Memo to:
Massachusetts Program Administrators New Construction Subcommittee

Prepared by: Nathan Caron, DNV GL

Copied to:
Chad Telarico, DNV GL

Date: January 24, 2020

Massachusetts non-residential new construction EUI baseline study – revised results

1 INTRODUCTION

This memo presents revised results, conclusions, and considerations for the Massachusetts Non-residential New Construction EUI Baseline Study. This study's primary objective was to assess whether the MA Data Warehouse maintained by DNV GL can be used to:

- Estimate energy use intensity (EUI) baselines for various building types
- Identify supplemental information that may be needed to support the creation of the EUIs
- Conclude whether the above methods will work

To perform this study, DNV GL used the Program Administrators' (PAs') billing and tracking data, MA L3 tax data, and Boston tax data housed in the MA Data Warehouse to estimate EUIs for buildings identified as non-residential true new construction.

1.1 Analysis revisions

DNV GL presented initial study results to the PAs and EEAC Consultants in a memo dated July 15, 2019. Upon review of these results, the PAs and EEAC Consultants requested the following revisions to the analysis:

- Substantially increasing the size of the analysis population by extending the time frame of the analysis to include new construction from 2010-2018. The initial analysis limited the population to buildings constructed from 2017-2018.
- Further breaking apart building use categories by three square footage categories to better target EUIs to buildings that would be eligible for the PAs' new construction initiative. These categories include <10,000 sqft, 10,000-50,000 sqft, and >50,000 sqft.
- Switching from separate kWh/sqft and therm/sqft EUIs to a combined kBTU/sqft metric
- Dropping very small accounts, including those consuming <2,000 kWh or <70 therms
- Adding residential consumption data to the calculation of EUIs for multifamily buildings to avoid EUIs capturing common area consumption only
- Handling outlier EUIs through two methods: capping of outliers and clustering of EUIs. We discuss these methods in detail in the body of the memo.

The current memo provides the results of the revised EUI analysis.

1.2 Conclusions

With revisions to this analysis, most notably increasing the analysis population from 2017-2018 to 2010-2018, and the handling of outlier EUIs, DNV GL was able to generate less variable EUI estimates for most building-use size categories. These results are also reasonably close to (and usually somewhat lower than) the average EUI estimates seen in the Northeast census region Commercial Buildings Energy Consumption Survey (CBECS) data from 2012.

Despite the changes, the analysis still contains very small (or no) populations of buildings that are most likely to participate in the PAs' initiative for large new construction buildings: hospitals, colleges, and other campuses. Unfortunately, it is unlikely that we would be able to develop defensible EUIs for those categories given the caveats discussed throughout this memo. Finally, while clustering EUIs within the building use categories substantially reduces variation and highlights differences within categories potentially based on operating hours, occupancy, and specific business activity, unfortunately, this level of detailed information is not available for all buildings without primary research.

1.3 Considerations

Based on the analysis revisions, DNV GL suggests that the PAs consider developing pilot baselines informed by these results for some building use size categories with less variable EUIs. Good candidates would include large (>50,000 sqft) education, large outpatient healthcare, large office, large multifamily, large food sales, large lodging, and large warehouse and storage. We also suggest that the PAs consider additional research to gather detailed information on specific buildings to allow tailoring of EUIs to more granular building use subsegments (e.g. medical imaging and primary care physician versus outpatient healthcare).

2 METHODOLOGY

The following sections describe the data sources we used to estimate EUIs for non-residential new construction buildings.

2.1 Data sources

2.1.1 PA billing and tracking data

This EUI analysis uses the PAs' 2010-2018 billing and tracking data following the extract, transform, and load (ETL) process under the 2018 C&I Data Intake project.

2.1.2 MA L3 and Boston tax data

The Massachusetts L3 tax database provides a standardized "Level 3" parcel mapping dataset containing property (land lot) boundaries and parcel attributes from each community's tax assessor. Boston does not participate under the L3 project but makes similar, and often additional, information available for public use; this information must be accessed and processed separately.

While providing the building square footage information necessary to calculate EUIs, the tax data also classifies parcels by land-use code. There are 10 one-digit (high-level) land use classifications: multiple-use, residential, open space, commercial, industrial, personal property, forest property, agricultural/horticultural, recreational property, and exempt property. Under these high-level classifications, there are 40 two-digit and 262 three-digit classification codes, each providing an increased level of detail.[1]

To allow for comparison across datasets, and to limit the number of building categories in our analysis, DNV GL crosswalked the land use codes to a modified CBECS building category. In addition to the standard CBECS building types, we split out K-12 and colleges/universities under education; we split out hotels and multi-family under lodging; and we added a mixed-use category for buildings without an obvious principal activity. Finally, we also include manufacturing as a building category, which is not captured by CBECS at all. Table 1 shows the complete list of building categories.

Table 1. Modified CBECS classes

Education (K-12)	Lodging (hotel/motel/dormitory)	Public order and safety
Education (colleges/univer**sities)**	Lodging (multi-family)	Religious worship
Food sales	Mercantile (retail other than mall)	Service
Food service	Mercantile (enclosed and strip malls)	Warehouse and storage
Health care	Office	Other

[1] https://www.mass.gov/files/documents/2016/08/wr/classificationcodebook.pdf

(inpatient**)**

Health care (outpatient**)**	Public assembly	Mixed-use

Manufacturing

2.1.3 Tax data caveats - vintage

Some towns have not refreshed their data in the MA L3 tax system for several years, so we are missing potential recent new construction for 81 out of the 351 towns in Massachusetts. Figure 1 lists the towns for which we do not have recent parcel data, and the year during which each town last refreshed its data in the MA L3 tax system. Several of these towns, like Boston, maintain their tax data in a publicly downloadable format separate from the L3 database that is updated yearly; however, because the formats differ from town to town, each would require its own cleaning and processing to be compatible with any analysis.

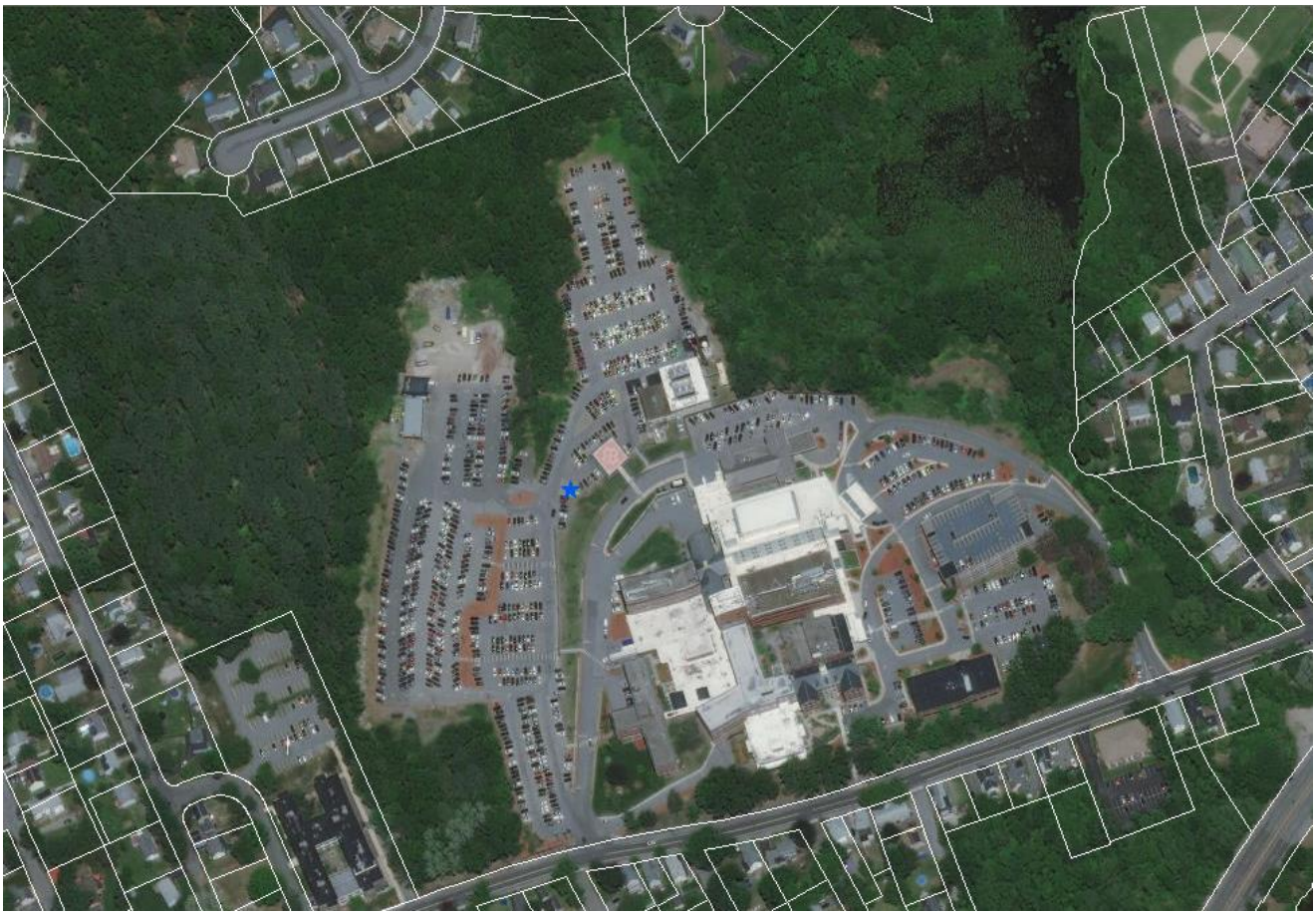
Figure 1. Towns without recent L3 tax updates, as of April 23, 2019

