Identifying Comparison Districts to Andover Pubic Schools

Marcus Waldman
marcus.waldman@andoverma.us

Executive Summary

This report provides my recommendations for selecting ten comparable school districts to Andover Public Schools. These ten districts include Arlington, Chelmsford, Franklin, Holliston, Natick, Needham, North Andover, Wellesley, Westwood, and Winchester. I arrive at my recommendations by comparing Andover's student demographic population to the student demographics at 24 other Boston-area school districts. Specifically, my analysis considers students' race and ethnicity, students' socioeconomic status, the proportion of the student population whose first language is not English, and the proportion of the student population that have a diagnosed disability or receive special education services. I compare the districts most similar according to statistical analysis with districts listed as "comparable" by the Massachusetts Department of Elementary and Secondary Education ("DESE") alongside a set of 16 districts labeled as the Sweet Sixteen comparison communities by the town. Encouragingly, I find agreement between districts identified as most "comparable" using the statistical techniques and the latter two sources.

1.0 Motivation

Many factors influence a school district's observed performance on key indicators such as achievement on standardized tests, graduation rates, dropout rates, college-persistence, etc. Perhaps the most studied factor is the income of the families that the district serves. School districts that tend to serve a high proportion of affluent students tend to have higher test scores, a greater proportion of students who graduate and persist through college, and decreased dropout rates.

In addition to family income or economic status, other factors influence student achievement as well. Language barriers when learning English are associated with lower standardized test scores, especially when measuring reading comprehension. Moreover, a school district may gain a positive reputation for serving students with diagnosed learning disabilities. If families of children with disabilities disproportionately choose to send their children to a district, then that school district may have lower average achievement scores simply because that district serves a larger share of students with disabilities.

Given that many of the factors that are predictive of student achievement and attainment are outside the district's control, *evaluating* a district's performance is a complicated exercise. Most education researchers would argue that comparing college persistence rates of a relatively affluent district like Andover Public Schools to a district where a high percentage of students are experiencing poverty would likely be an unfair comparison. A direct comparison in evaluating a district's performance would not be appropriate because students in poverty face more barriers towards obtaining a college degree. Colloquially stated, the flaw in comparing Andover and a

high-poverty district is that we are "comparing apples-to-oranges" because the student populations are different on a measure (i.e. family income) that is predictive of educational attainment but is outside the control of the district itself.
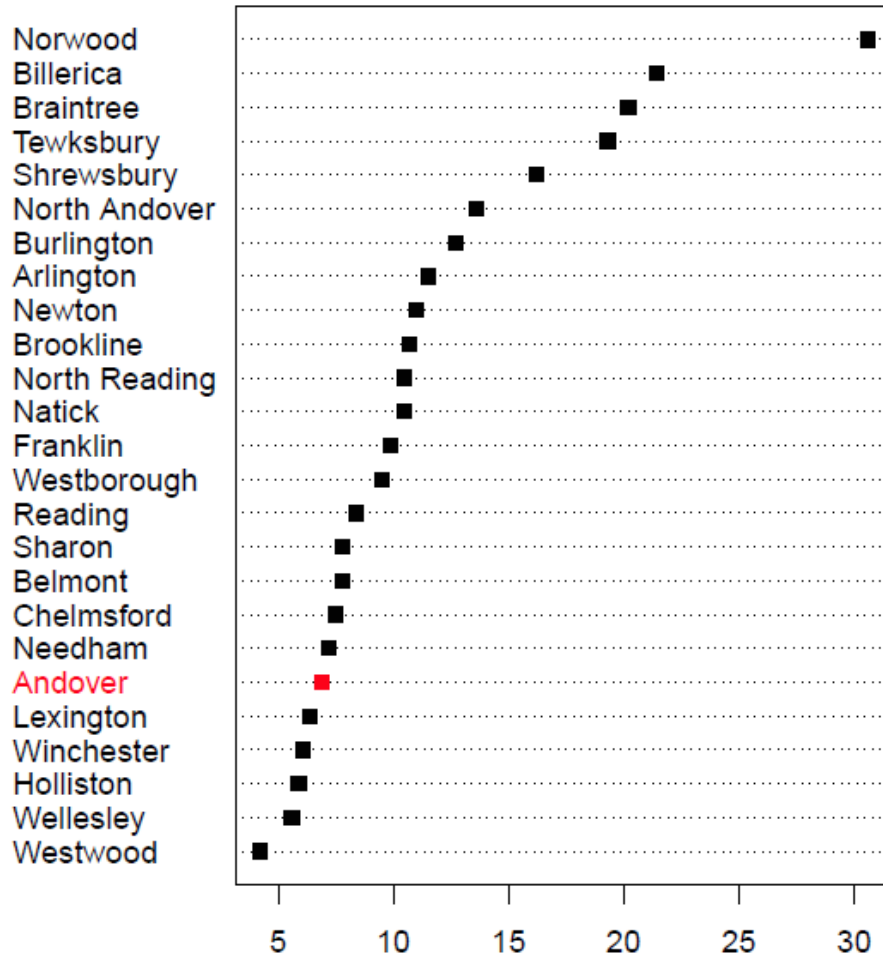
To evaluate Andover's performance on key student achievement and educational attainment indicators by comparing Andover Public Schools to other districts, we must first ensure that comparisons are "apples-to-apples." Specifically, the comparison districts should have similar student populations in terms of school district size, economic status, race and ethnicity, percent of students with disabilities, and percent of students whose first language is not English. In this analysis, I compare Andover Public Schools to a set of 24 possible comparison districts.[1]

2.0 Measures and Preliminary Analysis

Economic status. I measure economic status by the percentage of students who are classified by DESE as low income or high needs. DESE classifies a student as low income if the student receives free or reduced price lunch. Moreover, a student is classified as high needs if the student is low income, is an English language learner ("ELL"), or has diagnosed learning disabilities. Additionally, the Massachusetts Department of Revenue provides an additional indicator of economic status through the Equalized Valuations Finalized per capita ("EQV per capita"). The EQV per capita measures the mean property value per thousand people and serves as a useful proxy for a community's wealth.

---

[1] These comparison districts include Arlington, Belmont, Billerica, Braintree, Burlington, Chelmsford, Franklin, Holliston, Lexington, Natick, Needham, Newton, North Andover, North Reading, Norwood, Reading, Sharon, Shrewsbury, Tewksbury, Wellesley, Westborough, Westwood, and Winchester. Note that Acton-Boxborough was originally included as a possible comparison district; however, it was removed because the towns of Acton and Boxborough report separate mean property values to the MA Department of Revenue.
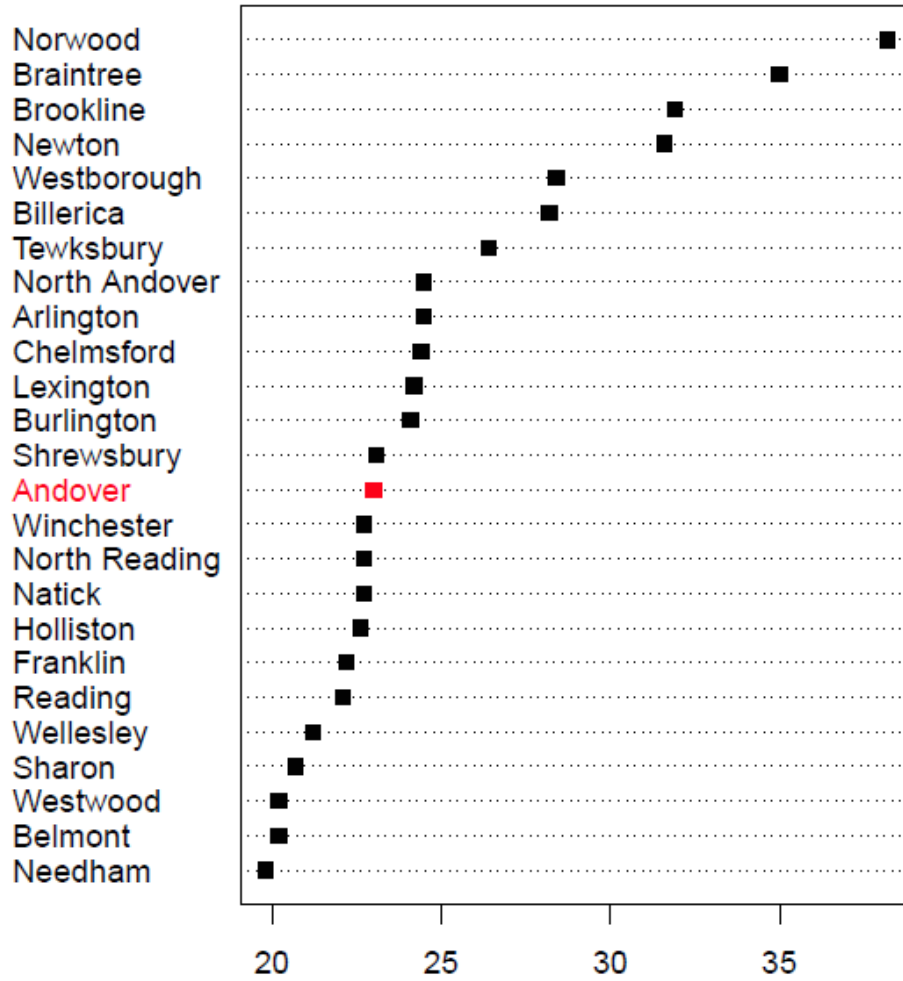
% Low Income

