CARRA Working Paper Series

Working Paper *#2014-07*

Evaluation of Commercial School and Teacher Lists to Enhance Survey Frames

Quentin Brummet
U.S. Census Bureau

Mark Masterton
U.S. Census Bureau

Damon Smith
U.S. Census Bureau

Center for Administrative Records Research and Applications
U.S. Census Bureau
Washington, D.C. 20233

Paper Issued: July 1, 2014

Introduction

The Census Bureau conducts the Schools and Staffing Survey (SASS) for the National Center for Education Statistics (NCES). SASS is undergoing a redesign and reorganization and will be called the National Teacher and Principal Survey (NTPS) during its next administration in the 2015-16 school year. The NTPS will be conducted every two years, rather than every four years, and will likely have two to three components (teachers, principals, and possibly schools) rather than the traditional five components (districts, schools, principals, teachers, and library media centers).

Past administrations of the SASS utilized the following process. First, a school sampling frame was built using the Common Core of Data (CCD) and Private School Survey (PSS) databases. Second, schools were sampled from this frame. A teacher sampling frame was then built for the sampled schools using a Teacher Listing Form (TLF) developed by the SASS survey team at the Census Bureau. Last, the final SASS teacher sample is drawn from the teacher sampling frame.

During the TLF operation, every school was asked to provide a list of teachers along with their primary subject, full-time or part-time status, and years of experience. These lists were obtained using an Internet upload application as well as from paper forms that were mailed to individual schools and districts.[1] In past SASS administrations, the teacher listing was a time-consuming and resource-intensive operation. Hence, in an effort to reduce costs while maintaining appropriate coverage of teachers, NCES suggested that NTPS consider alternative sources for a teacher frame. In order to determine the viability of vendor teacher lists as an alternative or enhancement to the traditional TLF, the research team purchased school and teacher lists from three different vendors. This paper analyzes the quality and coverage of these vendor lists for the 2011-2012 school year relative to the school and teacher frames from the 2011-2012 SASS.

This paper assesses the ability of the vendor lists to provide the same coverage as the traditionally used SASS school and teacher frames. Potential coverage limitations of the traditionally used SASS school and teacher frames are not addressed and are considered outside the scope of this paper.

[1] While schools were permitted to use the Internet upload application, only school districts were encouraged to respond via Internet. The resulting distribution for mode of data collection was 39.6% district respondents via Internet, 11% school respondents via Internet, and 49.4% school respondents via paper TLF.

Description of the data sources

Vendor school files

The analysis considered the following variables based on their importance to NTPS sampling and availability in the various sources: school name, mailing address, NCES ID,[2] public/private school status, full time equivalent (FTE) teacher count, number of teachers linked to the school on the teacher file, and urbanicity.[3] For each public school record, the analysis also considered school type[4] and a flag indicating if the school was a charter school. The U.S. Census Bureau requested these variables because they are all included on the 2011-2012 SASS school universe file and each variable has traditionally been used during sampling. All three vendors were able to provide all of the variables that the Census Bureau requested. While a classification of schools into categories varies across data sets, each vendor provided NCES IDs that allow the files to be linked with the SASS school universe file.[5] The vendors provided school lists at the building level with each record representing a unique address. These lists were composed of schools in the 50 US States and the District of Columbia as of the 2011-2012 school year.

School universe file

The 2011-2012 SASS school universe file (also referred to as the 2011-2012 SASS school sampling frame) is a modified version of the 2009-10 Common Core of Data (CCD) and 2009-10 Private School Survey (PSS). As SASS is a mail-based survey, the modifications to the CCD and PSS focused on collapsing records to the address level. For example, on the CCD and PSS, a single address may contain multiple schools that, under the SASS definition, are only a single school.[6] The SASS school universe file was chosen for this analysis instead of the CCD and PSS in order to use address-level information.

Vendor teacher files

In addition to school lists, vendors were asked to provide a full teacher universe, composed of all available records for teachers in the 50 US States and the District of Columbia as of the 2011-2012 school year. The Census Bureau requested the following variables for each record of the teacher lists: Teacher name, school NCES ID, subject matter taught, full-time/part-time status,

[2] NCES ID is a unique identifier assigned to each public or private school in the United States.
[3] Defined using NCES urban-centric locale code.
[4] Regular school, special education, vocational, other/alternative school.
[5] The rate at which vendors were able to provide valid NCES IDs varied between data sets. This issue is discussed further below.
[6] Census did not perform any collapsing on the vendor files.

and years of teaching experience. These variables were requested because they are traditionally used during sampling for SASS and other education surveys.[7]

All three vendors provided a text description of "subject taught." None of the vendors provided full-time/part-time status. Vendor A provided "years at current school" and Vendor B provided "years teacher has been on the file." For sampling purposes, neither measure is an acceptable substitute for years of teaching experience. Vendor C was not able to provide any data on teaching experience. This limitation is an important consideration when evaluating the potential of vendor lists for use in future administrations of the NTPS.

Teacher sample file

The teacher lists provided by the vendors were compared to the teacher sample file used for the 2011-2012 SASS. The teacher sample file was built through a teacher listing operation. Teacher lists were collected through an online upload application or a paper TLF that was mailed to individual schools. This file contained teacher name, a code for subject matter taught, full-time/part-time status, years of teaching experience, and NCES ID of the teacher's school.

Definitions and key concepts

Coverage and ineligibility

The primary focus of this paper is to compare the coverage of vendor-provided lists with the traditional SASS school and teacher frames. The traditional SASS school and teacher frames are considered benchmarks for the purposes of this analysis.[8] "School coverage rate" refers to the proportion of schools on the universe file that are contained on each vendor file. "Teacher coverage rate" refers to the proportion of teachers on the SASS teacher sample file that are contained on each vendor file. Subsequent sections of this paper discuss the calculation of these rates in further detail.

"School ineligible rate" refers to the proportion of schools on each vendor file that are not contained on the school universe file. These schools are deemed "ineligible" for the purposes of this analysis; the term has no further meaning regarding the existence or operations of these schools. "Teacher ineligible rate" refers to the proportion of teachers on each vendor file that are not contained on the SASS teacher sample file. "Teacher ineligible rate by state" refers to the

[7] In the past, including during the 2011-2012 administration, SASS has oversampled new teachers in order to have a sufficient sample for the Teacher Follow-up Survey (TFS) and Beginning Teacher Longitudinal Survey (BTLS) in succeeding years. This practice is likely to continue in the future NTPS.