2017	2016	2015	2014	2013	2012	2011	2010
ADAMS	AVON	BOURNE	BOYLSTON	BLANDFORD	CAMBRIDGE	AMHERST	COHASSET
BRIDGEWATER	BROCKTON	DEDHAM	CARVER	BOXFORD	CHESTER	BROOKLINE	HEATH
CHARLEMONT	HOLDEN	GOSNOLD	EAST BRIDGEWATER	BRAINTREE	FRAMINGHAM	GROTON	RUSSELL
CONWAY	MONTGOMERY	LAKEVILLE	GEORGETOWN	HANSON	HANOVER	HINGHAM	
DANVERS	NANTUCKET	LAWRENCE	NEWBURY	IPSWICH	HARWICH	NATICK	
DIGHTON	NEW SALEM	METHUEN	NORTHAMPTON	MARSHFIELD	HOLYOKE	NORWELL	
DRACUT	ORANGE	NEW BEDFORD	ROWLEY	MEDFIELD	LONGMEADOW	SALEM	
EVERETT	PALMER	NEW MARLBOROUGH	WEST NEWBURY	MERRIMAC	MATTAPOISETT	SANDWICH	
FALL RIVER	SHUTESBURY			NORTH ANDOVER	MEDWAY	WESTON	
GARDNER	WASHINGTON			NORTHBRIDGE	OXFORD		
HALIFAX				SALISBURY	SOMERSET		
HATFIELD					WESTFORD		
HAVERHILL					WILMINGTON		
MALDEN							
MONTAGUE							
NEW ASHFORD							
ROCKPORT							
ROYALSTON							
WOBURN							

2.1.4 Tax data caveats - sub-parcels

In addition to potential issues arising from stale data, we are limited to calculating EUIs for only those tax parcels that have no sub-parcels. Occasionally, parcels housing a campus of buildings or condominiums will have separate records in the tax data for each building or unit with separate square footage and year built. Since all the buildings within a parcel share the same tax address, and since matching between the tax and billing data relies on address, it is usually not possible to associate the correct consumption with the correct square footage for EUI calculations. To provide a specific example of one such parcel not permitting accurate EUI calculations using the tax data, we present Lowell General Hospital in Figure 2 with a discussion below.

Figure 2. Lowell General Hospital tax parcel example



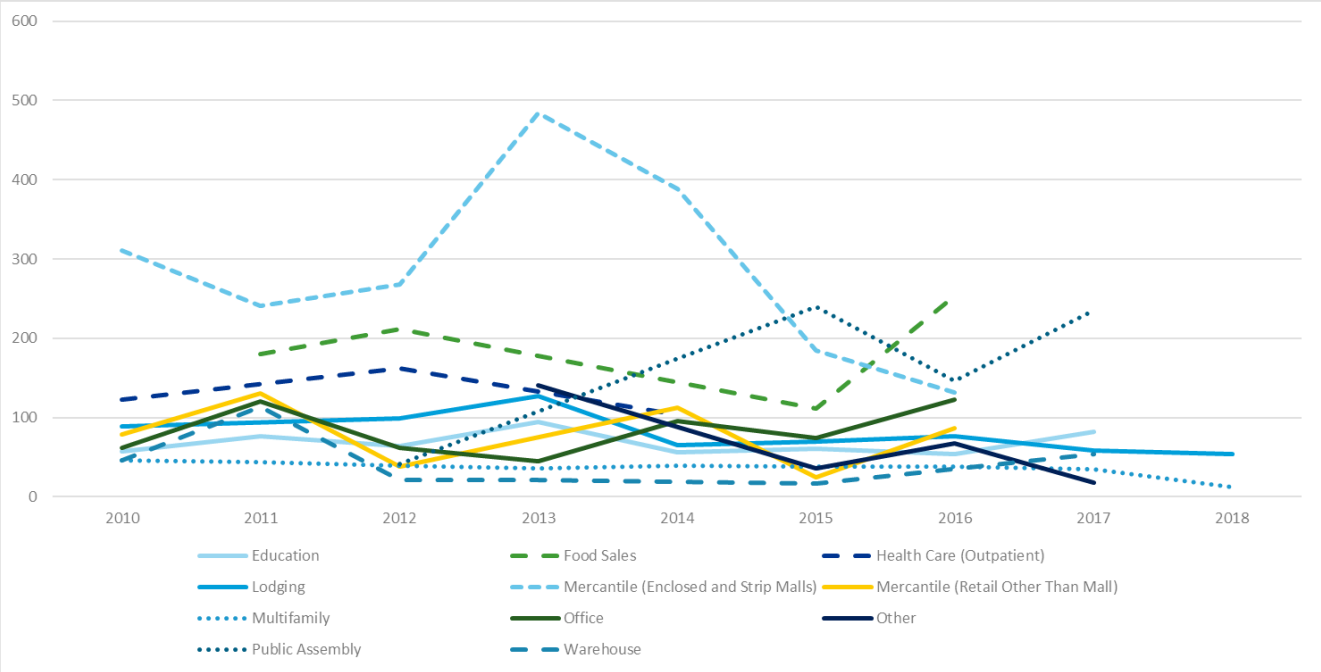
During our analysis, on Lowell General Hospital's tax parcel, we identified a single recent non-residential new construction building of 1,980 sqft built in 2015. However, we were unable to generate an accurate EUI for this building due to the sub-parcel and billing data structure for this site. In all, this tax parcel contains 20 sub-parcels of varying land use and vintage. These include the primary hospital building, built in 1949 with 46,488 sqft. They also include a fast-food restaurant that is listed in the tax data as a building separate from the hospital, but that further research found to be in the hospital's lobby. While there are 20 sub-parcels associated with the hospital, there are only 5 electric and 5 gas accounts in the PAs' billing data, and each account has an address of either 275 or 295 Varnum Avenue, with no additional detail connecting a building or group of buildings to each meter. This makes it impossible to calculate an EUI for the square footage constructed in 2015.