SY2014−2015
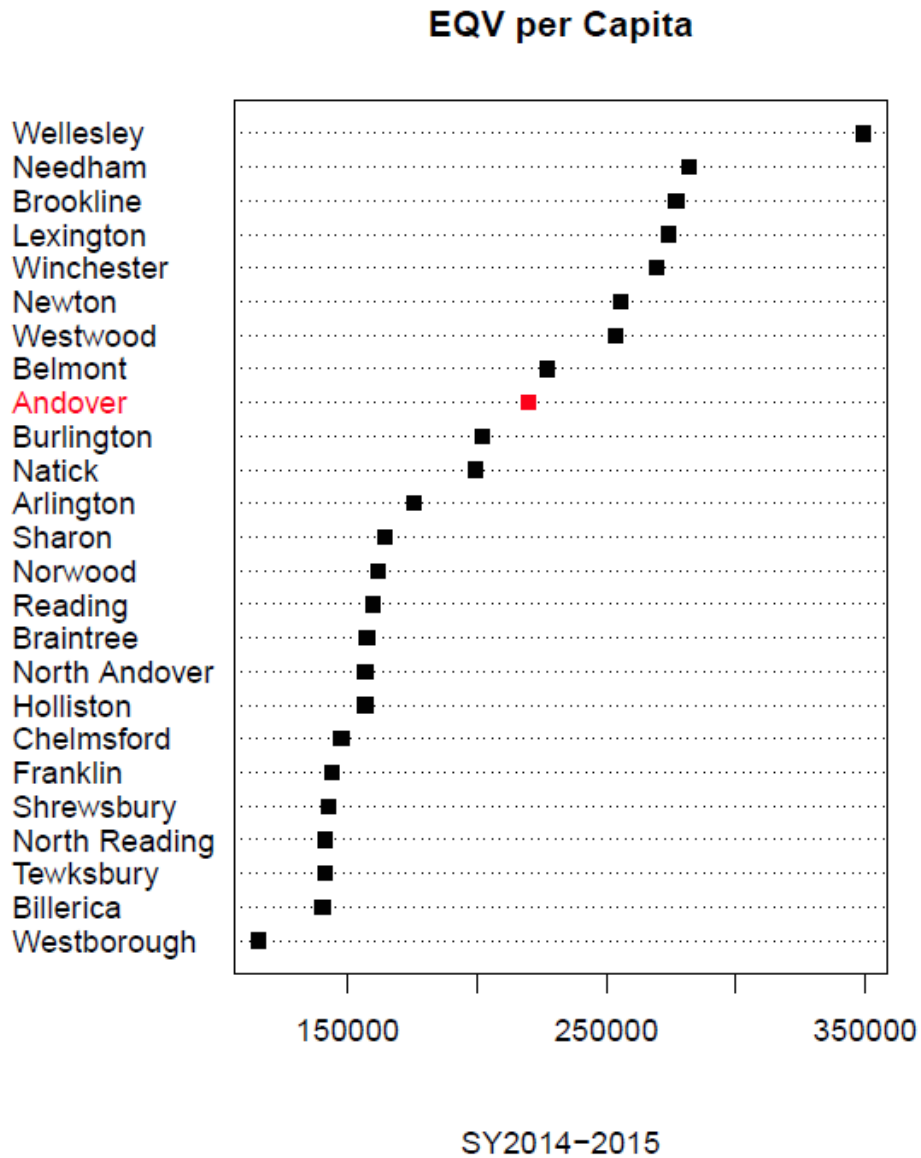
Figure 1

% High Needs

SY2014-2015

Figure 2

EQV per Capita

Wellesley
Needham
Brookline
Lexington
Winchester
Newton
Westwood
Belmont
Andover
Burlington
Natick
Arlington
Sharon
Norwood
Reading
Braintree
North Andover
Holliston
Chelmsford
Franklin
Shrewsbury
North Reading
Tewksbury
Billerica
Westborough

150000   250000   350000

SY2014–2015

Figure 3

Figures 1 and 2 report the proportion of low-income and high-needs students, respectively. Andover ranks at the bottom third in terms of the district's low-income student population but near the median in terms of the district's high needs population. Communities with high property values may continue to serve a high proportion of low-income students. According to Figure 3, eight of the comparison school districts are situated in communities with greater property values, implying that Andover is near the top third in terms of community wealth.

Disability classification. Figures 4 and 5 display the proportion of students with disabilities and/or enrolled in special education in each of the 24 school districts.
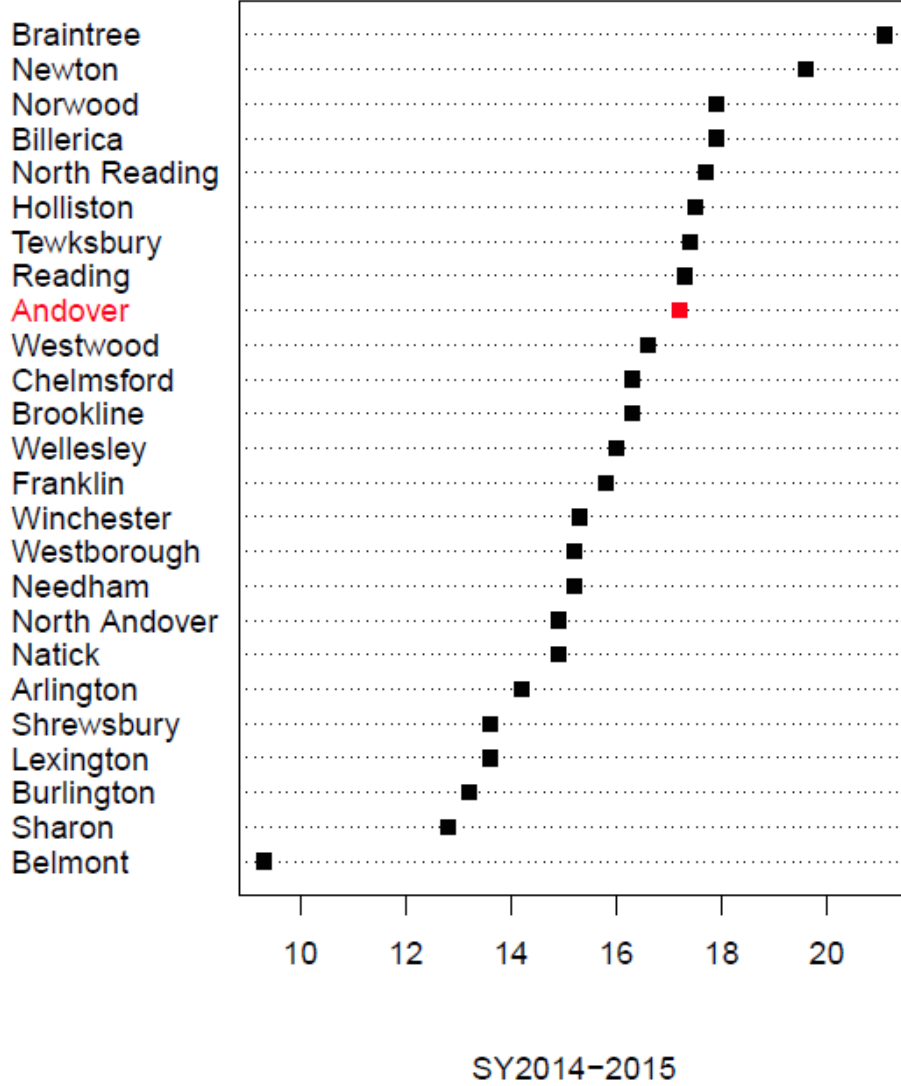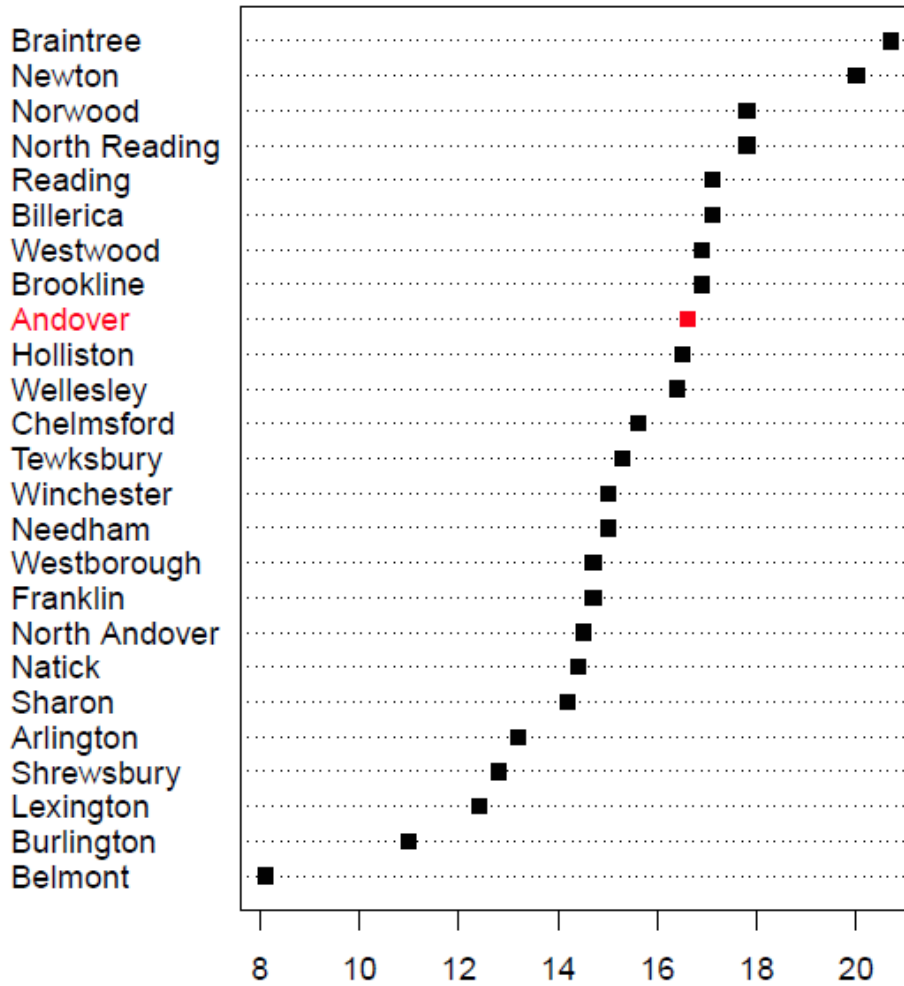
% Students w/ Disabilities