[8] For a further analysis on the traditional SASS public school frame, see NCES Working Paper 2000-12: Coverage Evaluation of the 1994-94 Common Core of Data.

average ineligibility rate for a given school within each state (i.e., a school-weighted average of the statewide ineligibility rate).

For the purposes of using vendor files to build a school frame, ineligibility and lack of coverage are both problematic if a completely vendor-based sample was desired. Coverage is a highly salient factor when evaluating potential sample frames. Low coverage rates for some types of schools would require the use of a dual sampling approach to draw a nationally representative sample. For example, if private schools are inadequately covered by the vendor lists, then a list of private schools from a different source would be necessary to yield a nationally representative sample. High ineligibility rates can lead to extra time and resources spent investigating and processing potential ineligible cases, and higher mailing costs. However, given that the CCD and PSS are available at extremely low cost, future administrations of the NTPS could still cut costs dramatically by using the traditional SASS school sampling frame from the CCD and PSS while using teacher lists from vendors to form a teacher sampling frame.[9] Hence, coverage rates are considered a more important factor than ineligibility rates when assessing the potential utility of a give vendor's teacher lists.

Definition of school and teacher

When comparing the vendor school and teacher lists to the sampling frames used for the 2011-2012 SASS, definitional issues may account for differences. These issues can affect both coverage and ineligibility. For example, the vendor lists may include non-teaching school personnel such as guidance counselors or administrative staff that would not be included on the traditional SASS teacher frame. During a production cycle, these differences would result in additional processing work in order to ensure that the vendor lists are sample-ready.

During administration of the 2011-2012 SASS, a school was defined as: an institution or part of an institution that has one or more teachers who provide instruction to students, has students in one or more of grades 1–12 (or the ungraded equivalent), has its own principal/administrator if it shares a building with another school or institution, is in operation during the 2011–12 school year, and is NOT primarily a postsecondary or adult basic education institution. The following are NOT considered a school: schools located exclusively in a private home, Department of Defense (DoD) schools located outside of the US, offices of special education in a Local Education Agency (LEA), tutoring services, homeschool clearing houses, and adult learning facilities.

[9] It is also feasible that schools deemed ineligible in this analysis should have been included in the CCD or PSS data used to create the SASS sample. Coverage errors in this direction are out of scope for this analysis.

During administration of the 2011-2012 SASS, a teacher was defined as a full-time or part-time teacher who teaches any regularly scheduled classes in any of grades K–12. This includes administrators, librarians, and other professional or support staff that teach regularly scheduled classes on a part-time basis. Itinerant teachers are included, as well as long-term substitutes who are filling the role of a regular teacher on a long-term basis. An itinerant teacher is defined as a teacher who teaches at more than one school (e.g., a music teacher who teaches 3 days per week at one school and 2 days per week at another). Short-term substitute teachers and student teachers are not included.

Vendor A did not provide a definition for teacher but did identify schools as educational organizations that provide instruction. These can be designed for "the teaching of elementary and secondary age children, adult education, career and technical education, and/or special needs."

Vendor B defines a school as "an institution that provides instruction for a group of students. These institutions may be private or public, are often sourced through the school district, governing agency, phone calls, NCES or web research." Vendor B defines a teacher as "those who are identified as teachers by the school either through surveys, state databases or the schools themselves."

Vendor C defines a school as "institutions identified as "schools" by the NCES, state department of education or the individual school district. The definition varies from state to state, and district to district." Vendor C defines a teachers through a job coding system that attempts distinguishes teaching personnel from other school staff such as Teacher Aides.

Vintage of vendor files

The vendor files that were purchased for this analysis were purchased in 2013 and meant to reflect universes as of the beginning of 2011-2012 school year. During the acquisition process, representatives from each vendor noted that due to the nature of their database updating[10], it would be difficult to draw a file for a precise time. This issue could alter both coverage and ineligibility rates.

[10] Each vendor stated that they can include a date stamp variable, indicating when the record was last modified, on future files.

Timing of acquisition

These files were acquired during the summer of 2013, and, as mentioned above, are retroactive in nature. Additionally, they are intended to be entire universes and cover all teachers in the U.S. In a live, production environment, the files would be acquired for specific schools during the fall of the school year for immediate teacher sampling. This analysis does not provide any insight into the ability of the vendors to provide sample-ready files that meet certain specifications within a defined timeframe. Once the data terms and conditions were agreed upon and the contracts were signed, the vendors delivered extracts very quickly.

Analysis of School Coverage

The initial assessment of the coverage of the vendor lists focused on the school coverage. This was based on the assumption that if the vendor lists had low coverage rates for schools, teacher coverage rates would be low as well. However, because the CCD and PSS universe files are publicly available, the expectation was that the vendor school lists would have high coverage rates for both public and private schools.

The following results display coverage and ineligible rates of schools matched on NCES ID. This analysis dropped schools that are not located in one of the 50 US States or the District of Columbia, as well as schools that were missing NCES ID. This analysis did not carry out an additional coding operation to assign missing NCES IDs (where applicable). NCES ID was a necessary linking variable during SASS sampling and data product creation; therefore, the vendors' ability to provide NCES IDs without additional burden on NTPS staff was a salient factor in assessing the vendors' ability to replace the previous SASS frame.

Results

Table 1 contains the overall coverage rate of each vendor file for both public and private schools. Coverage rate is calculated by dividing the number of schools contained in both the vendor file and the SASS school universe file by the number of schools in the SASS school universe file. The table shows that the vendors have a higher coverage rate in public schools (85 percent and higher) compared to private schools (less than 66.5 percent.) The PSS, the private school universe source, is collected by the Census Bureau for NCES, and is publically available. The relatively low coverage rate for private schools suggests that the vendors' lists will not adequately serve as a sampling frame for private schools. Therefore, the majority of the subsequent analyses in this paper focus on public schools.

None of the vendors exhibit perfect coverage, but at 95 percent, Vendor C has a higher coverage rate for public schools than Vendor B (86 percent) and Vendor A (85 percent.). The following

analyses explore the characteristics of matching and non-matching schools to describe additional indicators of vendor list quality.

Table 1. School coverage rate, by control of school and vendor

	Coverage Rate		
Control of school	Vendor A	Vendor B	Vendor C
Public	85.43%	86.41%	92.25%
Private	-†	66.46%	60.08%

Source: Vendor school files linked to 2011-2012 SASS school universe file.
N = 95,606 for public schools, 27,295 for private schools.
† Vendor A did not provide NCES ID for private schools

Vendor A has much lower ineligible rates than Vendor B, with Vendor C providing the highest ineligible rates. Comparing these ineligible rates to the coverage rates shown in Table 1 reveals that the vendors with the highest coverage rates also include the highest ineligible rates.
In addition, the public school ineligible rate for each vendor file is lower than the ineligible rate for private schools. For example, Vendor B's private school ineligible rate was almost double its public school ineligible rate (8.09 percent compared to 4.60 percent). The ineligible rate for Vendor C's public schools (5.6 percent) is less than a quarter of the rate for private schools (28.1 percent.). Note again that high ineligible rates may not be a concern for school lists, as the CCD and PSS can be used as school sampling frames instead of vendor school lists.