2.1.5 Tax data population

Figure 3 shows the process used to filter the full tax population into our analysis population. Once we filtered the tax data to only non-residential, non-vacant parcels with no sub-parcels, with which we were able to associate some billing data, we were left with the breakdown of building use categories for buildings constructed from 2010-2018 shown in Table 2.[2]

The initial EUI analysis relied on tax data from 2017-2018. This was done to ensure that any EUIs used to inform future baselines would be based on energy use in buildings with the most current new construction practices. However, this decision left an analysis population of only 146 buildings, and consequently, several empty building-use size categories. To try to expand the analysis population, we investigated whether there was any obvious positive or negative trend at the the building-use level in tax-data-calculated EUIs, which might prevent developing defensible averages for current new construction using a longer time series of recent new construction. We show these EUI trends in Figure 4 below. This figure shows that some building use segments increased in EUIs from 2010-2018 while others decreased, and EUIs can vary substantially within a building use segment from year to year. Based on the lack of a definite trend, we felt comfortable expanding the analysis population to include buildings constructed from 2010-2018; this yielded an analysis population of 1,059 buildings.[3]

Figure 3. Average EUIs based on tax data, buildings >50k sqft



[2] Note that there were no colleges/universities or inpatient health care built in 2017 or 2018 based on the L3 tax data. This does not mean that none of these buildings were constructed in MA during that time frame, only that the subset of towns for which we have recent tax data do not contain those building types.

[3] We acknowledge that there is some risk to assuming that the change in energy consumption among buildings between 2003-2012 is the same as any change in energy consumption bewteern 2012-2018, especially given new building codes and energy efficiency programs over the more recent time period; however, this assumption is necessary to provide a reasonable set of new construction buildings to estimate EUIs for each building-use size category.

Figure 4. Restriction of full tax population to 2010-2018 analysis population

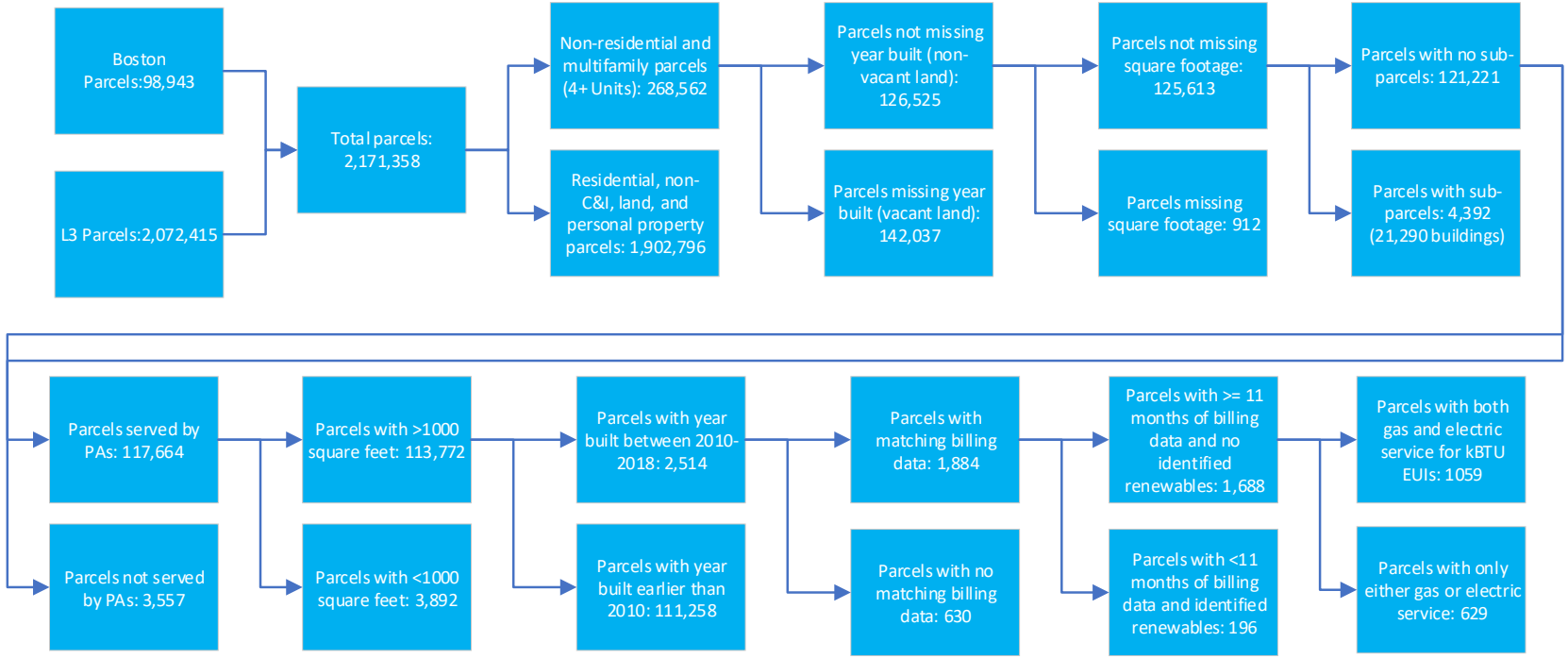


Table 2. 2010-2018 MA L3 non-residential new construction by building category[4]

Building Use	N
College/University	7
Education	35
Food Sales	14
Food Service	72
Health Care (Inpatient)	5
Health Care (Outpatient)	51
Lodging	36
Manufacturing	15
Mercantile (Enclosed and Strip Malls)	26
Mercantile (Retail Other Than Mall)	160
Mixed-Use	17
Multifamily	240
Office	152
Other	12
Public Assembly	41
Public Order and Safety	6
Religious Worship	10
Service	62
Unknown	4
Warehouse and Storage	94
Total	1,059

2.2 EUI calculations

For this analysis, DNV GL estimated site EUIs only, dividing kBTUs based on the total kWh and therm consumption for all accounts associated with a building into that building's square footage.[5] To avoid over- or underestimating EUIs resulting from combining kWh-only and therm-only estimates, we calculated EUIs only for buildings served by both electricity and gas, to produce a major fuel estimate for each building category. For example, from the data, we do not know which buildings have electric heat. If we were to calculate average kWh EUIs for a building category where some buildings used electric heat, and then combined that value with the average building category therm/sqft EUI as a stand-in for the delivered fuel or district steam consumption for which we lack data, we would potentially greatly overstate the average EUI for that category by double-counting heating energy consumption.[6]

[4] Due to a focus on commercial buildings in this analysis, we do not include manufacturing throughout the remainder of this memo.

[5] Site EUIs do not consider the primary energy consumed to produce the secondary energy used by each building; for example, electricity or district steam.

[6] We may still underestimate the energy consumed by some buildings if they utilize both utility gas and a delivered fuel, since we have no information regarding the consumption of other delivered fuels (oil, propane, district steam, etc.). Our simplifying assumption for the EUI calculations is that if a building is consuming both electricity and utility gas, they are not consuming a delivered fuel.

2.2.1 Outliers

One complication we encountered during the initial analysis was the presence of several outliers in many of the building categories, identified due to the extreme variablility in EUIs for many of these categories. Additional research found that some of these outliers were the result of incorrect data (mismatch between tax address and billing address for the same customer, incorrect square footage in the tax data, miscategorized tax land use), while others were the result of our inability to perform EUI calculations at a more granular level than the modified CBECS classes described above. For example, the CBECS class of "outpatient healthcare" may include medical imaging centers, which have a much higher EUI than other outpatient healthcare facilities and appear as an outlier. However, without primary research into each such facility, we cannot identify something like a medical imaging center in order to move it to its own sub-category.

As a first step to reduce outliers, we dropped buildings with very low annual consumption, using the same criteria as the previous and future existing building baseline studies (<2,000 kWh or <70 therms). Further, we added residential consumption to the C&I consumption for multifamily buildings, since we had been capturing only common area energy use for many of these buildings during our initial analysis. To reduce the leverage of the remaining outliers on the EUI estimates, and without a way to tell whether each outlier was the result of data error or overly diverse industry categories, we elected to cap, rather than remove, the outliers using the interquartile (IQR) range rule:

$$\begin{cases} if\ EUI > Q3 + 1.5\ x\ IQR\ then\ EUI = Q3 + 1.5\ x\ IQR \\ if\ EUI < Q1 - 1.5\ x\ IQR\ then\ EUI = Q1 - 1.5\ x\ IQR \end{cases}$$

Identifying and capping the outliers had the effect of substantially reducing variation for most industry sector/size categories. The outcome of this effort is visible in the Findings section in Table 5.

