same.

2.2.2 *Name of SABINS shapefiles for grades 00 to 12*

There are 13 polygon layers (i.e., shapefiles) for each grade 00 to 12. The names of these shapefiles are **PY_SABINS_2010_00** to **PY_SABINS_2010_12**, where PY represents “polygon” and the last two digits of each features class designates whether it delineates grades 00 through 12 (where 00 represents Kindergarten). Descriptions of Primary Key Fields in featureclasses PY_SABINS_2010_00 through PY_SABINS_2010_12

2.2.2.1 **LEAID** This is the abbreviation for Local Education Agency Identification which is derived from the National Center for Educational Statistic’s designation for **school districts**. This is a seven (7) digit string field in which the first two digits indicate the state FIPS code and the remaining five digits indicate a school district identification number that is unique within each state. Therefore, the LEAID value assigned to each school district is unique nationally and is derived from the identification keys developed by the National Center for Education Statistics (NCES).

2.2.2.2 **BOUNDARYID** This is a unique ID that is assigned to every school attendance boundary polygon in the SABINS data set. It is a string field that is 16 characters long.

2.2.2.3 **SABINS_YEAR** is a two digit string field that tracks the vintage of the polygons.

2.2.2.4 **GRADE** is a two digit string field representing grade (i.e., 00, 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12) where “00” represents Kindergarten. Designating the grade Kindergarten as 00 preserves the inherent ordinal nature of the data for purposes of computational manipulation.

2.2.3 *Additional Fields in PY_SABINS_2010_00 to PY_SABINS_2010_12 shapefile attribute tables*

2.2.3.1 **IND_multipart** Polygons of school attendance boundaries in the 13 PY_SABINS_SABINSYR_GRADE features classes can be multipart. The “IND_multipart” attribute indicates whether a feature is multipart. Its data type is short integer and default value is NULL. *This is accomplished by counting the number of exterior rings in the polygon geometry.*

2.2.3.2 **IND_Open_Enrollment** This field indicates whether a school attendance boundary is an open enrollment zone. When open enrollment is true (value 1),



students can choose to enroll in one of the schools in the choice set. Its data type is short integer and default value is NULL.

- 2.2.3.3 ***Special_Case_Code*** This is a binary field where a value of one (1) indicates that this polygon represents a special case. Special cases are those school attendance boundary polygons that: (1) represent deviations from the dominate configuration of school attendance boundaries that are delineated by nearly all school districts and; (2) have been confirmed by SABINS staff as a special case to the dominant school boundary configuration. The codes are described in section 2.4.
- 2.2.3.4 ***IND_Coincident_W_District*** This field indicates whether or not a Local Educational Agency (typically called a school district) is coincident with all school attendance boundaries within a school district. This indicator is a Boolean field and its default value is null. A description of how SABINS determines whether a school district is coincident with a school attendance boundary is found in the “definitions section” above.
- 2.2.3.5 ***Object_ID*** This field is a unique identifier for ESRI's Geodatabase format. This is useful for the management of spatial data.
- 2.2.3.6 ***Shape*** This field Contains Spatial Information of a feature.
- 2.2.3.7 ***Area*** This field records the area of a polygon feature.
- 2.2.3.8 ***Perimeter*** This field records the perimeter of a polygon feature.

2.3 Companion Tables for PY_SABINS_2010_00 to PY_SABINS_2010_12

2.3.1 *Name and general description of companion tables to PY_SABINS_2010_00 through PY_SABINS_2010_12*

There are other attributes of the PY_SABINS_2010_00 through PY_SABINS_2010_12 polygon layers. These attributes are not properties of the geometry of school attendance boundary polygons and are thus placed in separate tables, named **NS_SABINS_2010_00** through **NS_SABINS_2010_12** (where NS represents “Non-Spatial”). The SABINSID field is the primary key in both tables. PY_SABINS_2010_00 and NS_SABINS_2010_00 are joined by the SABINSID and have a one-to-one relationship.