Table 2. School ineligible rate, by control of institution and vendor

	Ineligible rate		
Control of school	Vendor A	Vendor B	Vendor C
Public	2.98%	4.60%	5.59%
Private	-†	8.09%	28.09%

Source: Vendor school files linked to 2011-2012 SASS school universe file.
N = 95,606 public schools, 27,295 private schools.
† Vendor A did not provide NCES ID for private schools

Table 3 provides shows the school coverage rate in public schools by charter status and school type. The results show that the vendor files have much higher coverage of regular and traditional public schools.

Table 3. **Coverage rate for public schools, by charter status, school type, and vendor**

	Coverage Rate		
	Vendor A	Vendor B	Vendor C
Charter Status			
Traditional Public School*	88.75%	87.93%	95.79%
Charter School	39.04%	66.59%	89.41%
School Type			
Regular School	91.01%	91.08%	97.21%
Non-regular school†	37.61%	46.74%	79.86%

Source: Vendor school files linked to 2011-2012 SASS school universe file.
N=95,606.
*Traditional Public School refers to any public school that is not a charter school.
† Non-regular schools encompass special education schools, vocational schools, and other/alternative schools

Table 4 shows that there is substantial variation in the public school coverage rate within each vendor across states (e.g., Vendor A ranges from 53.15% in the District of Columbia to 95.89% in Maryland). Note that states with relatively low coverage rates in one vendor file are likely to have relatively low coverage rates in the other vendor files. Arizona, Minnesota, and the District of Columbia all have among the lowest coverage rates of any state, and trend is consistent across all three vendor files, though coverage rates vary across vendors. Vendor C has a higher coverage rate than Vendor A and Vendor B in every state, while Vendor B has a higher coverage rate than Vendor A in 33 states.

Table 4. School coverage rate, by selected states† and by vendor

State	Coverage Rate		
	Vendor A	Vendor B	Vendor C
Hawaii	97.93%	98.97%	99.31%
Maryland	95.89%	93.80%	98.96%
Massachusetts	92.37%	92.48%	98.47%
District of Columbia	53.15%	77.93%	84.68%
Minnesota	59.89%	67.29%	85.57%
Arizona	56.48%	72.79%	87.93%

Source: Vendor school files linked to 2011-2012 SASS school universe file.
N=95,606.
† Top/bottom 3 states by coverage rate only. For full table, see Appendix Table A.1.

The results of the school level matching indicate that the vendor school lists have very high coverage for regular, non-charter public schools. Specifically, regular, non-charter public schools in most states have greater than 90 percent coverage in all three vendor lists. Vendor C consistently has the highest rate of coverage; however, it also has the highest rate of ineligible schools. Analyst review of the ineligible schools did not reveal any obvious patterns. While some schools appear to be virtual instruction programs or non-existent schools, some ineligible schools in the vendor files appear to be genuine schools that were not included on the 2011-2012 SASS sample frame.

Analysis of Teacher Coverage

The primary assessment of the quality of vendor lists comes from calculating coverage and ineligibility rates for teacher lists. This section first discusses the matching process for linking teachers between vendor lists and the SASS teacher sample frame, then presents coverage and ineligibility rates for the vendor teacher lists.

Matching methodology

The first step in the process of matching teachers from vendor to SASS files was to standardize teacher names. This process accounts for any nicknames and differential spellings of names. The match is done using a probabilistic approach that gives more weight to rare names. Hence, it is easier to match teachers with relatively rare names when compared to teachers with relatively common names.

Teachers were only matched within schools. If one of the vendor lists contained a teacher's name at a different school, this teacher was not considered a match.

Four passes were made in an attempt to match teachers between the two files. If a teacher matched in a particular pass, they were set aside and only unmatched teachers were included in the next pass.

1. Within matched schools based on NCES ID, teachers were matched based on name.[11]
2. Within matched schools based on NCES ID, teachers were matched based on teacher names and first and last names were flipped in an attempt to uncover data entry errors.
3. Within address, teachers were matched based on name.
4. Within address, teachers were matched based on names, flipping first and last names in an attempt to uncover data entry errors.

Prior to passes 3 and 4, addresses on both files were standardized. In addition to parsing addresses for use in the probabilistic match process, this standardization accounts for different abbreviations of roads (e.g., "Rd." and "Road").

Note that because we use address as a matching field, schools that did not match on NCES ID in the section above may still have teachers who match between SASS and vendor files.

Results

Table 5 shows the overall coverage rates of each vendor for public and private schools. As shown in the table, the coverage rate for private schools is extremely low. For the rest of this section, the tables will focus on public schools only.

Table 5. Teacher coverage rate, by control of institution and vendor

	Coverage Rate		
Control of school	Vendor A	Vendor B	Vendor C
Public	70.60%	58.37%	55.80%
Private	-†	34.56%	22.87%

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file. N=390,999.

† Vendor A did not provide NCES ID for private schools.

[11] Matching based on names only removes what are considered fake names (e.g., "Parent", "Mr. Miss"). Initials and very short names are not altered, and a string comparator is used to evaluate the similarity of first and last names separately.

Table 6 shows the coverage rate in public schools by charter status and school type. As in the school-level analysis, the coverage rate of the vendor files is much higher in regular and non-charter public schools. For example, 72.58 percent of all public, non-charter teachers in the 2011-2012 SASS teacher sample frame can be found in Vendor A's teacher lists. Additional tables showing coverage rates for all schools are available in Appendix B.

Table 6. Teacher coverage rate for public schools only, by charter status, school type, and vendor

	Coverage Rate		
	Vendor A	Vendor B	Vendor C
Charter Status			
Traditional Public School*	72.58%	59.48%	56.58%
Charter School	27.93%	25.83%	32.64%
School Type†			
Regular School	72.02%	59.75%	57.05%
Non-regular school	38.88%	27.36%	27.81%

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=351,418.
*Traditional Public School refers to any public school that is not a charter school.
† Non-regular schools encompass special education schools, vocational schools, and other/alternative schools

Vendor coverage rates in public, non-charter schools were generally higher in cities and suburbs than in rural areas (see Table 7a). Urbanicity is defined here by the urban centric locale code created for CCD. Table 7b shows the distribution of schools by locale code in the 2011-2012 SASS, showing that the SASS contained a relatively even distribution of schools across different locales.