2.2.2 Renewables

Another complication with EUI calculations using the PAs' billing data involves renewables, or on-site generation. Due to the potential presence of renewable generation at each site, especially solar resulting from the solar renewable energy certificate (SREC) program, the solar Massachusetts renewable target (SMART) program, and net metering, net consumption may be driven down for both new and existing buildings. This skews the EUIs and makes it difficult or impossible to assess true building performance with PA consumption data.[7]

DNV GL first received a flag from one PA indicating renewable interconnection agreements or net metering rate codes for some buildings in 2017. As of 2018, two PAs provide this information. [8] For this analysis, we removed sites with likely renewables activity when possible based on the flag provided by the PAs or negative monthly consumption.[9] We cannot reliably identify sites that the PAs have not flagged and whose consumption is never reduced below zero by renewables/on-site generation.

[7] As of 2018, nearly 10% of electricity in MA was produced by solar: https://www.eia.gov/state/analysis.php?sid=MA.

[8] These interconnection flags and rate codes do not necessarily cover the entire billing population., requiring additional analysis to identify potential renewables.

[9] Using negative monthly consumption to identify buildings with renewables is not ideal since some negative bills may be corrections and unrelated to renewables activity. To avoid removing bill corrections, we limited our flagging of negative monthly consumption to bill periods with a single negative read, i.e., if an account had two reads for 1/1/18 – 1/31/18, one negative and one positive, these would not be removed.

3 FINDINGS

This section presents the findings of our revised EUI analysis tasks. After a discussion of EUIs from secondary sources, we present EUI results based on the PAs' billing data and MA tax data.

3.1 Secondary EUI sources

In addition to the EUIs calculated using the PAs' billing data and MA tax data, we explored whether there were secondary data sources available to produce baselines for non-residential new construction. While we found no secondary data sources that capture EUIs for a wide range of recently constructed buildings, we found that data available from the EIA/CBECS for the Northeast census region and New England census division may provide a good check against EUIs calculated with primary data sources.[10] We use the Northeast region alone in the final analysis for comparison due to the limited number of building-level estimates available from the New England census division.

3.1.1 EIA/CBECS

According to the EIA, "The Commercial Buildings Energy Consumption Survey (CBECS) is a national sample survey conducted on a quadrennial basis that collects information on the stock of U.S. commercial buildings, including their energy-related building characteristics and energy usage data (consumption and expenditures). Commercial buildings include all buildings in which at least half of the floor space is used for a purpose that is not residential, industrial, or agricultural. By this definition, CBECS includes building types that might not traditionally be considered commercial, such as schools, hospitals, correctional institutions, and buildings used for religious worship, in addition to traditional commercial buildings such as stores, restaurants, warehouses, and office buildings."[11]

While the most recent CBECS survey was conducted in 2012, contains limited new construction for our analysis time frame (2010-2018), and consists primarily of existing building stock, we believe CBECS is the best source for confirming that our tax-data-calculated EUIs are not unreasonable. Before using the CBECS data for this purpose, we wanted to confirm that industry sector EUIs have not changed substantially over time. To do this, we compared the 2012 CBECS results to the 2003 CBECS results. We found that there were relatively small changes (0-10%) in average major fuel EUIs between the two periods for more than half of the building categories. Vacant buildings were the only category with a change in EUI greater than 30%. Because of this, and because our review of time-series EUIs in section 2.1.5 did not identify a particular trend in EUIs for new construction from 2010-2018, we believe that these values provide a way to confirm the EUIs calculated with primary data are in a reasonable range for each building category. Note that while EUIs for most building use categories decreased from 2003-2012, food sales, food service, non-mall retail, and outpatient health care all saw EUI increases during that time period.

Table 3 presents major fuel kBTU per sqft EUIs for the 2012 and 2003 CBECS. CBECS provides only median and 25th percentile EUIs in addition to averages at the national level, and provides regional and division EUIs only as averages.

[10] The Northeast census region includes Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont, New Jersey, New York, and Pennsylvania. The New Engalnd census division drops the Middle-Atlantic states from this list and onily includes Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont.

[11] https://www.eia.gov/consumption/commercial/about.php

Table 3. CBECS major fuel kBTU EUIs – 2012 and 2003 results[12]

CBECS Building Use	2012						2003			2012 vs 2003 Difference in Average
	Average	25th Pctl	Median	75th Pctl	Average	Average	Average	Average	Average	
	All Buildings				Northeast	New England	All Buildings	Northeast	New England	
Education	68.8	31.1	50.8	81.1	82.1	69.7	83.1	101.6	Q	-17.2%
Food Sales	209.5	128.1	222.8	321.6	227.8	Q	199.7	Q	Q	4.9%
Food Service	282.7	94.4	241.1	449.0	249.0	Q	258.3	Q	Q	9.4%
Health care overall	172.7	42.4	63.2	90.9	165.0	176.5	187.7	212.2	Q	-8.0%
Health Care (Inpatient)	231.1	163.8	205.7	272.6	230.4	237.1	249.2	Q	Q	-7.3%
Health Care (Outpatient)	94.8	33.5	61.7	82.1	98.5	Q	94.6	Q	Q	0.2%
Lodging	96.9	48.1	69.1	119.1	97.8	Q	100.0	Q	Q	-3.1%
Mercantile overall	88.9	33.4	63.6	107.6	92.8	75.2	91.3	96.5	Q	-2.6%
Mercantile (Enclosed and Strip Malls)	66.9	25.6	52.2	88.3	66.9	Q	73.9	65.0	Q	-9.5%
Mercantile (Retail Other Than Mall)	109.3	59.0	100.8	177.3	119.2	Q	102.2	Q	Q	6.9%
Office	77.8	32.6	52.7	80.2	99.8	82.4	92.9	101.2	114.6	-16.3%
Public assembly	86.3	22.2	49.3	82.5	97.1	69.9	93.9	Q	Q	-8.1%
Public order and safety	92.2	32.9	62.1	107.7	108.4	Q	115.8	Q	Q	-20.4%
Religious worship	38.0	12.5	25.1	49.8	41.5	Q	43.5	52.1	Q	-12.6%
Service	58.7	18.5	42.2	83.7	59.3	54.9	77.0	79.8	Q	-23.8%
Warehouse and storage	32.8	2.5	13.8	35.6	35.8	Q	45.2	41.6	Q	-27.4%
Other	142.9	16.5	44.1	126.5	236.4	Q	164.4	Q	Q	-13.1%
Vacant	12.7	0.0	0.0	6.4	Q	Q	20.9	Q	Q	-39.2%

[12] Q: Data withheld either because the relative standard error (RSE) was greater than 50% or because fewer than 20 buildings were sampled.

3.2 Primary data source EUIs

The following tables present primary data source EUIs for buildings constructed between 2010 and 2018, calculated based on square footage information from the MA L3 and Boston tax data and from consumption information in the MA Data Warehouse (both C&I data and residential data for the multifamily building use category), broken out by building use and size.

During discussions about this analysis, the PAs requested that DNV GL investigate EUIs for new buildings that had completed CHP projects. We only identified a single CHP project associated with a building in our 2010-2018 new construction population. Since there was only a single instance, we do not provide a breakout for buildings with CHP versus buildings without CHP.

3.2.1 Tax data EUIs (buildings constructed in 2010-2018)

Table 5 shows the overall EUIs for all buildings constructed in 2010-2018, based on 2018 consumption data. We provide boxplots of these results in Appendix A – EUI Boxplots. Table 4 shows a subset of the "best" EUIs (those based on 5 or more buildings, with the lowest coefficients of variation (CVs) and with estimates closest to CBECS) sorted from least variable to most variable. Many of the building-use size categories highlighted here are <10,000 sqft, which are not the target of the PAs' new construction initiative.