SY2014–2015

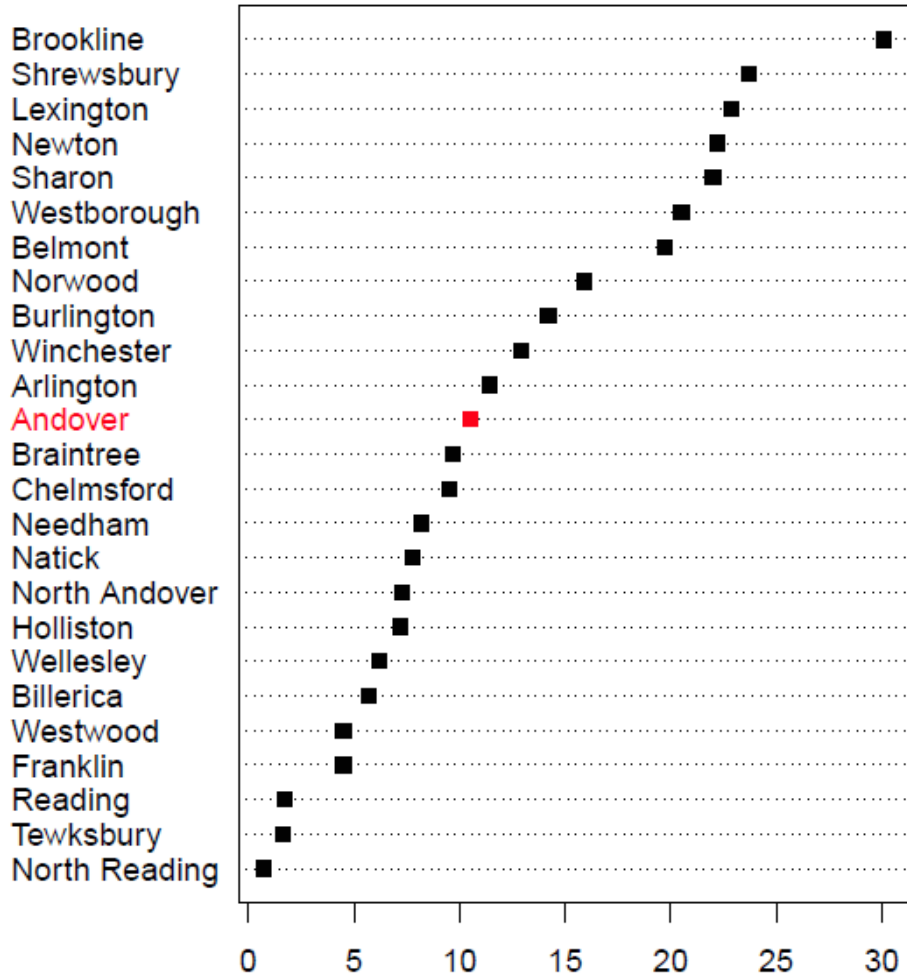*Figure 4*

% Enrolled in Special Education

SY2014−2015

Figure 5

Andover ranks above the median in the proportion of students with a documented disability and/or enrolled in special education. Newton and Braintree, however, serve a disproportionately large share of students with documented disabilities.

English language status. I use three measures provided by DESE to classify students based on English language status. These include the percentage of students classified as "limited English proficient" ("LEP"), the percentage classified as an "English language learners" ("ELL"), and the percentage of students whose first language is not English.
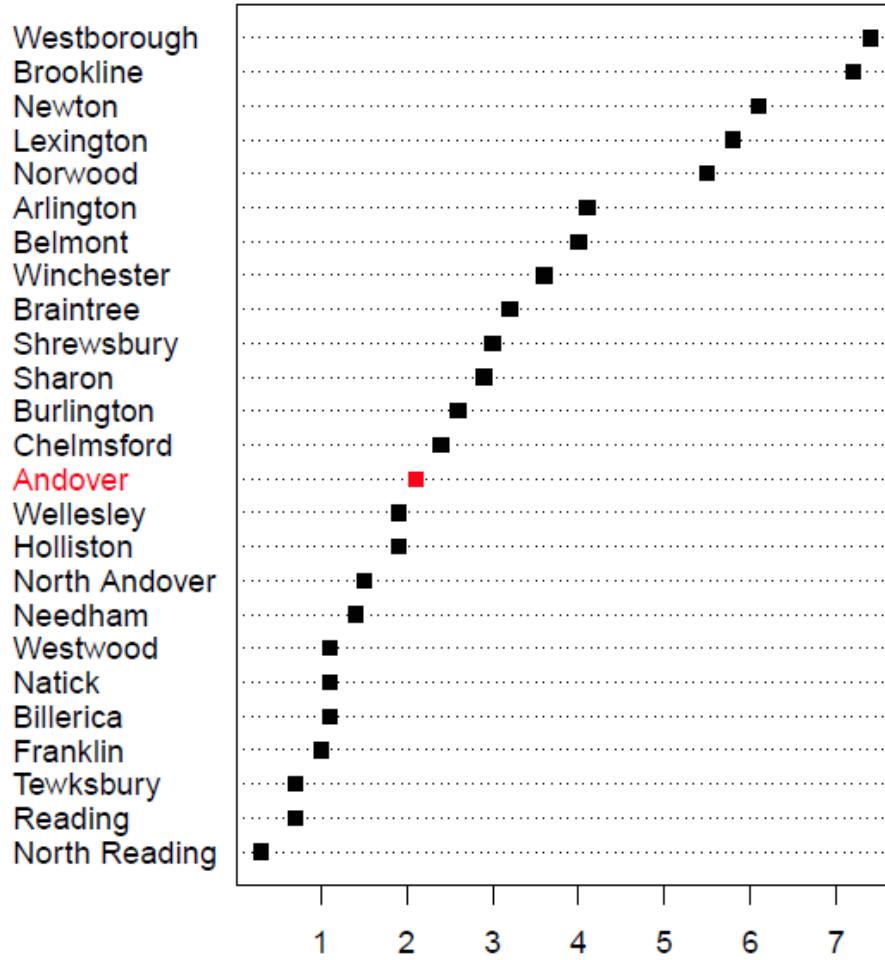
% First Language is not English