2.3.2 *Field Descriptions of NS_SABINS_2010_00 through NS_SABINS_2010_12*

- 2.3.2.1 ***SABINSID (Primary key)*** consists of LEAID, BOUNDARYID, SABINSYR and GRADE fields. Their descriptions are same as above for PY_SABINS_2010_00 through PY_SABINS_2010_12.
- 2.3.2.2 ***CNT_Regular_School*** This field shows number of traditional, neighborhood schools that serve children who reside within a school attendance boundary.
- 2.3.2.3 ***CNT_Magnet_Inside*** This field shows number of magnet schools inside a school attendance boundary for a particular grade. For example, if the a user is exploring the first-grade polygon layer, the CNT_magnet_inside variable will only count the number of magnet schools with a first grade. Magnet schools are designated as such by the CCD and their locations are also derived from the CCD. This field's data type is a three-digit integer. Its default value is NULL.
- 2.3.2.4 ***CNT_Charter_Inside*** This field shows number of charter schools inside the attendance boundary for a particular grade. For example, if the a user is exploring the first-grade polygon layer, the CNT_charter_inside variable will only count the number of magnet schools with a first grade. The point locations of charter schools are derived from the CCD. This field's data type is a three-digit integer. Its default value is NULL.
- 2.3.2.5 ***CNT_Special_Ed_Public_Inside*** This field shows number of other special education schools inside the attendance boundary for a particular grade. For example, if the a user is exploring the first-grade polygon layer, the CNT_special_ed_public_inside variable will only count the number of magnet schools with a first grade. The point locations of special education schools are derived from the CCD. This field's data type is a three-digit integer. Its default value is NULL.
- 2.3.2.6 ***CNT_Private_Inside*** This field shows number of private schools inside the attendance boundary. These data are from Private School Universe Survey—which supplies the latitude and longitude of each private school's location. This field's data type is a three-digit integer. Its default value is NULL.
- 2.3.2.7 ***Link_Raw_Path*** This field shows a path where raw data are stored. The raw data include shapefiles, CAD files, PDF, JPG and GIF images, narrative descriptions, and other source information used to digitize school attendance boundaries.
- 2.3.2.8 ***CNT_Open_Enrollment*** This field records number of schools that participate open enrollment for an area. Its value is equal to or greater than 2.



2.3.2.9 CNT_Multipart This field records number of parts for a multipart feature.

Users should note that there are some public schools in the CCD that do not fit one of the classifications we describe above. For example, schools for “homeless children” or “teenage parents” are examples of schools that are not classified as traditional neighborhood schools, nor as magnet schools, charter schools, or special education schools. Moreover, in some cases, there are discrepancies between the CCD’s classification of schools and a local district’s classification. For example, the CCD may classify a school as a non-magnet school while a local district defines a school as a magnet school.

2.4 Special Cases

2.4.1 Name and general description of Special Case Table

Some school districts delineate school attendance boundaries in relatively rare configurations. SABINS has created special case code lookup table called NS_SABINS_SCC_LOOKUP. This table describes special cases of school attendance boundary polygons that represent exceptions to the dominant boundary configuration delineated by nearly all school districts. The legitimate exception table provides information that allows SABINS users the flexibility of configuring the spatial and tabular information in a way that is useful to their particular research needs.

2.4.1.1 Special_Case_Code consists of a string field that is two characters long and that identifies the type of legitimate exception the polygon represents.

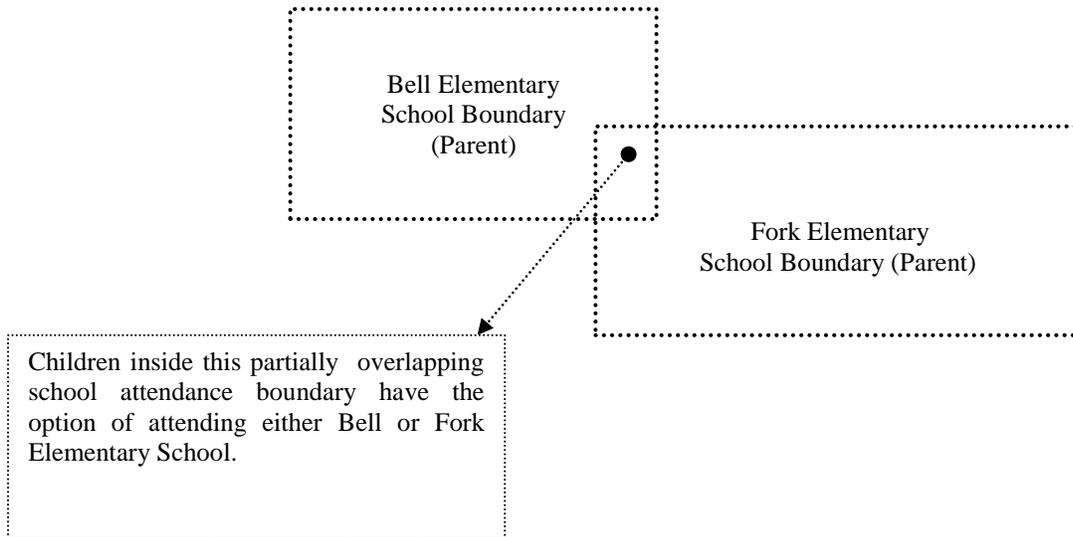
2.4.1.2 Description This is a 200 character string field. This field provides description explaining the particularities of the exception.