Table 7a. Teacher coverage rate for public, non-charter schools, by locale and vendor

CCD Urban-Centric Locale	Coverage Rate		
	Vendor A	Vendor B	Vendor C
City, Large	63.88%	43.41%	41.89%
City, Midsize	72.19%	55.10%	53.91%
City, Small	74.39%	58.48%	57.05%
Suburb, Large	73.69%	62.91%	58.89%
Suburb, Midsize	74.37%	60.48%	58.82%
Suburb, Small	74.98%	67.26%	62.46%
Town, Fringe	68.77%	61.15%	60.80%
Town, Distant	71.92%	63.57%	60.57%
Town, Remote	70.35%	61.28%	56.12%
Rural, Fringe	72.86%	59.90%	58.01%
Rural, Distant	69.55%	61.78%	58.58%
Rural, Remote	67.08%	54.66%	48.49%

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=335,864.

Table 7b. SASS schools by locale code

CCD Urban-Centric Locale	Total[†]	Percent
City, Large	995	10.28%
City, Midsize	537	5.55%
City, Small	740	7.64%
Suburb, Large	2149	22.20%
Suburb, Midsize	290	3.00%
Suburb, Small	175	1.81%
Town, Fringe	187	1.93%
Town, Distant	690	7.13%
Town, Remote	646	6.67%
Rural, Fringe	1501	15.50%
Rural, Distant	1032	10.66%
Rural, Remote	739	7.63%

Source: 2011-2012 SASS teacher sample frame file.

[†] "Total" refers to total number of schools in the teacher sample file for the 2011-2012 SASS.

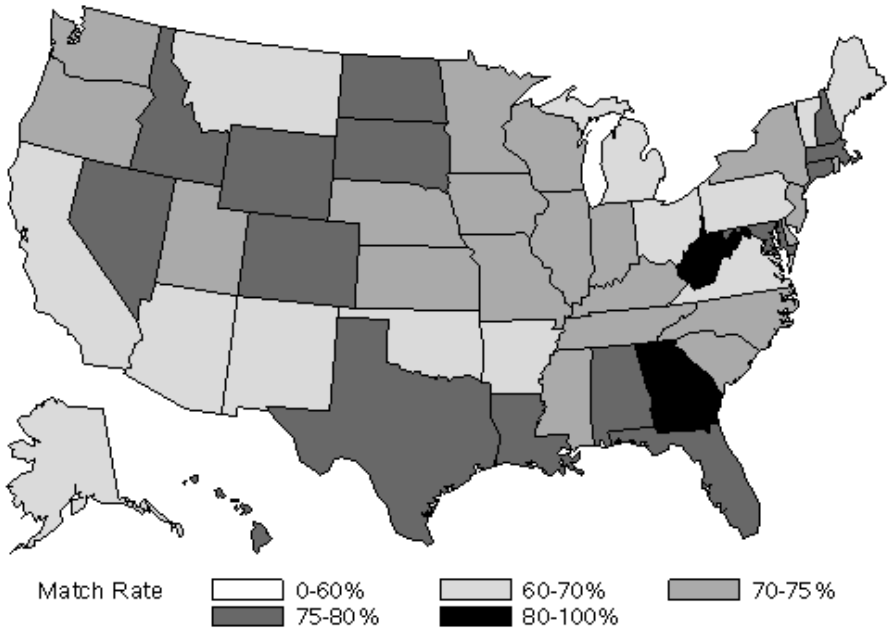
Figures 1a-1c display teacher coverage rates across U.S. States. As can be seen, teacher coverage rates follow a similar pattern to the school coverage rates. There is significant variation across states within each vendor file and there is strong correlation across the teacher files within each state. As expected, there appears to be a relationship between teacher coverage and school coverage across states. Vendor A has a higher teacher coverage rate than Vendor B and Vendor C in almost every state (See Figures 1a-c). Vendor B has a higher teacher coverage rate than Vendor C in 43 states. The exact figures for a handful of states are shown below in Table 8.

Table 8. Teacher coverage rate for public, non-charter schools, by selected state and vendor †

	Coverage Rate		
Selected State	Vendor A	Vendor B	Vendor C
Idaho	78.99%	66.11%	62.65%
Connecticut	77.96%	71.69%	68.98%
New Hampshire	76.27%	70.00%	70.69%
Oklahoma	64.15%	54.54%	50.80%
New Mexico	66.32%	49.94%	40.57%
District of Columbia	45.61%	13.16%	30.10%

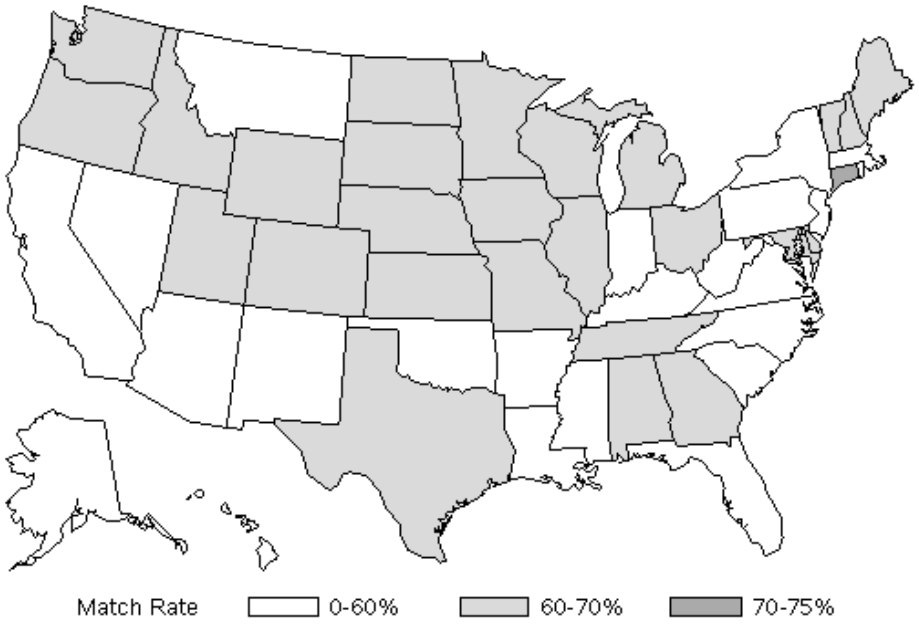
Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=335,864.
† Top/bottom 3 states by coverage rate only. For full table, see Appendix Table A.2.