SY2014−2015

Figure 6

Figure 7

% English Language Learners

Brookline
Westborough
Newton
Norwood
Lexington
Belmont
Arlington
Braintree
Winchester
Shrewsbury
Burlington
Chelmsford
Sharon
Andover
Holliston
Wellesley
Needham
Natick
North Andover
Franklin
Westwood
Billerica
Tewksbury
Reading
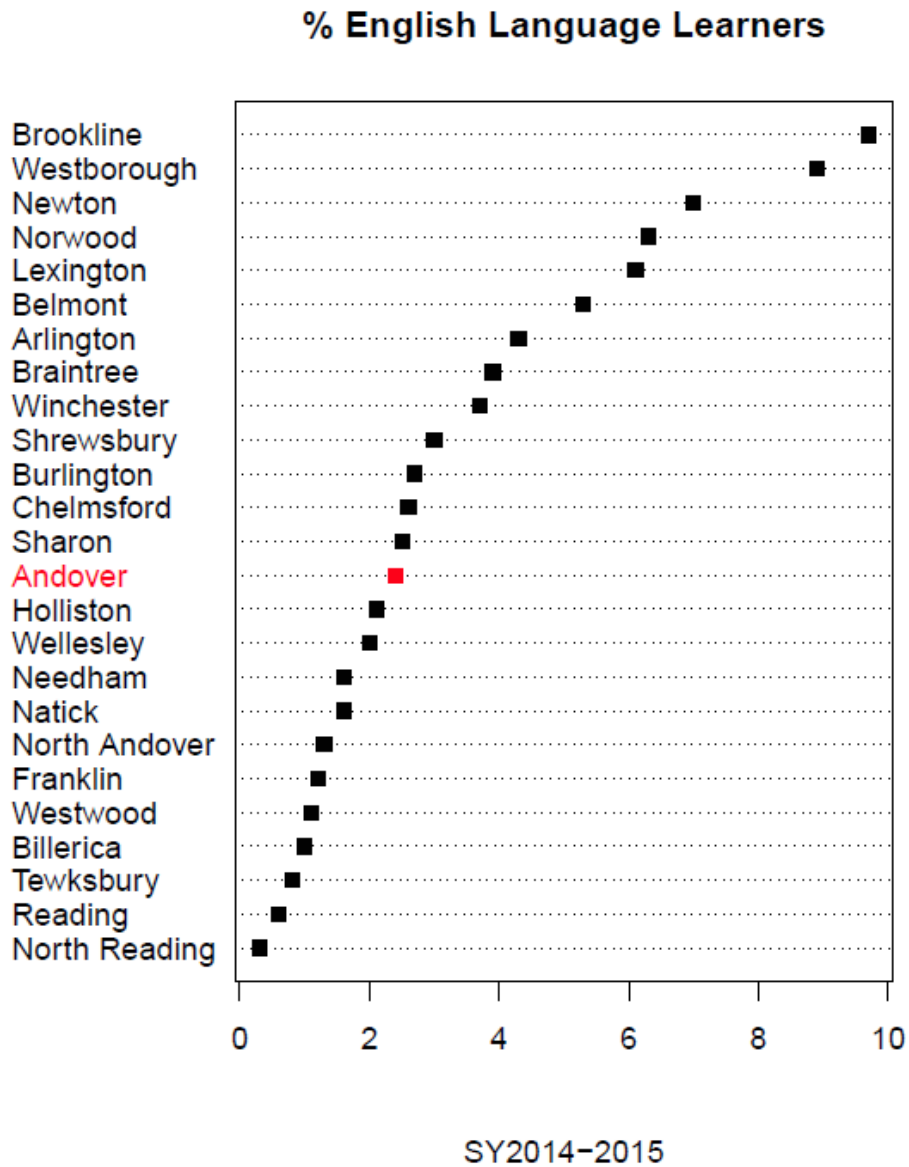North Reading

0   2   4   6   8   10

SY2014−2015

Figure 8

Figures 6 and 7 display the proportion of students who are classified as LEP or ELL, while Figure 8 displays the proportion of the student population whose first language is not English. Andover ranks near the median across all three measures. On the other hand, Brookline serves a disproportionately large share of students whose first language is not English. Additionally, Westborough serves a substantial number of students who are classified as LEP.

Race and ethnicity. Figure 9-12 display the proportion of the district populations who identify as African-American, Asian, White, and Hispanic, respectively.
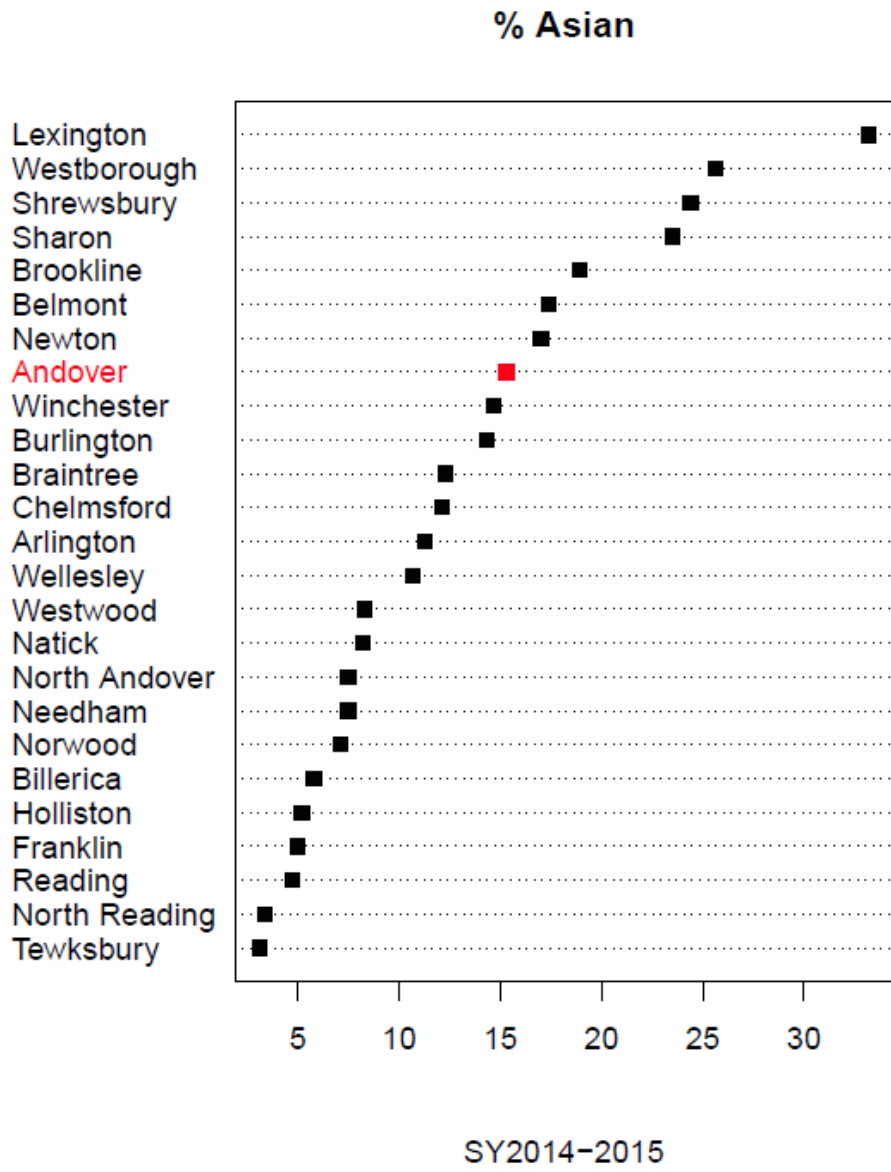
10