The size categories in Table 4 and Table 5 are as follows:

1: small (<10,000 sqft)

2: medium (10,000 - 50,000 sqft)

3: large (>50,000 sqft)

Based on these results, among large buildings (> 50,000 sqft), it appears that food sales, outpatient health care, education, office, multifamily, lodging, and warehouse and storage may be good candidates for baseline development. Note that the presence of certain energy-intensive facilities within new education (swimming pools) and healthcare (diagnostic imaging) buildings may mean that these average EUIs would underestimate these facilities' energy use.

Table 4. A subset of EUIs sorted by CV and difference from 2012 Northeast CBECS

Building use	Size category	Capped outliers average	CBECS Northeast average	MA vs CBECS % difference	N	Capped outliers CV
Public Order and Safety	2	90.68	108.40	-16%	5	16%
Religious Worship	2	41.06	41.50	-1%	5	26%
Food Sales	3	195.79	227.80	-14%	10	29%
Health Care (Outpatient)	3	123.58	98.50	25%	6	33%
Lodging	3	77.59	97.80	-21%	23	35%
Multifamily	2	45.97	N/A	N/A	87	39%
Multifamily	3	39.94	N/A	N/A	68	41%
Multifamily	1	51.24	N/A	N/A	85	41%
Public Assembly	2	79.41	97.10	-18%	21	43%
Education	3	68.28	82.10	-17%	25	43%
Health Care (Outpatient)	2	88.73	98.50	-10%	18	50%
Lodging	2	85.62	97.80	-12%	9	51%
Education	2	74.69	82.10	-9%	7	52%
Public Assembly	3	123.93	97.10	28%	8	53%
Mercantile (Retail Other Than Mall)	3	83.43	119.20	-30%	12	54%
Health Care (Outpatient)	1	63.76	98.50	-35%	27	58%
Office	1	98.70	99.80	-1%	88	63%
Office	2	86.81	99.80	-13%	38	64%
Office	3	78.87	99.80	-21%	26	65%
Mercantile (Retail Other Than Mall)	2	117.02	119.20	-2%	63	71%
Public Assembly	1	74.45	97.10	-23%	12	72%
Mercantile (Retail Other Than Mall)	1	148.78	119.20	25%	85	81%
Warehouse and Storage	3	37.61	35.80	5%	11	102%

Table 5. Overall kBTU EUIs for MA buildings constructed in 2010-2018, using PA data (part 1)[13]

Building use	Size category	Median		Mean				N		CV	
		All records	Capped outliers	All records	Capped outliers	CBECS Northeast average	MA vs CBECS % difference	All records	Outliers	All records	Capped outliers
College/ University	2	91.83	91.83	791.39	177.72	N/A	N/A	3	1	154%	88%
College/ University	3	155.01	155.01	148.91	148.91	N/A	N/A	4	1	30%	30%
Education	1	42.80	42.80	42.69	42.69	82.10	-48%	3	1	21%	21%
Education	2	66.52	66.52	81.59	74.69	82.10	-9%	7	2	61%	52%
Education	3	65.01	65.01	93.91	68.28	82.10	-17%	25	4	121%	43%
Food Sales	1					227.80		0	0	0%	0%
Food Sales	2	117.89	117.89	160.43	160.43	227.80	-30%	4	0	84%	84%
Food Sales	3	203.14	203.14	195.79	195.79	227.80	-14%	10	0	29%	29%
Food Service	1	565.26	565.26	554.89	523.24	249.00	110%	68	4	69%	46%
Food Service	2	94.62	94.62	178.14	178.14	249.00	-28%	4	0	125%	125%
Food Service	3					249.00		0	0	0%	0%
Health Care (Inpatient)	1	822.04	822.04	822.04	822.04	230.40	257%	1	1	0%	0%
Health Care (Inpatient)	2	3822.78	1896.56	3822.78	1896.56	230.40	723%	1	1	0%	0%
Health Care (Inpatient)	3	105.70	105.70	141.34	141.34	230.40	-39%	3	0	50%	50%
Health Care (Outpatient)	1	56.79	56.79	66.86	63.76	98.50	-35%	27	1	74%	58%
Health Care (Outpatient)	2	77.54	77.54	96.70	88.73	98.50	-10%	18	4	72%	50%
Health Care (Outpatient)	3	112.69	112.69	152.04	123.58	98.50	25%	6	1	70%	33%
Lodging	1	83.22	83.22	97.21	84.85	97.80	-13%	4	3	67%	51%
Lodging	2	69.36	69.36	99.40	85.62	97.80	-12%	9	3	67%	51%

[13] Note: small Ns and the removal of campus sites limit our ability to generate realistic EUIs for inpatient healthcare.

Building use	Size category	Median		Mean				N		CV	
		All records	Capped outliers	All records	Capped outliers	CBECS Northeast average	MA vs CBECS % difference	All records	Outliers	All records	Capped outliers
Lodging	3	72.90	72.90	82.92	77.59	97.80	-21%	23	3	55%	35%
Mercantile (Enclosed and Strip Malls)	1	312.02	312.02	238.02	238.02	66.90	256%	5	1	63%	63%
Mercantile (Enclosed and Strip Malls)	2	124.06	124.06	124.93	124.93	66.90	87%	10	0	56%	56%
Mercantile (Enclosed and Strip Malls)	3	310.75	310.75	297.31	295.40	66.90	342%	11	2	50%	49%
Mercantile (Retail Other Than Mall)	1	91.30	91.30	164.21	148.78	119.20	25%	85	30	107%	81%
Mercantile (Retail Other Than Mall)	2	88.36	88.36	148.14	117.02	119.20	-2%	63	9	180%	71%
Mercantile (Retail Other Than Mall)	3	86.62	86.62	83.43	83.43	119.20	-30%	12	1	54%	54%
Mixed Use	1	44.18	44.18	378.23	149.23	N/A	N/A	7	2	155%	130%
Mixed Use	2	64.86	64.86	119.13	119.13	N/A	N/A	10	2	111%	111%
Mixed Use	3					N/A	N/A	0	0	0%	0%
Multifamily	1	46.31	46.31	238.52	51.24	N/A	N/A	85	20	258%	41%
Multifamily	2	43.82	43.82	51.37	45.97	N/A	N/A	87	10	83%	39%
Multifamily	3	37.73	37.73	42.09	39.94	N/A	N/A	68	4	58%	41%
Office	1	84.44	84.44	324.72	98.70	99.80	-1%	88	19	477%	63%
Office	2	71.51	71.51	114.64	86.81	99.80	-13%	38	4	140%	64%

Building use	Size category	Median		Mean				N		CV	
		All records	Capped outliers	All records	Capped outliers	CBECS Northeast average	MA vs CBECS % difference	All records	Outliers	All records	Capped outliers
Office	3	64.78	64.78	78.87	78.87	99.80	-21%	26	3	65%	65%
Other	1	66.49	66.49	75.86	75.86	236.40	-68%	5	2	106%	106%
Other	2	110.97	110.97	138.43	138.43	236.40	-41%	3	0	43%	43%
Other	3	51.39	51.39	65.37	65.37	236.40	-72%	4	0	83%	83%
Public Assembly	1	55.47	55.47	81.63	74.45	97.10	-23%	12	5	89%	72%
Public Assembly	2	67.54	67.54	79.71	79.41	97.10	-18%	21	3	44%	43%
Public Assembly	3	145.76	145.76	202.96	123.93	97.10	28%	8	3	119%	53%
Public Order and Safety	1	134.13	125.09	134.13	125.09	108.40	15%	1	2	0%	0%
Public Order and Safety	2	92.79	92.79	84.38	90.68	108.40	-16%	5	1	33%	16%
Public Order and Safety	3					108.40		0	0	0%	0%
Religious Worship	1	47.45	47.45	84.47	59.82	41.50	44%	5	2	71%	44%
Religious Worship	2	42.41	42.41	41.06	41.06	41.50	-1%	5	0	26%	26%
Religious Worship	3					41.50		0	0	0%	0%
Service	1	260.85	260.85	226.73	226.73	59.30	282%	59	7	61%	61%
Service	2	66.65	66.65	61.47	61.47	59.30	4%	3	0	73%	73%
Unknown	2	101.99	101.99	101.99	101.99	N/A	N/A	1	0	0%	0%
Unknown	3	90.85	90.85	108.27	108.27	N/A	N/A	3	1	49%	49%
Warehouse and Storage	1	50.07	50.07	69.89	55.68	35.80	56%	51	13	107%	49%
Warehouse and Storage	2	48.28	48.28	184.34	53.95	35.80	51%	32	11	385%	60%
Warehouse and Storage	3	21.12	21.12	40.83	37.61	35.80	5%	11	3	112%	102%