Figure 1a. Public, Non-Charter Teacher Coverage Rate across U.S. States: Vendor A



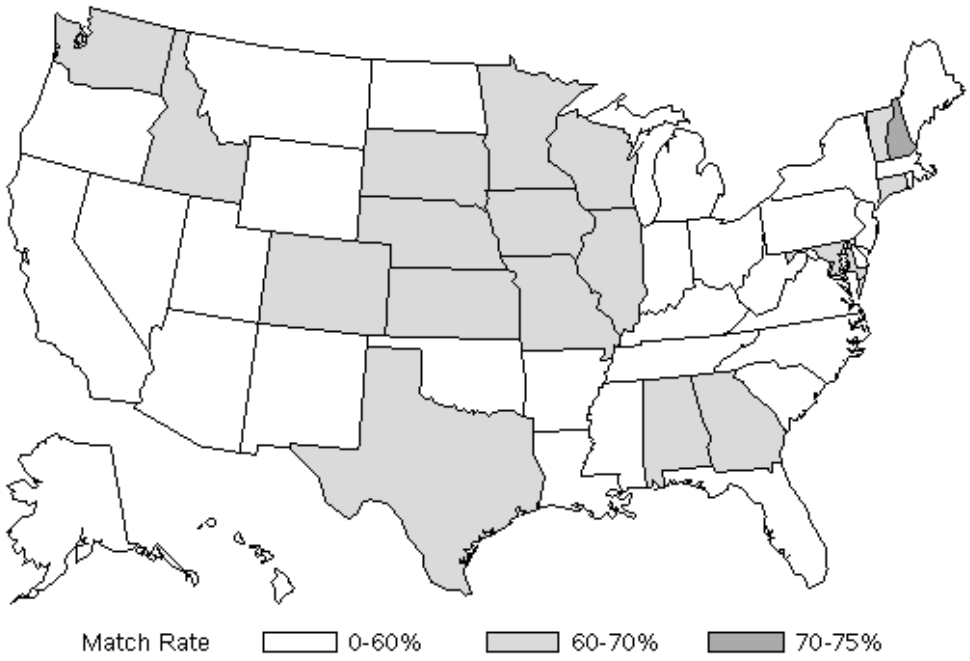
Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.

Figure 1b. Public, Non-Charter Teacher Coverage Rate across U.S. States: Vendor B



Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.

Figure 1c. Public, Non-Charter Teacher Coverage Rate across U.S. States: Vendor C



Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.

Table 9 shows the distribution of matching rates within public, non-charter schools for each vendor file. Over 58 percent of schools on the Vendor A file have a greater than 75 percent coverage rate, while Vendors B and C only have 37.81 percent and 33.02 percent of schools with a greater than 75 percent coverage rate, respectively.

Table 9. Distribution of teacher coverage rates within public, non-charter schools, by vendor

Match Rate	Vendor A	Vendor B	Vendor C
Less than 50%	15.37%	32.55%	32.58%
50-75%	25.95%	29.64%	34.40%
76-90%	42.55%	27.24%	26.35%
Greater than 90%	16.13%	10.57%	6.67%

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=335,864.

Teachers who were reported to have more teaching experience on the TLF displayed higher coverage rates than teachers with less reported experience (see Table 10). This difference is important for NTPS sampling, where new teachers are regularly oversampled for use in other surveys (the Teacher Follow-up Survey and the Beginning Teacher Longitudinal Study).

Table 10. Teacher coverage rate, by teacher years of experience and vendor, public non-charter schools only

Years of teaching experience	Coverage Rate		
	Vendor A	Vendor B	Vendor C
1st year	56.22%	23.59%	25.88%
2-3 years	68.15%	43.51%	47.77%
4-19 years	74.75%	62.73%	60.00%
20 or more years	73.22%	67.51%	60.39%

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=335,864.

Differences in teacher coverage rates were observed depending on the teachers' full-time/part-time status reported on the TLF (see Table 11). Full-time teachers had higher coverage rates on each vendor file.

Table 11. Teacher coverage rate, by full-time status and vendor, public non-charter schools only

Teacher status	Coverage Rate		
	Vendor A	Vendor B	Vendor C
Full-time	74.20%	60.60%	57.74%
Part-time	50.81%	49.10%	46.41%

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=335,864.

On over 95 percent of the schools on all three vendor files, at least 10 percent of the listed teachers were ineligible (See Table 12). Over half of Vendor B's teachers were ineligible in 52.35 percent of the schools. Less than 10 percent of Vendor A's schools contained high (above 50 percent) ineligibility rates, while less than 1 percent of schools on Vendor A's file contained very low (less than 10 percent) ineligibility rates. However, compared with the other two vendors, Vendor A has lower ineligibility rates. Nearly 49 percent of Vendor A's schools have

ineligibility rates of 25 percent or less, compared with less than 7 percent of Vendor B's schools and just over 14 percent of Vendor C's schools.

Table 12. Distribution of teacher ineligible rates within public, non-charter schools, by vendor

Ineligible rate	Vendor A	Vendor B	Vendor C
Less than 10%	0.96%	4.30%	3.58%
10-25%	48.00%	2.61%	10.55%
26-50%	41.97%	40.74%	58.16%
Greater than 50%	9.06%	52.35%	27.71%

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=335,864.

Many schools on the vendor file contained a higher total number of teachers than the SASS teacher sample file. These differences in counts explain a great deal of the ineligible rates that were shown in Table 12. There are multiple potential explanations for these differences. First, the vendor lists include school personnel that do not fit the NCES definition of a teacher. Brief analyst review suggests that inclusion of administrative and other personnel (nurses, custodians, guidance counselors, etc.) accounts for much of these differences. Second, the vendor files include duplicate teachers. This analysis left potential duplicates in the file, as extra time would be needed during a production cycle to identify and account for these duplicates.

Figures 2a-2c show that there is substantial variation across states within each vendor file and there is strong correlation across the teacher files within each state. Exact numbers for a few select states are shown below in Table 13. Vendor A has a lower teacher ineligibility rate in every state than Vendor C, and Vendor A has a lower ineligibility rate in every state than Vendor B except for Montana. However, Vendor C has a lower ineligibility rate than Vendor B in all states except for Montana and Connecticut. While Vendors B and C tended to have similar teacher coverage rates, Vendor C typically has lower ineligibility rates than Vendor B, which is a notable difference between the two files.