% African-American

SY2014−2015

Figure 9

% Asian

SY2014−2015

Figure 10

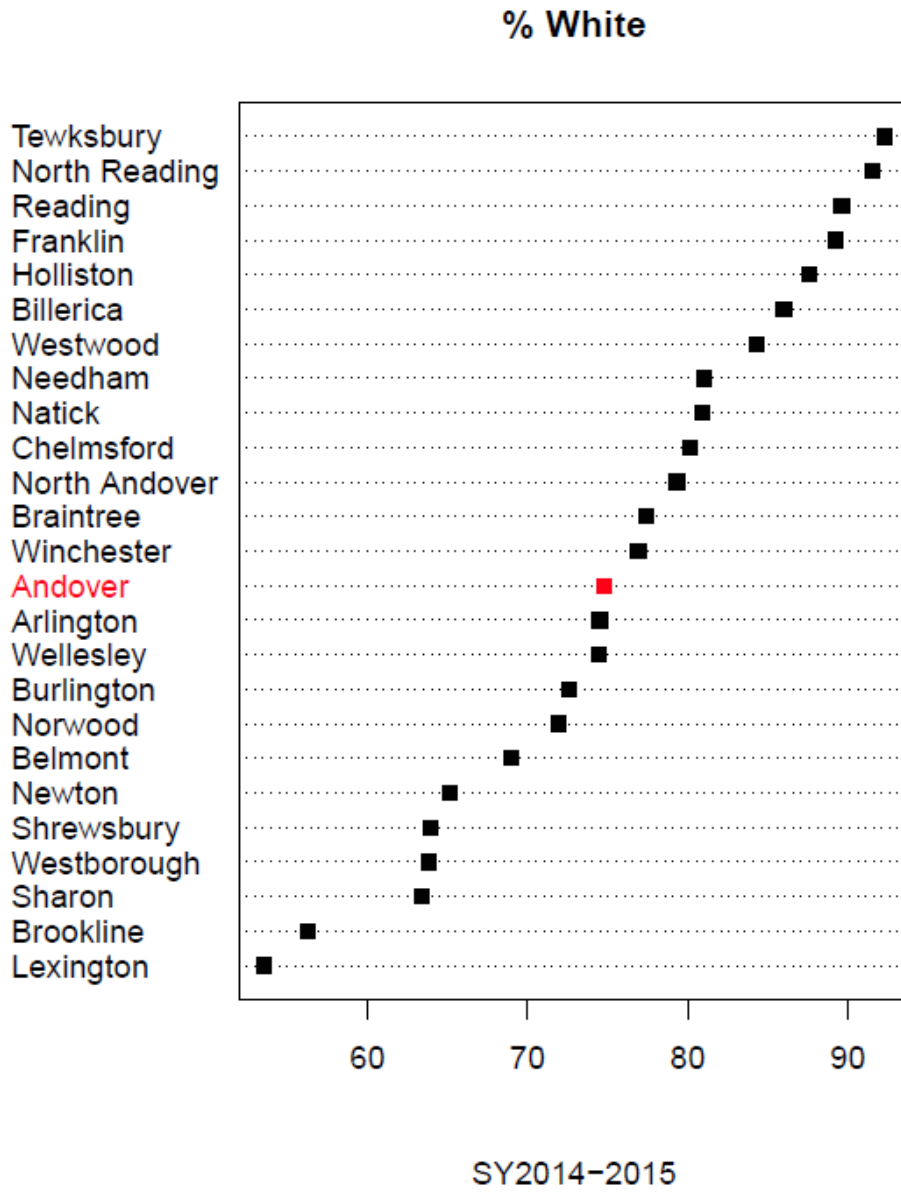## % White

SY2014-2015

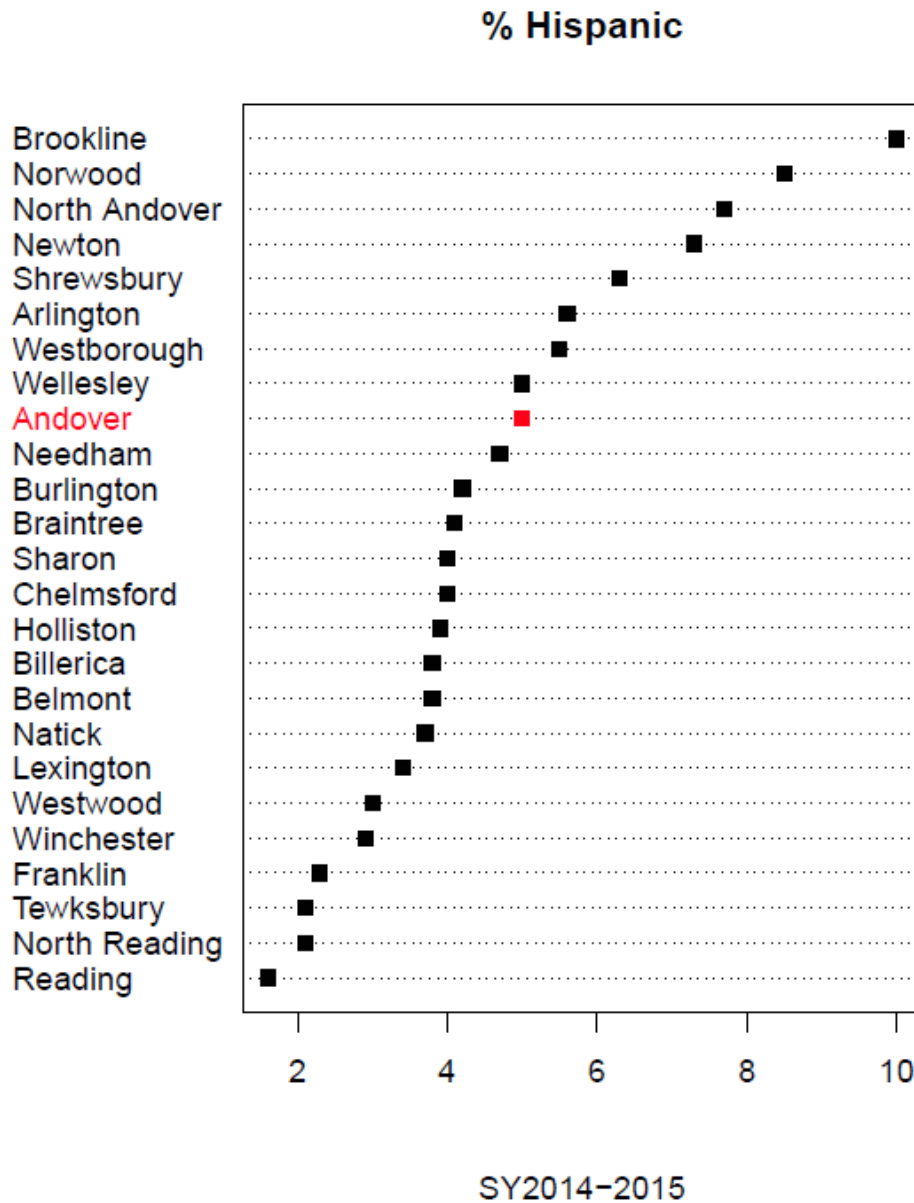*Figure 11*

% Hispanic
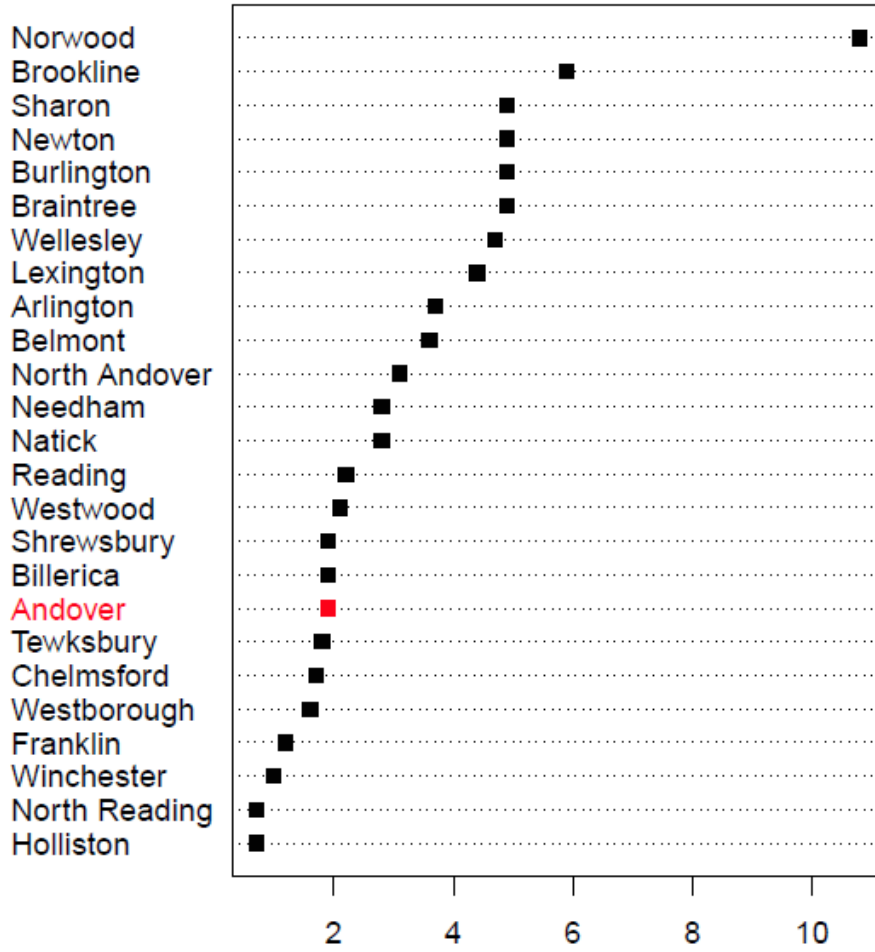
SY2014-2015

Figure 12

As with the racial composition of the other comparison districts, Andover is a majority white district with nearly three-quarters of students identifying as white. Moreover, Andover serves a relatively large share of Asian students and a relatively small share of African-American students. The district's student composition ranks in the top- and bottom-third in both of these measures. The proportion of students who identify ethnically as Hispanic is also above the median.

The student composition in Lexington and Norwood differ substantially from the remaining comparison district. Lexington serves a much larger share of Asian students, while Norwood serves a much larger share of African-American students. Indeed, nearly 33% of students

14

identify as Asian in Lexington, while about 10% of students identify as African-American in Norwood.