2.4.2 Special Case Code A: Open Enrollment

Open enrollment school attendance boundaries are those served by more than one school. Students can choose any of these schools or school districts assign students to one of the schools the provide services to the school attendance boundary.

2.4.3 Special Case Code B: Partially Overlap

Some school attendance boundaries are “partially overlapping” school attendance boundaries and have the following characteristics: 1) they are unique geographic objects (and are not simply “multipart objects”); 2) students who live within partially overlapping school attendance boundaries have a choice of schools from which they can select; 3) the schools from which students can choose have a “primary” school attendance boundary and the students who live within this primary boundary can only attend the “primary” regular service school. The following figure represents a partially overlapping school attendance boundary:



2.4.1 Exception Code C: Language Zones

Exception Code C: represents cases in which students who speak English as a first language attend the zone that we provide to users. Students who speak English and another language (i.e., who are bilingual) exception are segregation zones. Some school districts have separate school attendance boundaries for bilingual and monolingual students. For example, Alief School District in Texas has one set of school attendance boundaries for “bilingual students” and second set of boundaries for other students. We provide the school attendance boundaries for non bilingual students and note it the attribute table of the school that the zone is different for bilingual zones.

2.5 Tables associating the Common Core of Data (CCD) school-level data with PY_SABINS_2010_00 through PY_SABINS_2010_12 featureclasses

2.5.1 Name and general description of the CCD

The purpose of the CCD school-level survey is to provide a listing of all schools providing free public elementary and secondary education, along with basic descriptive statistical information on each school. The CCD data are downloadable without cost to the public (from: <http://nces.ed.gov/ccd/pubschuniv.asp>) and from the SABINS web page.

Each record in the CCD dataset represents a *public* educational facility. The CCD provides variables describing the (approximate) geographic locations of school facilities (i.e., geographic coordinates in GCS NAD 83). Therefore, a digital GIS point file can be generated with these location data. Each school record also provides the name of the school.



The CCD also counts the number of students enrolled in each grade and its operational status (i.e., whether it is active, it closed the previous year, or it will open during the following school year). The CCD describes whether a school is a magnet school, a charter school, or a special educational school. These attributes allow us to determine whether a particular school *or schools* serve the children in a particular school catchment area (i.e., is a regular service school). This process is described in section 3.5.3.

2.5.2 Relationship between schools in the CCD and SABINS school attendance boundary polygons

As noted, schools and school attendance boundaries do not share a one-to-one relationship. Their relationship is often complex (as described by exception codes A and B). To tame this complexity, SABINS provides a relational table that links regular servicing schools with the school attendance boundaries from which they draw their students. This is a many-to-many relationship in which a school attendance boundary polygon ID (SABINSID) is associated with at least one and often many schools (and vice versa).

2.5.3 Names of bridge tables between SABINS and CCD

The SABINS project provides 13 bridge tables between the CCD and SABINS—one for each grade. These tables will be called NS_SABINS_CCD_2010_00 through NS_SABINS_CCD_2010_12 for grades Kindergarten through 12th. Every record in NS_SABINS_CCD_2010_00 will contain a regular service school that enrolls children in grade 00. So do other tables.

2.5.4 Field descriptions of bridge tables NS_SABINS_CCD_2010_00 through NS_SABINS_CCD_2010_12

There are only two fields in each of the bridge tables. The sole purpose of the of bridge tables is to enable users to associate regular service schools in the CCD with school attendance boundaries.



2.5.4.1 *NCESSCH* is a string field of length 12 that uniquely identifies every school within the CCD's Public Elementary/Secondary School Universe Survey. This field is created by the National Center for Educational Statistics and allows users to link SABINS polygons to all variables in the CCD. A school's unique ID (NCESSCH) assigned by the National Center for Educational Statistics does not change over time (although the geometry of the boundary it serves may change over time).

2.5.4.2 *SABINSID* uniquely identifies a school attendance boundary. Their descriptions are same as above for PY_SABINS_2010_00 through PY_SABINS_2010_12.