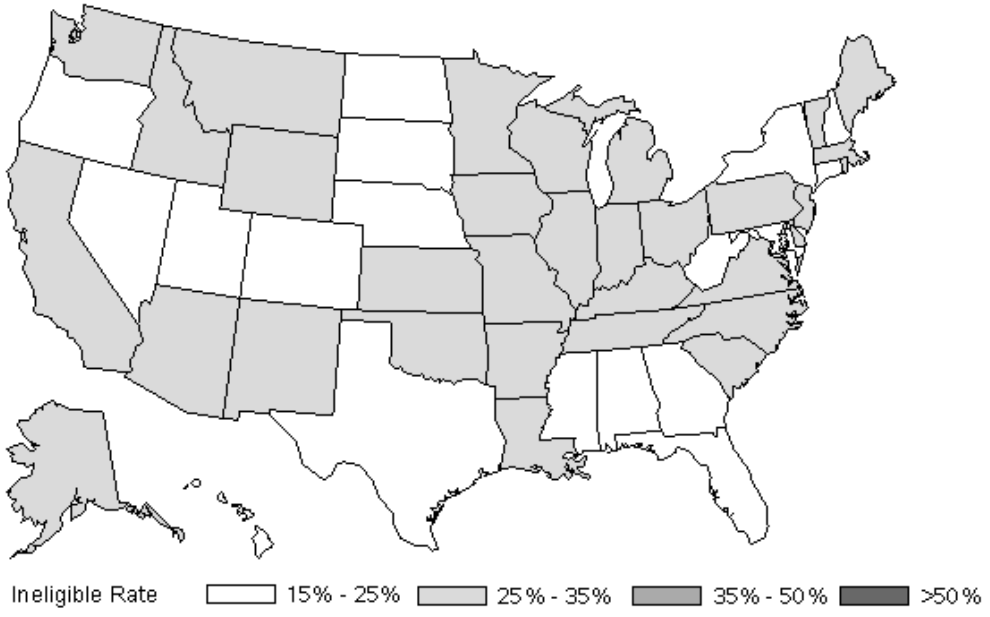
Table 13. Teacher ineligible rate by selected state and vendor †

	Ineligible Rate		
Selected State	Vendor A	Vendor B	Vendor C
Connecticut	21.47%	32.37%	39.07%
North Dakota	21.48%	48.07%	39.08%
Texas	23.07%	51.20%	37.01%
Arizona	29.83%	58.24%	49.10%
Alaska	33.77%	60.17%	50.97%
District of Columbia	40.64%	72.83%	58.05%

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=335,864.

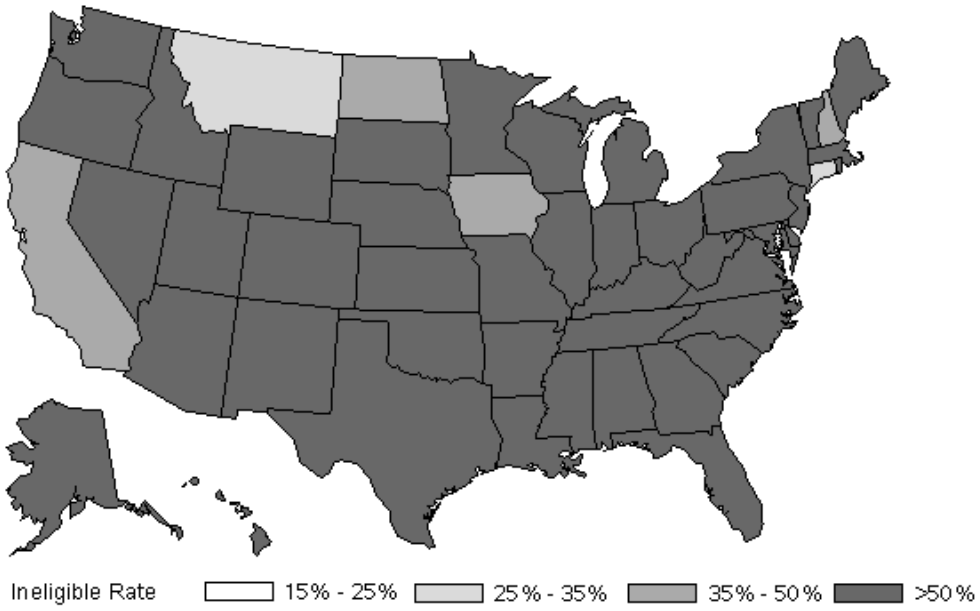
† Top/bottom 3 states by ineligible rate only. For full table, see Appendix Table A.3.

Figure 2a. Teacher Ineligible Rate across U.S. States: Vendor A



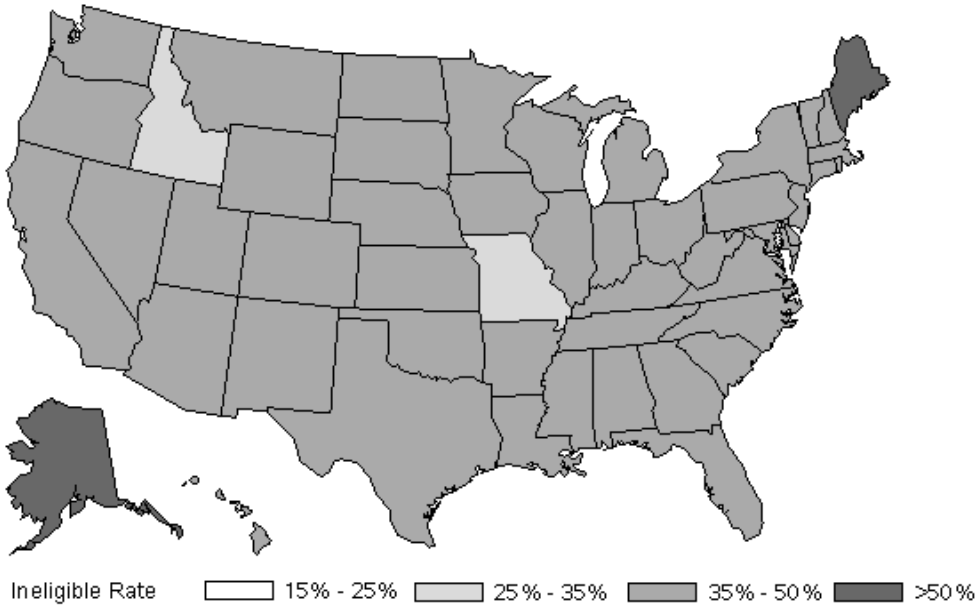
Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.

Figure 2b. Teacher Ineligible Rate across U.S. States: Vendor B



Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.

Figure 2c. Teacher Ineligible Rate across U.S. States: Vendor C



Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.