School district size and per-pupil expenditures. Figure 13 displays the total student population in each district, and Figures 14 compares Andover Public Schools' per pupil expenditures to the 24 comparison districts.
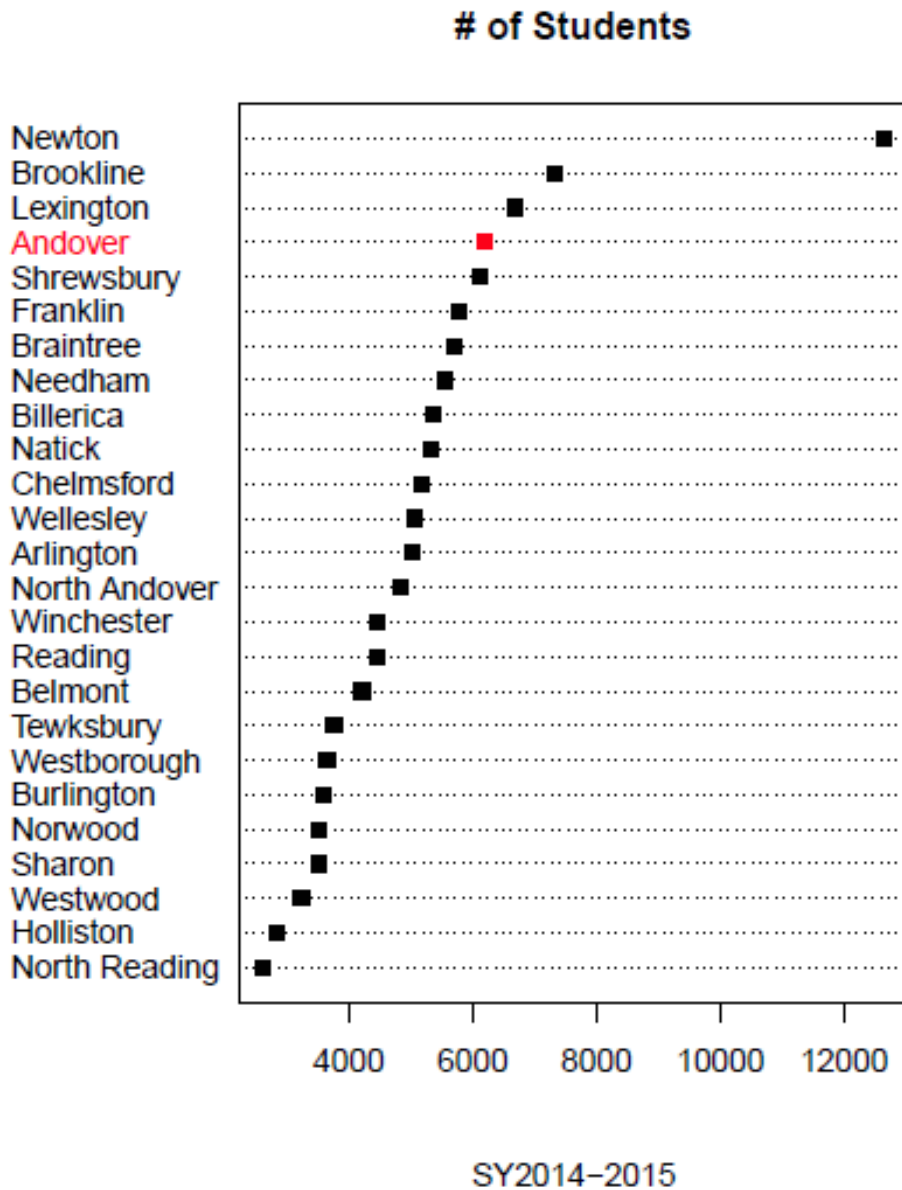

Figure 13

Andover Public Schools is a relatively large school district, ranking in the top four based on the number of students the district serves. Newton is a clear outlier, serving nearly 50% more students than the next largest school district in the comparison set.
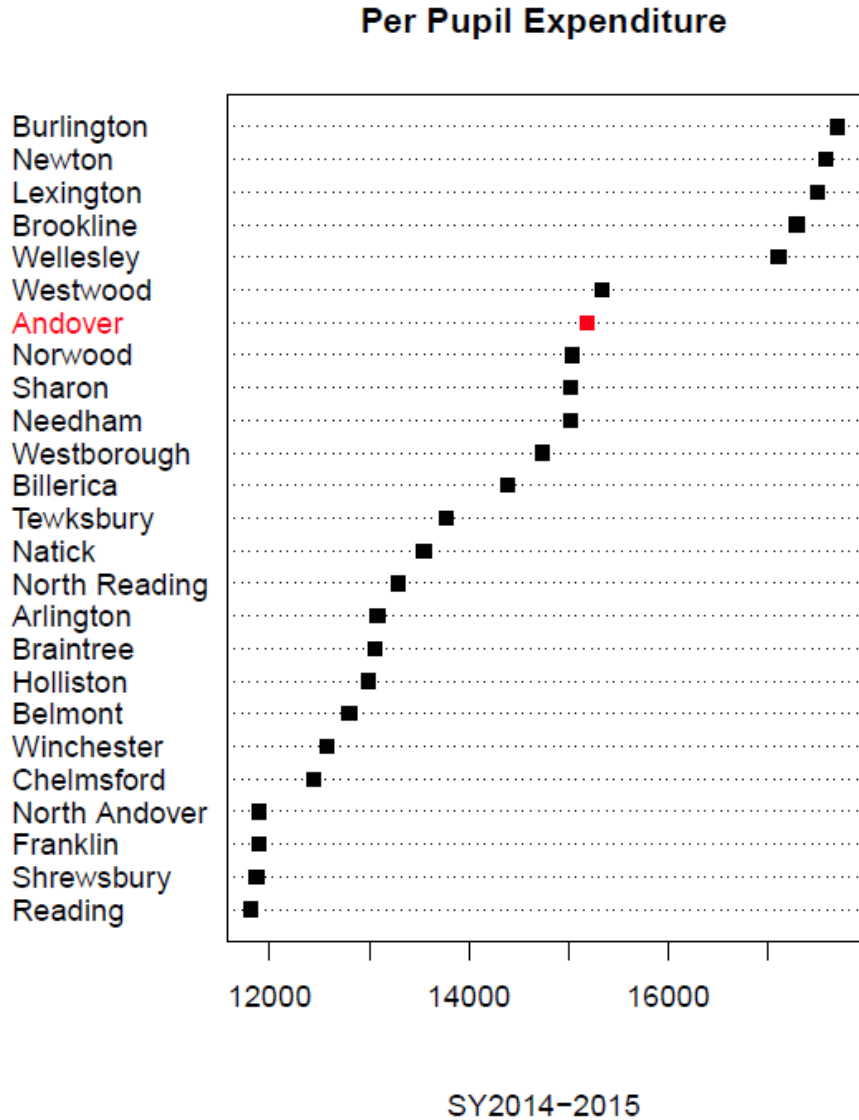
Per Pupil Expenditure

SY2014−2015

Figure 14

Figure 14 shows that Andover ranks in the top-third in per-pupil expenditures. The following districts spend nearly 10% more than Andover: Wellesley, Brookline, Lexington, Newton, and Burlington. On the other hand, North Andover, Franklin, Shrewsbury, and Reading each have per- pupil expenditures that are only 80% of Andover Public Schools' per-pupil expenditures.

Per-pupil expenditures correspond to local property values. Figure 15 displays the per-pupil expenditures as a function of EQV per capita. As expected, I find a positive correlation between

16

per pupil expenditures and property values. Moreover, it is worthwhile noting that Andover Public Schools is located near the best fit line, suggesting that the district devoted a commensurate amount of funds given the average property values of the district. Natick and Belmont were two districts with similar property values, but both spent less per pupil than Andover. On the other hand, Burlington and Newton are two districts with similar property values to Andover, but these districts spent over $2,000 more per pupil.
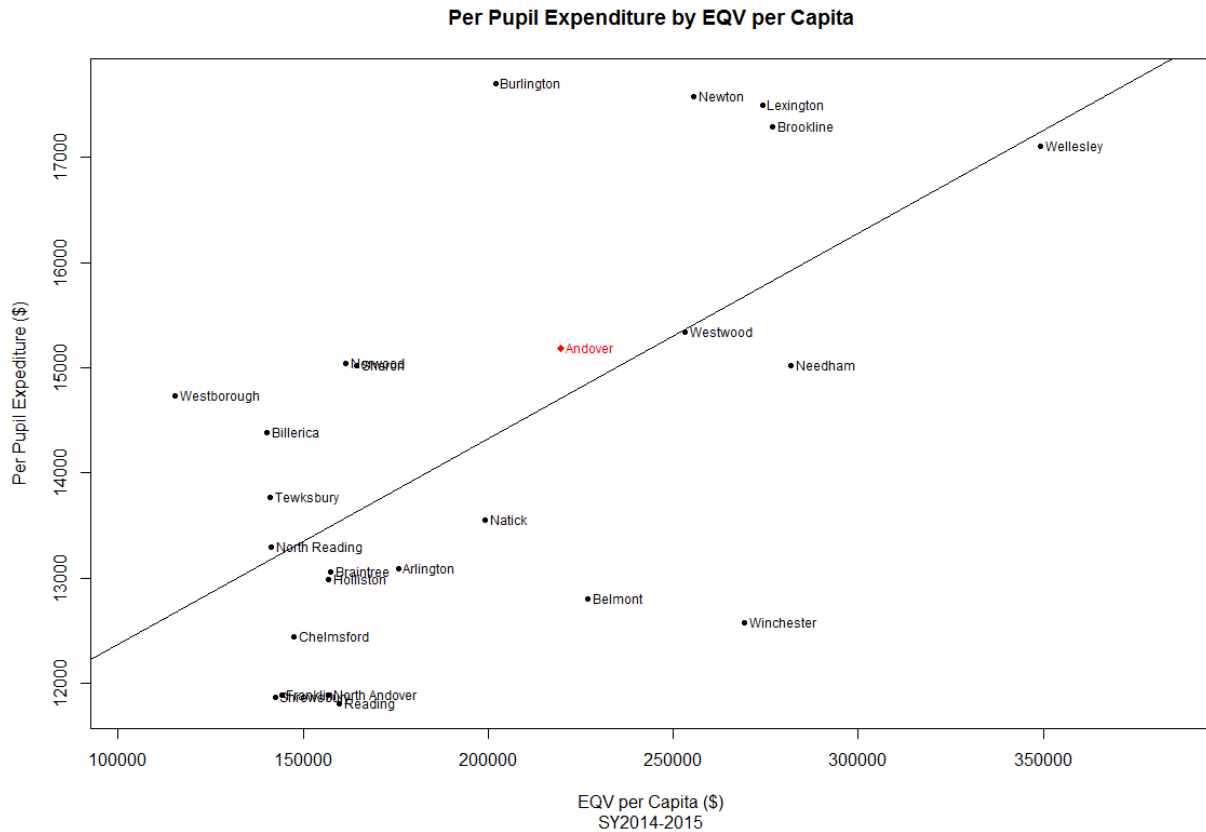

Figure 15

To summarize, Figures 1-15 suggest that Andover Public Schools is a large school district that serves relatively affluent families. The student populations at Lexington, Norwood, and Brookline do not reflect the student population at Andover and, therefore, should not be used to make comparisons across performance indicators.