3.2.2 Clustering

In order to further refine the EUI estimates, and under the assumption that many of the identified outlier EUIs resulted from a lack of granularity in the building use categories, we attempted to break building use categories into smaller groups by applying a k-means clustering algorithm to the EUI calculation. We show the results of this effort in Appendix B – EUI Clusters, with a subset of the data below in Table 6. While the clustering of EUIs highlighted the existence of distinct groups within each building use type, we lack information to assign descriptions to these groups without additional primary research. We conducted some additional research on two of the categories to highlight EUI differences that are potentially based on activity and hours of use.

Education breaks into two distinct groups, one containing a mix of schools serving K-12 students (elementary, middle, and high schools) and one containing only elementary and pre-K schools/daycare, with three additional outlier clusters. Two of these outlier clusters appear to be reasonable due to the presence of swimming and outdoor athletic complexes that add to the building's energy consumption without increasing square footage. The final cluster (3) appears to be a true outlier, as the building is simply a combined middle and high school, but has an EUI nearly 10 times greater than other similar buildings.

Within the outpatient healthcare sector, differences in EUIs appear to be based on building size and hours of use. For example, specialist outpatient centers with limited hours who see patients on an appointment basis appear to have lower EUIs than extended-hour urgent care centers. Note that two buildings that would appear to be outliers are actually medical imaging centers, which we would expect to have considerably higher EUIs.

Table 6. Example of building use broken out by EUI cluster with auxiliary information

Building use	Cluster	Median sqft	Median EUI	N	CV	Cluster notes
Education	1	95,842	65.77	22	17%	K-12 Schools
	2	51,524	268.00	1	0%	High School and Community Pool
	3	235,150	584.25	1	0%	Combined middle and high school
	4	104,987	34.06	8	41%	Elementary schools and pre-k/day care
	5	42,420	151.60	3	7%	Athletics complex, center for science and innovation
Health Care (Outpatient)	1	12,205	80.00	14	15%	Medical groups, vetrinary, eye care
	2	38,256	348.25	2	6%	Imaging centers
	3	5,388	274.73	1	0%	Parcel also contains hot yoga studio. No sub-parcel information in tax data
	4	14,588	127.12	9	11%	Urgent care, rehab and wound care, oral surgery, family practice
	5	5,440	46.20	25	30%	Specialists: ENT, OB/GYN, Urologist, Ophthalmology. Also a few medical marijuana dispensaries

4 CONCLUSIONS AND CONSIDERATIONS

4.1 Conclusions

With revisions to this analysis, most notably increasing the analysis population from 2017-2018 to 2010-2018, and the handling of outlier EUIs, DNV GL was able to generate less variable EUI estimates for most building-use size categories. These results are also reasonably close to (and usually somewhat lower than) the average EUI estimates seen in the Northeast census region CBECS data from 2012.

Despite the changes, the analysis still contains very small (or no) populations of buildings that are most likely to participate in the PAs' initiative for large new construction buildings: hospitals, colleges, and other campuses. Unfortunately, it is unlikely that we would be able to develop defensible EUIs for those categories given the caveats discussed throughout this memo. Finally, while clustering EUIs within the building use categories substantially reduces variation and highlights differences within categories potentially based on operating hours, occupancy, and specific business activity, unfortunately, this level of detailed information is not available for all buildings without primary research.

4.2 Considerations

Based on the revised results, DNV GL suggests the PAs consider developing pilot baselines informed by these results for some building-use size categories with less variable EUIs, using the values shown in Table 4. Good candidates would include large (>50,000 sq ft) education, large outpatient healthcare, large office, large multifamily, large food sales, large lodging, and large warehouse and storage. We also suggest that the PAs consider additional research to gather detailed information on specific buildings to allow tailoring of EUIs to more granular building use subsegments (e.g. medical imaging and primary care physician versus outpatient healthcare).

5 APPENDIX A – EUI BOXPLOTS

The figures on the following pages show paneled boxplots of EUIs by industry sector and building size. The outer edges of each box represent the 25th and 75th percentile EUIs, while the center line in each box represents the median EUI. The diamond represents the average EUI. The ends of each whisker represent the minimum and maximum observations within 1.5 times the interquartile range. Outliers appear as points outside of the whiskers. We have limited the y-axis of the boxplot panels to 500 kBTU/sqft for readability, so any extreme outliers beyond this range will not be visible. Each box also contains scattered points that show the actual EUIs for each building in our dataset, providing an indication of the distribution of EUIs that make up each box. The individual points in each panel are colored to correspond to an EUI cluster assigned using a k-means algorithm.

Figure 5. Dual fuel kBTU/sqft EUIs by building use (part 1)