In general, the results of the teacher-level matching demonstrate that there are school subgroups where the vendor teacher lists fail to provide adequate coverage. Specifically, private, charter, and non-regular schools have extremely low teacher coverage rates. In addition, the coverage rates of Vendor A are reliably higher than the coverage rates of both other vendors. While ineligibility rates for Vendor A are also higher, in cases (as described above) where non-teaching personnel are listed, automated programs could be employed during production to search for specific keywords (provided in the subject matter/job description character fields) and drop these records from the file.

The low school coverage rates for private, charter, and non-regular schools partially explain the low teacher coverage for these subgroups. This was expected prior to the teacher analysis, given the lower than average TLF response rates associated with these types of schools during previous production cycles of SASS.

Conclusions and Recommendations

Based on the analysis of the lists provided by vendors, the Census Bureau has two recommendations for the National Teacher and Principal Survey (NTPS) Field Test.

1. Continue Using the CCD and PSS to Build a School Frame

The Census Bureau recommends that NTPS continue to use the Common Core of Data (CCD) and Private School Survey (PSS) as the basis for developing the school sampling frame. For public, non-charter schools, each vendor had a school match rate of over 87 percent. However, for private schools, charter schools, and non-regular schools, these match rates were lower. The inclusion of private, charter, and non-regular schools is an important characteristic of SASS, traditionally, and will continue to remain so for the NTPS. Thus, a dual-frame approach would be necessary for school sampling. Because CCD and PSS are developed independently of NTPS, the potential cost savings of using vendor school lists is offset by the added complexity of developing a dual frame sampling approach for schools at this time.

2. Evaluate Vendor A's Teacher Lists Further During the 2014-15 NTPS Field Test

The Census Bureau recommends conducting further analysis of the teacher lists provided by Vendor A. Vendor A had higher teacher matching rates than Vendor B and Vendor C across nearly every dimension, despite having lower school matching rates than each of the other vendors by some characteristics. Additionally, table 9 shows that within 48.68 percent of

schools, Vendor A matches over 75 percent of teachers. This suggests that Vendor A may be able to provide a close approximation of the teacher lists gathered by the Teacher Listing Form (TLF) operation, given certain characteristics. Specifically, Vendor A matched approximately 72 percent of teachers in public, non-charter regular schools (Table 6). Due to the issues highlighted previously in this paper that could affect the teacher matching rates, the Census Bureau considers further analysis of Vendor A's ability to provide teacher lists necessary

While a dual-frame approach for schools was not recommended above, it is a possible option for developing a teacher frame. This is because, unlike CCD and PSS, the TLF operation is not conducted independently of NTPS. Purchasing teacher lists for only certain schools (public, non-charter, regular) could still result in significant cost savings compared with collecting these teacher lists via the TLF operation.

The Census Bureau proposes acquiring a "live" teacher list concurrently with the TLF collection scheduled to take place in fall 2014. This process can begin as soon as the field test school sample is complete (scheduled to occur in late spring 2014). The Census Bureau will request that Vendor A provide teacher lists for each public, non-charter regular school that is included in the field test sample. These lists will be compared with those collected by the 2014-15 TLF operation. This will allow analysis of the following features that were impossible to evaluate during this paper's analysis:

- The cost charged by Vendor A to provide a "live" teacher list during a production cycle.
- Vendor A's capacity to provide teacher lists in a timely manner during a production cycle.
- The ability of Vendor A to provide teacher lists for schools that are non-respondents during the TLF collection. The Census Bureau would have the opportunity to conduct follow-up analysis on these teachers.
- Vendor A's proficiency in providing a teacher list for a specific set of schools and the requirements associated with this type of acquisition.

The Census Bureau concludes that, with a field test scheduled for 2014-15, the added cost and work associated with this additional analysis is low. The field test is an opportunity to evaluate the viability of vendor-provided teacher lists in a production environment. During the acquisition of the vendor lists, vendor representatives indicated that their databases were more precise "live," rather than representing a specific point of time in the past. If this is true, then the Census Bureau would expect an increase in teacher matching rates. Depending on the magnitude of this increase, Vendor A's teacher lists could be a viable alternative for developing a teacher sampling frame.

Appendix A: Results by State

Table A.1 Public School Coverage Rate by State

State	Vendor A	Vendor B	Vendor C
AL	0.8723	0.8748	0.9368
AK	0.9050	0.9496	0.9690
AZ	0.5648	0.7279	0.8793
AR	0.8956	0.9191	0.9519
CA	0.8775	0.8374	0.9725
CO	0.9048	0.9072	0.9659
CT	0.8411	0.8848	0.9384
DE	0.8380	0.8750	0.9444
DC	0.5315	0.7793	0.8468
FL	0.8467	0.8341	0.9456
GA	0.9082	0.8824	0.9673
HI	0.9793	0.9897	0.9931
ID	0.8627	0.8516	0.9626
IL	0.8900	0.9187	0.9802
IN	0.9028	0.9090	0.9624
IA	0.8878	0.8894	0.9558
KS	0.9090	0.8893	0.9614
KY	0.8398	0.8543	0.9703
LA	0.8784	0.8634	0.9406
ME	0.8040	0.6750	0.9207
MD	0.9589	0.9380	0.9896
MA	0.9237	0.9248	0.9847
MI	0.7558	0.8000	0.9152
MN	0.5989	0.6729	0.8557
MS	0.9068	0.8967	0.9753
MO	0.8577	0.8846	0.9750

MT	0.7053	0.9298	0.9737
NE	0.7514	0.8570	0.9593
NV	0.8704	0.8567	0.9253
NH	0.9336	0.9646	0.9690
NJ	0.8905	0.8974	0.9509
NM	0.8596	0.8727	0.9659
NY	0.9120	0.8996	0.9531
NC	0.8970	0.9219	0.9623
ND	0.8598	0.9180	0.9788
OH	0.8380	0.8819	0.9441
OK	0.9076	0.9257	0.9685
OR	0.9136	0.8848	0.9751
PA	0.8884	0.8869	0.9622
RI	0.8818	0.9073	0.9457
SC	0.9390	0.9323	0.9691
SD	0.7676	0.7739	0.9544
TN	0.9575	0.9441	0.9784
TX	0.8021	0.8197	0.9342
UT	0.7290	0.7640	0.9290
VT	0.9446	0.9569	0.9692
VA	0.8844	0.8747	0.9399
WA	0.8443	0.8452	0.9579
WV	0.9418	0.9444	0.9735
WI	0.8938	0.8785	0.9603
WY	0.9032	0.8827	0.9619

Source: Vendor school files linked to 2011-2012 SASS school universe file.
N=95,606.