3.0 Methods

3.1 Dissimilarity metric using PCA. I developed a dissimilarity metric to contrast Andover with the 24 comparison districts. The dissimilarity metric is calculated using a statistical algorithm called principal components analysis ("PCA"). The PCA method produces a set of scores for each district using the socioeconomic and demographic variables in the data.[2] The goal of the

---

[2] I excluded the EQV per capita measure in conducting the PCA analysis because the EQV measure distorted the dissimilarity metric.

PCA method is to explain the greatest amount of variance by developing the fewest number of scores.
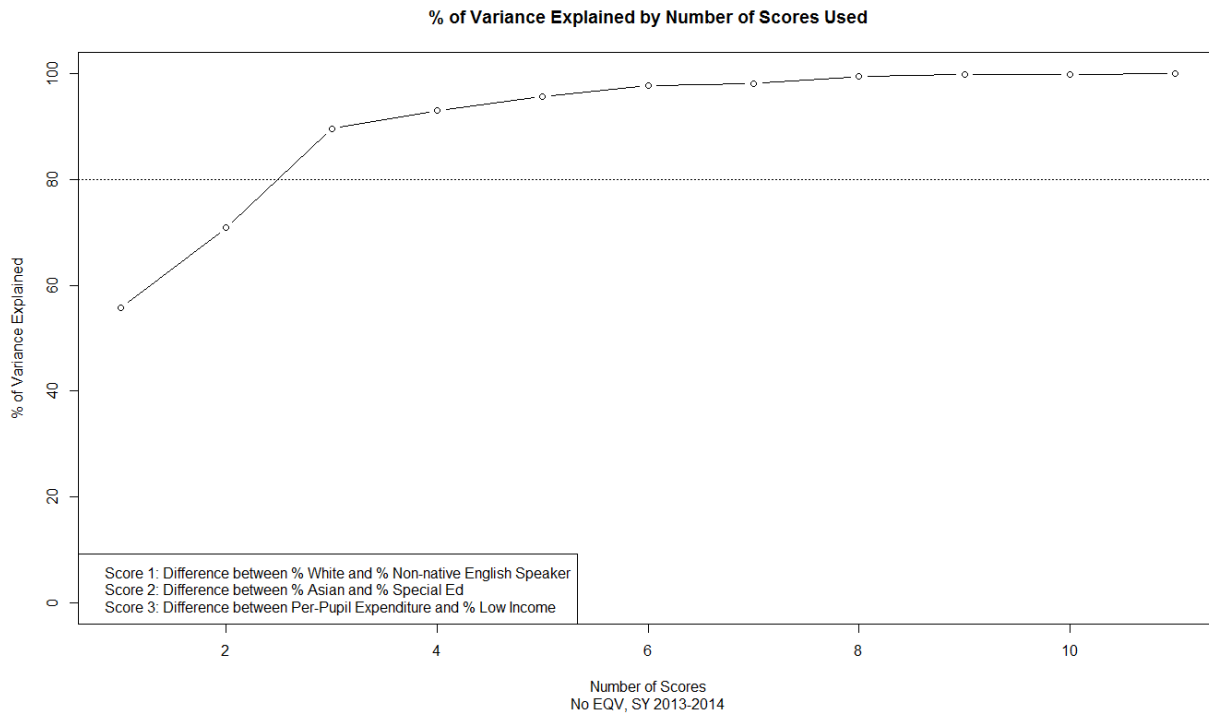

% of Variance Explained by Number of Scores Used

Score 1: Difference between % White and % Non-native English Speaker
Score 2: Difference between % Asian and % Special Ed
Score 3: Difference between Per-Pupil Expenditure and % Low Income

Number of Scores
No EQV, SY 2013-2014

Figure 16

For the PCA method to be valid, we must include a sufficient number of scores to explain a minimum threshold of the variance observed in the data. Explaining 80% of the variance is an accepted lower bound. Figures 16 shows that we must include three scores to reach the 80% threshold; therefore, three scores are included in calculating a composite dissimilarity metric. A rule-of-thumb interpretation of each score is provided in Table 1 below.[3]

Table 1: Interpreting the three scores produced by the Principal Components Analysis

Score #	Interpretation
1	Difference between a district's % white and % non-native English speaking student population
2	Difference between a district's % Asian and % special education
3	Difference between a district's per pupil expenditure and the proportion of low-income students in the district

Once the three scores for each district were known, I created the dissimilarity metric by calculating the Euclidean distance between that district's scores with Andover's scores, weighted

---

[3] This is a rough interpretation and should not be taken too literally. The exact formula is more complex.

by the percentage of variance in the data explained by the score (see Equation 1 below). This weighting scheme ensures that the scores that discriminate most across districts receive the most weight.

$$\text{PCA: } dissimilarity = \sqrt{w_1 \Delta Score_1^2 + w_2 \Delta Score_2^2 + w_3 \Delta Score_3^2} \qquad (1)$$

3.2 Dissimilarity metric using propensity scores. In addition to the dissimilarity metric created using the PCA method, I also developed a dissimilarity metric using a technique called propensity score analysis. The propensity score approach attempts to predict the probability that a randomly sampled district is Andover Public Schools, given the district's student demographic variables. The more likely that the district student population reflects Andover, the greater the probability that the randomly selected district is Andover. Thus, the districts that are more similar to Andover will have propensity scores near one, while districts that are least similar to Andover will have propensity scores near zero.

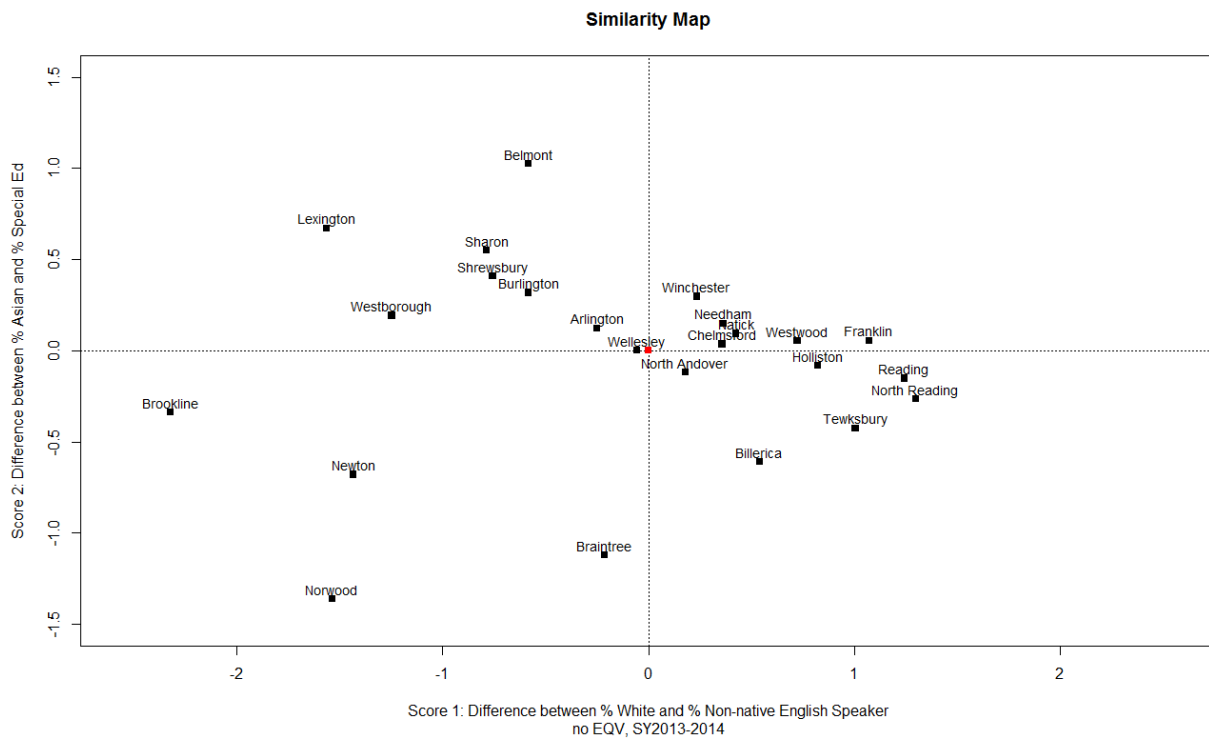## 5.0 Similarity Map Analysis

Figure 17

Figure 17 plots a district's second PCA score against the district's first PCA score. Districts near the origin—indicated by the red dot—have student populations that are similar to Andover. In contrast, districts furthest from the origin are least similar to Andover. Thus, I call Figure 17 a similarity map because a district's spatial relationship to the origin depicts that district's similarity to Andover. Consistent with the findings in Section 2, Norwood, Brookline and