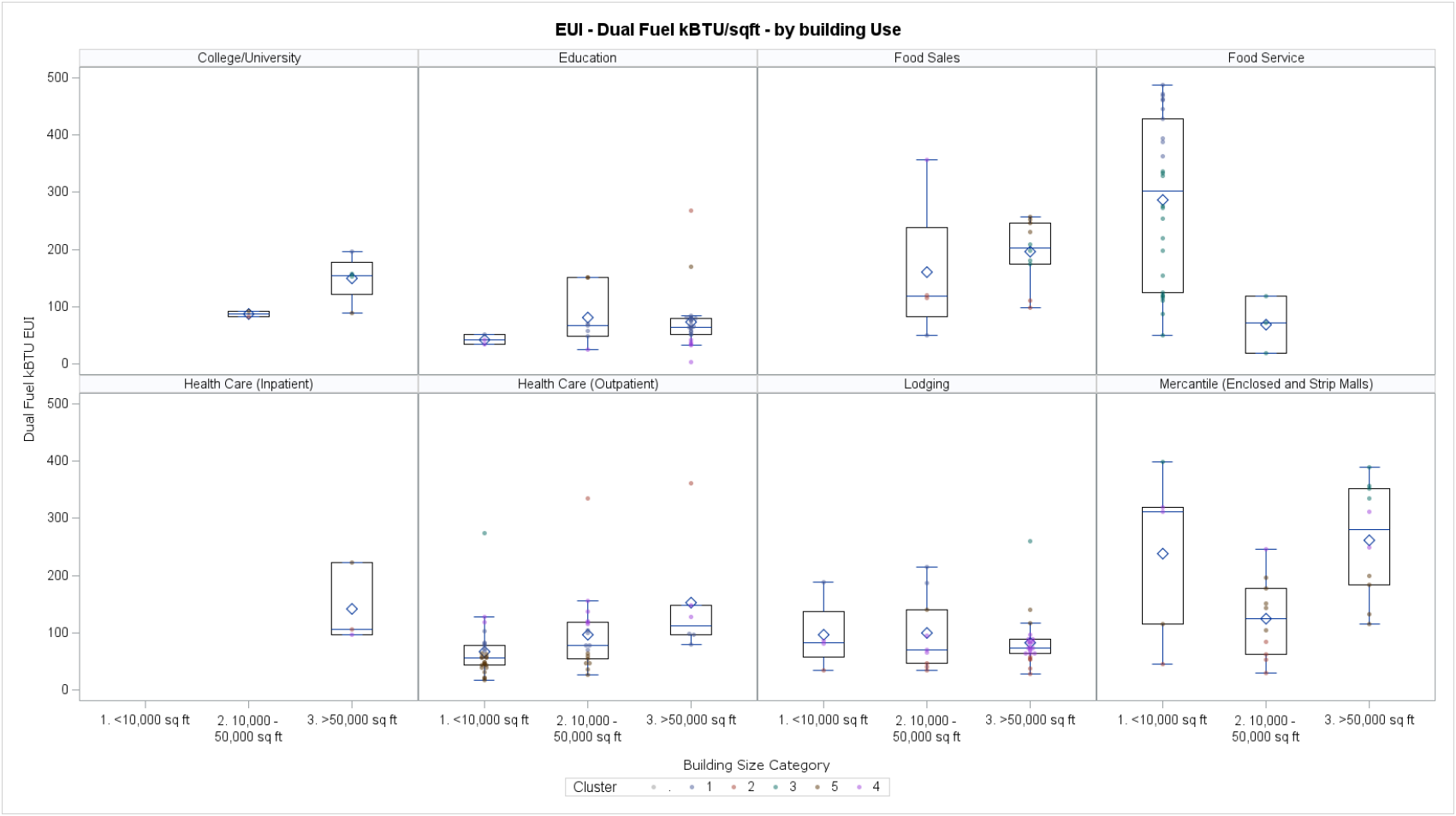


Figure 6. Dual fuel kBTU/sqft EUIs by building use (part 2)

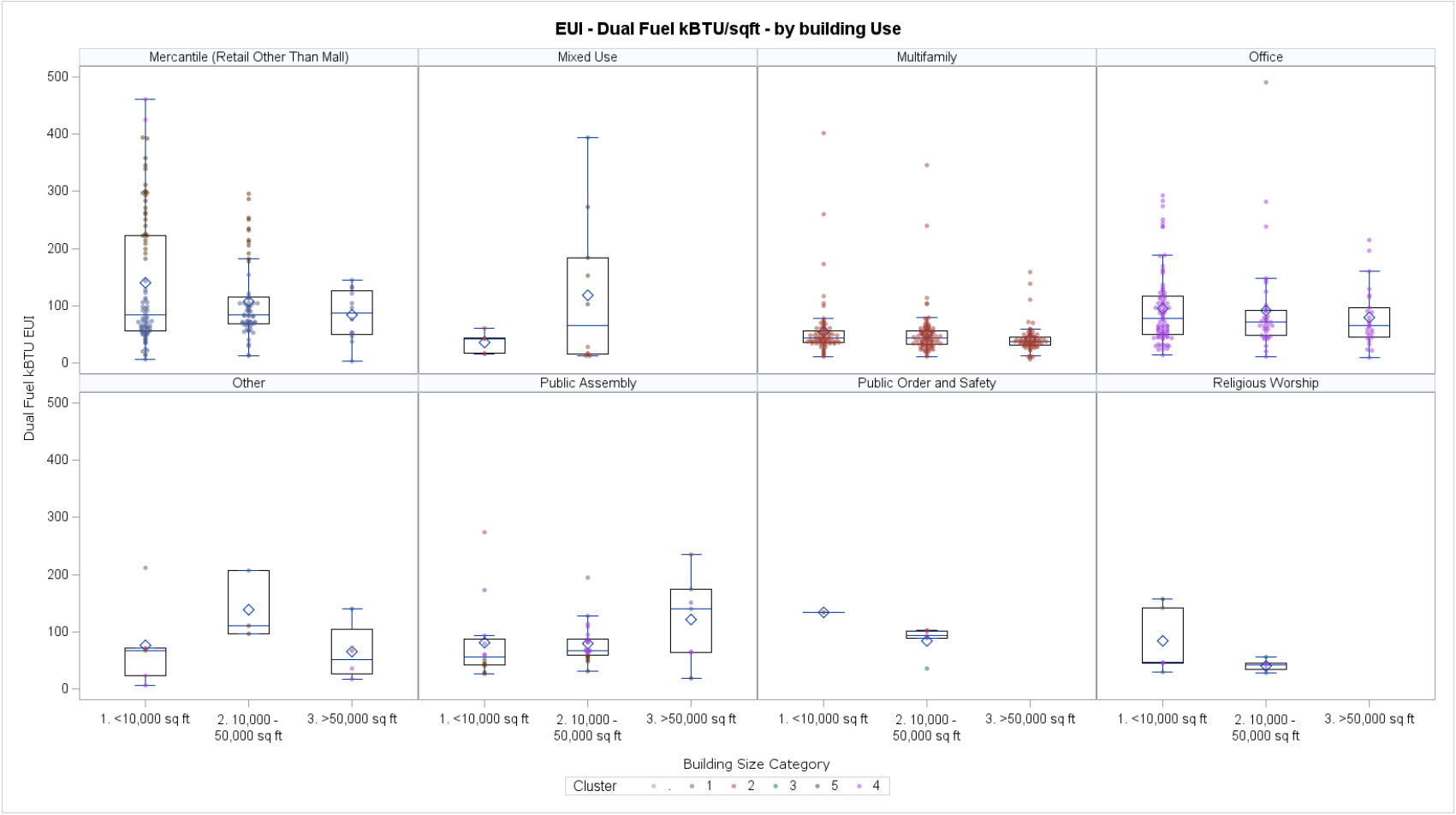
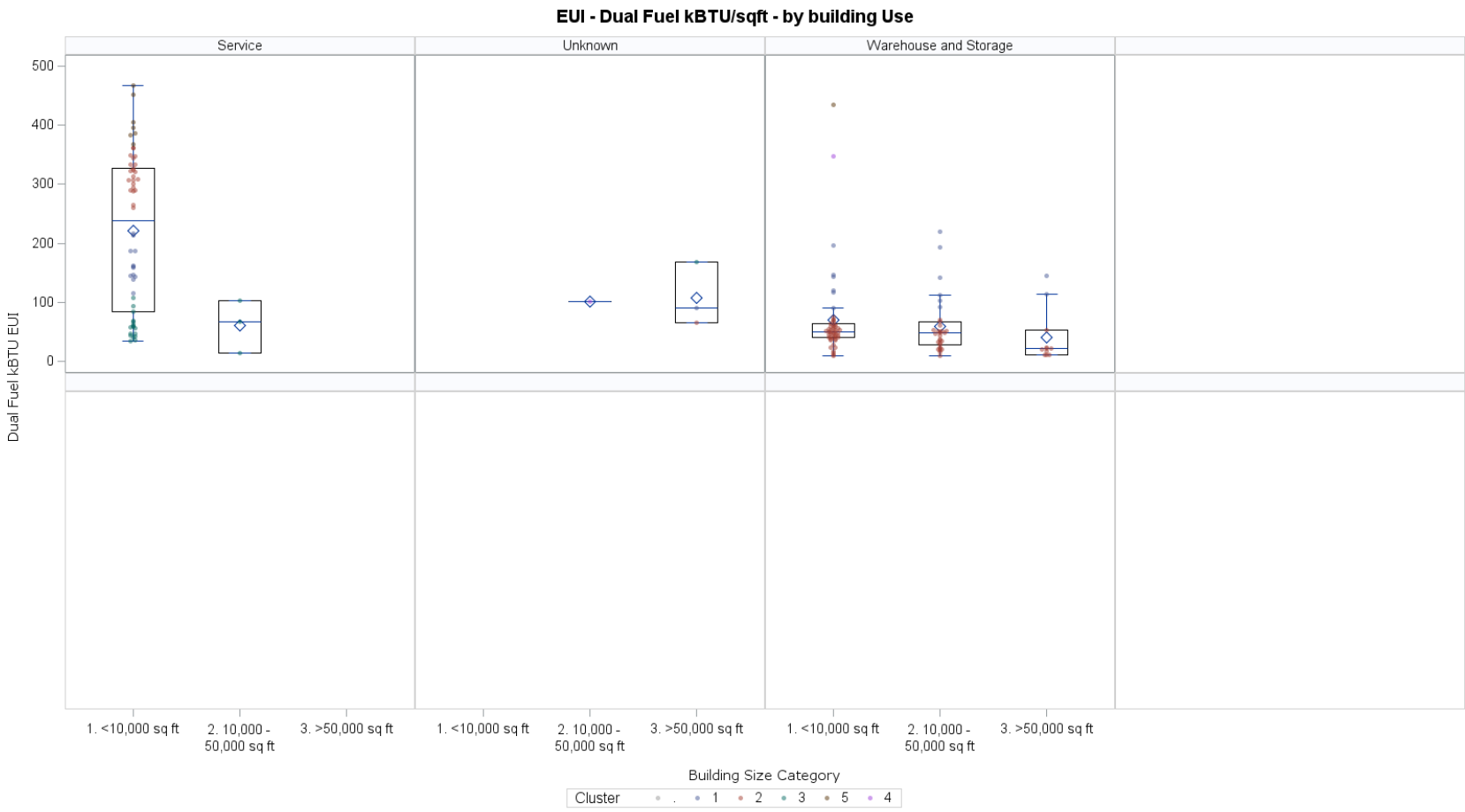