Table A.2 Public, Non-Charter Teacher Coverage Rates by State

State	Vendor A	Vendor B	Vendor C
AL	0.7623	0.6510	0.6512
AK	0.6735	0.5978	0.4317
AZ	0.6594	0.5156	0.5619
AR	0.6888	0.5413	0.4646
CA	0.6787	0.5637	0.5483
CO	0.7692	0.6548	0.6201
CT	0.7796	0.7169	0.6898
DE	0.7476	0.6257	0.5972
DC	0.4561	0.1316	0.3010
FL	0.7565	0.5758	0.5549
GA	0.8255	0.6355	0.6493
HI	0.7851	0.5793	0.3395
ID	0.7899	0.6611	0.6265
IL	0.7327	0.6602	0.6027
IN	0.7125	0.5891	0.5748
IA	0.7311	0.6999	0.6491
KS	0.7439	0.6825	0.6443
KY	0.7307	0.5454	0.5109
LA	0.7588	0.5060	0.5098
ME	0.6394	0.6572	0.5788
MD	0.7568	0.6364	0.6091
MA	0.7924	0.5730	0.5180
MI	0.6572	0.6109	0.5553
MN	0.7185	0.6657	0.6172
MS	0.7271	0.4544	0.4427
MO	0.7211	0.6263	0.6057
MT	0.6826	0.5886	0.5912
NE	0.7303	0.6587	0.6200
NV	0.7997	0.3911	0.4171

NH	0.7627	0.7000	0.7069
NJ	0.7128	0.5731	0.5647
NM	0.6632	0.4994	0.4057
NY	0.7122	0.5004	0.4256
NC	0.7290	0.5573	0.5651
ND	0.7728	0.6540	0.5749
OH	0.6998	0.6323	0.5883
OK	0.6415	0.5454	0.5080
OR	0.7437	0.6288	0.5828
PA	0.6929	0.5878	0.5532
RI	0.7068	0.5896	0.5046
SC	0.7040	0.5569	0.5401
SD	0.7512	0.6821	0.6523
TN	0.7373	0.6338	0.5935
TX	0.7840	0.6407	0.6278
UT	0.7343	0.6113	0.5916
VT	0.6549	0.6447	0.6201
VA	0.6705	0.5832	0.5671
WA	0.7201	0.6557	0.6296
WV	0.8157	0.5039	0.4489
WI	0.7462	0.6890	0.6573
WY	0.7568	0.6319	0.5980

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=335,864.

Table A.3 Teacher Ineligible Rates by State

State	Vendor A	Vendor B	Vendor C
AL	0.2289	0.5344	0.4223
AK	0.3377	0.6017	0.5097
AZ	0.2983	0.5824	0.491
AR	0.3206	0.5698	0.4556
CA	0.3026	0.5054	0.4413
CO	0.2325	0.5193	0.4509
CT	0.2147	0.3237	0.3907
DE	0.2923	0.5692	0.4684
DC	0.4064	0.7283	0.5805
FL	0.247	0.5534	0.449
GA	0.2308	0.5395	0.4193
HI	0.302	0.5212	0.4346
ID	0.268	0.5646	0.3442
IL	0.2892	0.5571	0.4653
IN	0.2872	0.5536	0.4734
IA	0.2504	0.4839	0.3872
KS	0.2906	0.517	0.4207
KY	0.3069	0.5912	0.4723
LA	0.2763	0.5999	0.4777
ME	0.2896	0.5519	0.5041
MD	0.2065	0.5745	0.4459
MA	0.2712	0.5392	0.4389
MI	0.3209	0.5191	0.4796
MN	0.2795	0.5006	0.4442
MS	0.2313	0.5977	0.4795
MO	0.2646	0.5165	0.3191
MT	0.3207	0.3132	0.3643
NE	0.2424	0.5322	0.4204
NV	0.2267	0.5999	0.4655
NH	0.2075	0.4935	0.4252
NJ	0.2987	0.5475	0.4505
NM	0.2862	0.593	0.4828
NY	0.24	0.5324	0.4668
NC	0.2989	0.5947	0.4629
ND	0.2148	0.4807	0.3908
OH	0.3014	0.5543	0.4465
OK	0.3109	0.5167	0.4672

OR	0.2548	0.5416	0.4908
PA	0.3031	0.5365	0.4544
RI	0.2892	0.5229	0.437
SC	0.2887	0.5616	0.4806
SD	0.2426	0.5136	0.4114
TN	0.2649	0.5091	0.448
TX	0.2307	0.512	0.3701
UT	0.1724	0.5157	0.4221
VT	0.3327	0.531	0.4882
VA	0.3334	0.5648	0.4657
WA	0.2956	0.5608	0.459
WV	0.1504	0.5744	0.4981
WI	0.2648	0.5069	0.4705
WY	0.2515	0.56	0.4658

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=335,864.

Appendix B: Teacher Coverage Rates for All Schools

Table B.1 **Teacher coverage rate for all schools, by locale and vendor**

CCD Urban-Centric Locale	Coverage Rate		
	Vendor A	Vendor B	Vendor C
City, Large	0.4885	0.3925	0.3611
City, Midsize	0.5937	0.5207	0.5000
City, Small	0.6538	0.5598	0.5313
Suburb, Large	0.6411	0.5897	0.5469
Suburb, Midsize	0.6829	0.5794	0.5515
Suburb, Small	0.6837	0.6237	0.5798
Town, Fringe	0.6465	0.6151	0.5889
Town, Distant	0.6836	0.6145	0.5843
Town, Remote	0.6834	0.6030	0.5549
Rural, Fringe	0.6645	0.5747	0.5492
Rural, Distant	0.6623	0.5942	0.5590
Rural, Remote	0.6431	0.5503	0.4877

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=390,999.

Table B.2 **Teacher coverage rate, by teacher years of experience and vendor, all schools**

Years of teaching experience	Coverage Rate		
	Vendor A	Vendor B	Vendor C
1st year	0.4596	0.2077	0.2181
2-3 years	0.5536	0.3736	0.4147
4-19 years	0.6567	0.5876	0.5558
20 or more years	0.6393	0.6481	0.5606

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=390,999.

Table B.3 **Teacher coverage rate, by full-time status and vendor, all schools**

	Coverage Rate		
Teacher status	Vendor A	Vendor B	Vendor C
Full-time	0.6534	0.5691	0.5351
Part-time	0.3596	0.4135	0.3732

Source: Vendor teacher files matched to 2011-2012 SASS teacher sample frame file.
N=390,999.