19

Lexington are furthest from the origin, indicating high dissimilarity with Andover Public Schools. Moreover, Braintree, Westborough, and Newton are also situated relatively far from the origin, suggesting that these districts are not good comparisons to Andover relative to the other districts.
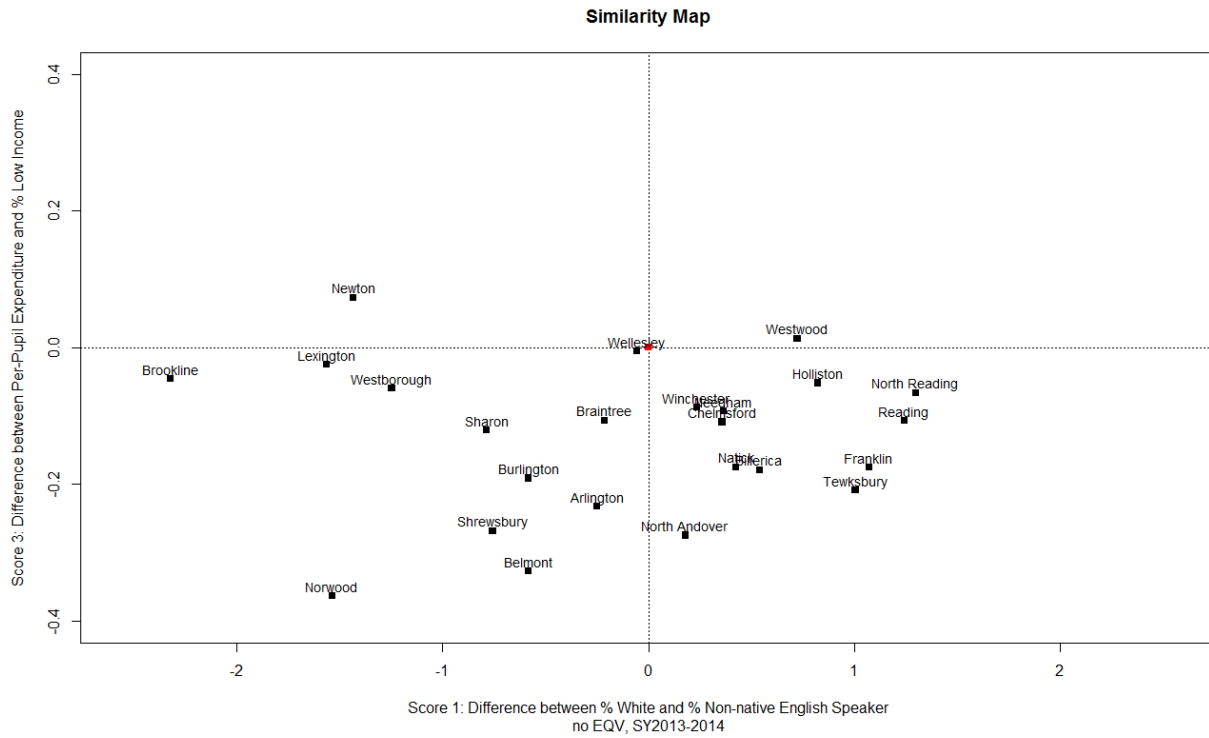

Figure 18

Figures 18 and 19 plot a district's third PCA score against the district's first and second PCA score, respectively. The distance of each district from the origin in Figure 18 is similar to that found in Figure 17. Belmont, however, is situated far from the origin in Figure 19, providing evidence that Belmont is not a good comparison district.
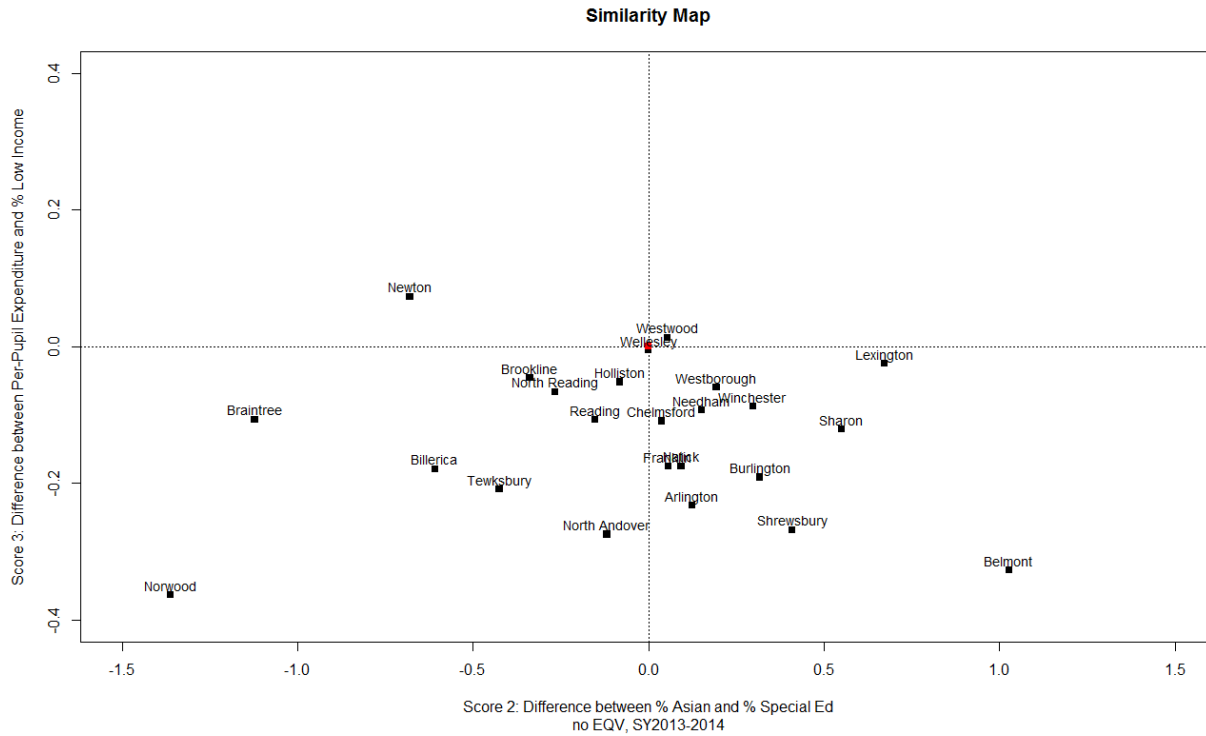
Figure 19

5.0 Dissimilarity Metric Analysis

The second and third column of Table 1 below display the rankings of the dissimilarity metric using the propensity score approach and the PCA approach. Higher ranking values imply greater dissimilarity to Andover, while smaller ranking values imply greater similarity to Andover. Specifically, the Spearman-rank correlation between the dissimilarity metric generated PCA is 0.96 (perfect correlation would be 1), suggesting considerable agreement between the two methods.

The fourth and fifth column in Table 1 indicate whether the district is considered a comparable district according to the Sweet Sixteen and DESE, respectively. Most of the districts listed in the Sweet Sixteen and DESE groups also have low dissimilarity rankings. This indicates that there is general consistency across the dissimilarity metrics and the districts included in the latter two sources.

In selecting which districts are most similar, I use the following criteria to balance the quantitative analysis with the Sweet Sixteen and DESE's lists**: A district is considered comparable if the district is either (a) both in the Sweet Sixteen list and listed as a comparison according to DESE and/or (b) ranked as the top five most similar according to both dissimilarity metrics.** Ten districts that meet this selection criterion include: 1) Arlington, 2) Needham, 3) Winchester, 4) Natick, 5) Chelmsford, 6) Westwood, 7) Holliston, 8) Franklin,

9) Wellesley, and 10) North Andover.

Table 2

District	Propensity Score Ranking	PCA Ranking	District is in Sweet Sixteen List	District is listed as Comparable according to DESE
Arlington	1	3	FALSE	FALSE
Needham	2	6	TRUE	TRUE
Wellesley	3	1	FALSE	TRUE
Winchester	4	5	TRUE	TRUE
North Andover	5	2	TRUE	FALSE
Natick	6	7	TRUE	TRUE
Chelmsford	7	4	TRUE	TRUE
Burlington	8	8	FALSE	FALSE
Westwood	9	9	TRUE	TRUE
Shrewsbury	10	12	FALSE	FALSE
Sharon	11	13	FALSE	FALSE
Holliston	12	10	TRUE	TRUE
Billerica	13	11	FALSE	FALSE
Westborough	14	19	FALSE	FALSE
Franklin	15	14	TRUE	TRUE
Braintree	16	16	FALSE	FALSE
Belmont	17	17	FALSE	FALSE
Tewksbury	18	15	FALSE	FALSE
Reading	19	18	FALSE	TRUE
North Reading	20	20	FALSE	FALSE
Newton	21	21	FALSE	FALSE
Lexington	22	22	FALSE	FALSE
Norwood	23	23	FALSE	FALSE
Brookline	24	24	FALSE	FALSE