Figure 7. Dual fuel kBTU/sqft EUIs by building use (part 3)



6 APPENDIX B – EUI CLUSTERS

Table 7. Dual fuel kBTU/sqft EUIs by building use and cluster (part 1)

Building Use	Cluster	Median Sq Ft	Mean Sq Ft	Median EUI	Mean EUI	N	CV
College/University	1	90,861	90,861	196.81	196.81	1	0%
	2	34,841	34,841	82.54	82.54	1	0%
	3	102,494	102,494	155.01	155.01	2	2%
	4	13,135	13,135	2199.81	2199.81	1	0%
	5	64,278	64,278	90.33	90.33	2	2%
Education	1	95,842	149,425	65.77	66.91	22	17%
	2	51,524	51,524	268.00	268.00	1	0%
	3	235,150	235,150	584.25	584.25	1	0%
	4	104,987	85,082	34.06	31.21	8	41%
	5	42,420	41,251	151.60	157.64	3	7%
Food Sales	1	17,746	17,746	49.48	49.48	1	0%
	2	83,341	142,208	113.58	111.37	4	8%
	3	65,563	64,048	188.68	190.23	4	8%
	4	16,835	16,835	356.47	356.47	1	0%
	5	79,903	76,128	249.77	246.81	4	5%
Food Service	1	3,150	3,677	554.85	529.35	33	13%
	2	2,280	4,042	1217.26	1146.96	3	12%
	3	4,194	8,181	124.32	173.92	19	58%
	4	3,532	3,987	681.48	704.34	16	11%
	5	1,824	1,824	2962.08	2962.08	1	0%
Health Care (Inpatient)	1	26,696	26,696	3822.78	3822.78	1	0%
	2	77,515	77,515	105.70	105.70	1	0%
	3	2,170	2,170	822.04	822.04	1	0%
	4	80,703	80,703	96.26	96.26	1	0%
	5	61,001	61,001	222.08	222.08	1	0%

Table 8. Dual fuel kBTU/sqft EUIs by building use and cluster (part 2)

Building Use	Cluster	Median Sq Ft	Mean Sq Ft	Median EUI	Mean EUI	N	CV
Health Care (Outpatient)	1	12,205	21,639	80.00	85.01	14	15%
	2	38,256	38,256	348.25	348.25	2	6%
	3	5,388	5,388	274.73	274.73	1	0%
	4	14,588	27,372	127.12	129.51	9	11%
	5	5,440	8,271	46.20	45.25	25	30%
Lodging	1	11,312	14,736	188.63	196.91	3	8%
	2	68,782	83,378	40.20	43.02	9	24%
	3	66,583	66,583	260.42	260.42	1	0%
	4	71,091	68,984	77.71	77.76	20	14%
	5	57,068	69,028	139.48	132.37	3	10%
Mercantile (Enclosed and Strip Malls)	1	67,344	67,344	646.13	646.13	1	0%
	2	16,748	18,330	52.73	54.86	5	37%
	3	71,268	80,442	357.21	366.53	5	7%
	4	44,300	42,359	310.75	287.37	5	13%
	5	33,650	108,289	147.40	151.99	10	23%
Mercantile (Retail Other Than Mall)	1	11,813	22,566	71.00	72.98	115	43%
	2	18,451	18,451	2093.32	2093.32	1	0%
	3	7,322	7,322	1165.97	1165.97	1	0%
	4	5,648	6,924	508.39	562.30	5	25%
	5	5,416	11,211	251.31	258.63	38	22%
Mixed Use	1	2,799	2,799	1290.48	1290.48	1	0%
	2	10,061	12,930	17.24	26.43	10	64%
	3	26,446	26,446	394.78	394.78	1	0%
	4	4,704	4,704	1177.77	1177.77	1	0%
	5	12,916	14,929	168.08	177.90	4	40%

Table 9. Dual fuel kBTU/sqft EUIs by building use and cluster (part 3)

Building Use	Cluster	Median Sq Ft	Mean Sq Ft	Median EUI	Mean EUI	N	CV
Multifamily	1	3,912	3,912	1656.95	1656.95	2	7%
	2	21,713	58,542	41.13	49.93	231	85%
	3	8,604	8,604	3903.10	3903.10	1	0%
	4	6,191	6,191	2885.61	2885.61	1	0%
	5	6,527	5,513	1250.47	1193.67	5	11%
Office	1	12,080	14,957	546.50	655.97	3	37%
	2	2,122	2,122	2156.51	2156.51	2	10%
	3	1,764	1,764	1548.42	1548.42	1	0%
	4	7,200	44,371	71.30	88.56	145	68%
	5	3,220	3,220	14311.90	14311.90	1	0%
Other	1	18,070	18,070	209.25	209.25	2	1%
	2	6,100	29,251	67.38	68.39	3	4%
	3	350,892	350,892	140.94	140.94	1	0%
	4	111,968	122,209	20.49	20.85	4	57%
	5	17,150	17,150	104.01	104.01	2	9%
Public Assembly	1	63,208	64,530	162.50	160.18	6	16%
	2	31,485	31,485	255.20	255.20	2	11%
	3	65,252	65,252	774.17	774.17	1	0%
	4	15,852	35,216	67.91	76.55	20	22%
	5	13,964	28,063	45.08	41.72	12	30%
Public Order and Safety	1	23,472	23,472	88.84	88.84	1	0%
	2	11,891	11,891	102.34	102.34	2	1%
	3	17,550	17,550	35.61	35.61	1	0%
	4	20,775	20,775	92.79	92.79	1	0%
	5	6,790	6,790	134.13	134.13	1	0%

Table 10. Dual fuel kBTU/sqft EUIs by building use and cluster (part 4)

Building Use	Cluster	Median Sq Ft	Mean Sq Ft	Median EUI	Mean EUI	N	CV
Religious Worship	1	34,560	27,247	30.34	30.81	3	12%
	2	4,320	4,320	142.46	142.46	1	0%
	3	24,376	24,376	55.96	55.96	1	0%
	4	8,342	8,782	45.03	44.98	4	5%
	5	8,907	8,907	156.89	156.89	1	0%
Service	1	4,973	5,199	160.05	164.61	12	18%
	2	3,804	3,696	317.56	315.99	22	9%
	3	4,686	5,653	57.99	59.53	20	40%
	4	3,903	3,903	586.55	586.55	1	0%
	5	4,286	4,142	395.49	408.18	7	9%
Unknown	1	138,964	138,964	90.85	90.85	1	0%
	2	95,548	95,548	65.57	65.57	1	0%
	3	98,000	98,000	168.39	168.39	1	0%
	4	47,362	47,362	101.99	101.99	1	0%
Warehouse and Storage	1	10,548	26,324	130.78	138.19	14	29%
	2	9,120	25,491	43.84	40.72	77	43%
	3	26,896	26,896	4060.30	4060.30	1	0%
	4	4,608	4,608	347.68	347.68	1	0%
	5	2,844	2,844	434.02	434.02	1	0%