2.5.5 Process of associating PY_SABINS_2010_00 through PY_SABINS_2010_12 features with traditional neighborhood schools in the Common Core of Data

SABINS produces a bridge table that associates (or links) schools in the CCD with the boundaries to which those schools supply services. This CCD bridge table has two fields: the first is the SABINSID that identifies school grade-specific school attendance boundary polygon. The second field is the NCESSCH code that uniquely identifies every school in the CCD.

Linking schools with their corresponding boundaries is a two-step process that requires the use of the CCD bridge table called "ns_sabins_ccd_00". (Users can refer to tutorials posted on the SABINS web page for specific instructions to undertake this task in ArcGIS.)

Step 1 entails joining (or matching) the attribute table of the school attendance boundary to the bridge table using the SABINSID. Note, the a specific SABINSID may appear more than once in the bridge table so users must ensure that every SABINSIS has information attached to it at this stage—a process that is automatically completed in ArcGIS.

Step 2 entails joining the CCD to the bridge table. The enables the user to determine the specific information about the schools that provide services to school attendance boundaries.

This approach is applied only to serving schools enrolling children who reside within a school attendance boundary. Magnet and charter schools can be matched to a school district using the LEAID.

3 Block Rectification

The primary set of geography SABINS supplies to the public are "block rectified" school attendance boundaries. This means that, in the SABINS database, school attendance boundaries are aggregates of Census blocks. The primary set of geography SABINS supplies to the public are "block rectified" school attendance boundaries. This means that, in the SABINS database, school attendance boundaries are aggregates of Census blocks. Most school attendance



boundaries closely follow TIGER/lines and, since these line features comprise Census blocks, most school attendance boundaries are, in fact, meant to entirely contain Census blocks. Still, some school districts delineate some of their school attendance boundaries such that a portion of a catchment area serves children on both sides of a street. In such cases, a school attendance boundary legitimately and intentionally splits a Census block. Still, the SABINS database assigns an entire Census block to an attendance boundary regardless of whether it is split by that attendance boundary. Thus, block rectified school attendance boundaries are not precisely the same as those delineated or described by a local school district. (SABINS allows users to obtain the original source information with a hyperlink stored in “link_raw_data” field so that they can have “pre-block rectified” files.)

The block–rectification process the SABINS project uses is straightforward. A point file is created that represents the geographic center (or centroid) of all U.S. Census blocks. (The Census Bureau’s block file contains the centroid coordinates for each block and these are used to create a point layer.) The point layer representing block centroids are then spatially joined with the school attendance boundaries—if a block point falls within a particular school attendance boundary polygon, the block centroid is assigned the school boundary identification code. Once the block points have been assigned to school attendance boundaries, the block points are then rejoined with the original block polygons from which these points were generated. After the block polygon file is associated with the identification codes of a school attendance boundary, the block polygons are dissolved into “block-rectified” school attendance boundaries. As discussed below, since most blocks have over 90 percent of their area within a school attendance boundary, the block rectification process has the effect of conflating school boundaries with Census blocks.

There are two reasons that school attendance boundaries are block-rectified. The first is to generate school attendance boundary population estimates using published Census data. The census block is the smallest unit for which the Bureau tabulates data. Thus, aggregating the block-level population totals to the block-rectified attendance boundaries provides accurate demographic information.

A second reason to block rectify school attendance boundaries is to conflate the line work of school attendance boundaries with Census blocks. Many school districts and states that supplied electronic GIS files to the SABINS project used local cadastral data to delineate their school attendance boundaries. For example, the streets and other line features used to digitize Delaware’s school attendance boundaries do not necessarily align with the MAF/TIGER lines. A school attendance boundary may follow a line segment such as a road but this road, as delineated in local data sources, is not aligned with MAF/TIGER files. After block rectification, a Census block along the periphery of a school attendance boundary will share geometry with the MAF/TIGER lines. This makes geocoding with MAF/TIGER line features more manageable.



4 Generating Population Estimates within SABINS Geography

At this writing, the Census Bureau has released the block-level redistricting files (also called the Public Law 94 files) that provide information describing the number of people for four basic tables: (1) race; (2) race by Hispanic origin; (3) race for people age 18 or over; (4) the number of households. We summarize this information to school attendance boundaries. This is a straightforward process. If a block has its centroid within a school attendance boundary then all of its people are assigned to the boundary. The metadata associated with the data download describe the attributes of the PL-94 that are tabulated to school attendance boundaries.