Earnings Adjustment Mechanisms to Support New York REV Goals

Outcome-Based, Program-Based, and Action-Based Options

Prepared for Advanced Energy Economy Institute

March 5, 2018

AUTHORS

Tim Woolf
Jamie Hall
Melissa Whited

Synapse Energy Economics, Inc.

485 Massachusetts Avenue, Suite 2
Cambridge, Massachusetts 02139

617.661.3248 | www.synapse-energy.com

CONTENTS

EXECUTIVE SUMMARY

The utilities of New York and the New York Public Service Commission are in the process of developing performance incentive mechanisms, referred to as Earning Adjustment Mechanisms (EAMs). These mechanisms generally fall under one of three broad categories:

- Outcome-based EAMs, which are based on the premise that regulators should define a desired outcome but should not specify explicit programs or actions that the utility must pursue in order to achieve that outcome.
- Program-based EAMs, which are based on the premise that the utility should achieve a desired outcome through a particular program or initiative, especially those that have been reviewed and approved by regulators.
- Action-based EAMs, a term that has not been used widely in New York EAM proceedings, but provides a useful distinction. These EAMs are focused on more specific utility actions but do not measure the impacts of those actions or the extent to which those actions lead to desired outcomes.

In this report, we examine some of the strengths and weaknesses of each of these types of EAMs and offer some general observations and recommendations. Because current efforts to implement EAMs in New York focus on outcome-based metrics, we have concentrated our analysis on outcome-based EAMs, whereas we address program-based and action-based EAMs in a qualitative manner.

Advantages

Each type of EAM offers advantages:

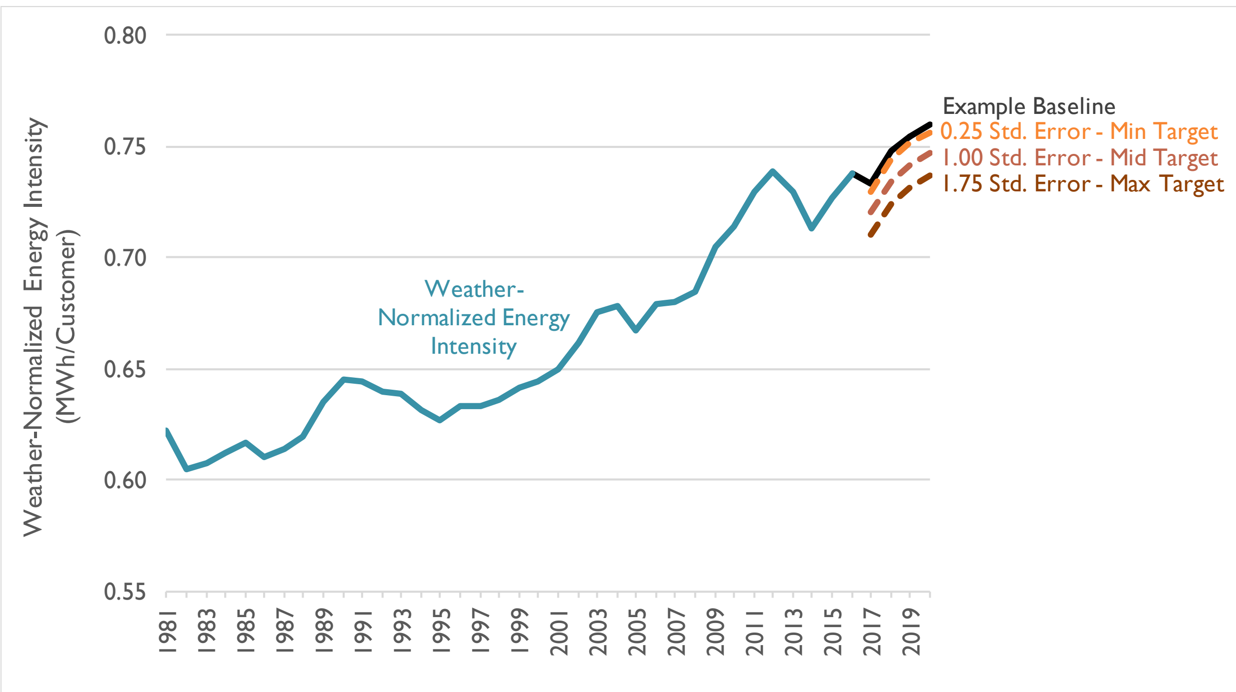
- Outcome-based EAMs provide the utility with flexibility to innovate and focus on the key high-level outcomes desired by regulators.
- Program-based EAMs provide clear and specific regulatory guidance and financial incentives regarding initiatives specifically designed to be cost-effective and in customers' interest.
- Action-based EAMs require the utility to undertake specific actions to assist in the transition to a new utility structure consistent with the goals of the New York Reforming the Energy Vision (REV) process.

Challenges

Each type of EAM also has its challenges. Outcome-based EAMs require careful development of counterfactual baselines and appropriate targets. As an example of some of these challenges, consider the proposed residential energy intensity EAM from New York State Electric and Gas Corporation (NYSEG). This EAM is an outcome-based EAM, as its compensation levels are set based on achieving a specific outcome (reductions in energy intensity), rather than based on the results of a particular utility program or action. Figure ES-1 below shows NYSEG's weather-normalized residential energy intensity since 1981. It includes an example counterfactual baseline for future years (2017-2019), derived using an econometric regression analysis, with potential targets for energy intensity reductions that are calculated using standard errors from the regression. This example reflects the methodology that NYSEG has proposed for setting its energy intensity EAM baseline and targets. The development of the baseline and targets is discussed more in Section 3.4.

This figure illustrates several challenges associated with outcome-based EAMs. First, customer energy intensity can vary significantly from year-to-year, even after accounting for weather variation, making it difficult to determine a reliable counterfactual baseline. The value from a recent year of actual data (2014, when there was no EAM) is lower than the baseline, the minimum, and mid-level targets for 2017, which calls into question whether the future energy intensity results will similarly deviate from the baseline without any effort on the part of the utility.

Figure ES-1. Annual Energy Intensity with Forecasts and Targets



Second, establishing targets for outcome-based EAMs can also be difficult. In this example, the targets were based on the standard error of the regression. The standard error is a measure of the accuracy of the model, based on the difference between the model's estimated values and the actual values. For example, 1.0 standard error is associated with an 83 percent level of confidence, while a standard error of 0.25 is associated with a 60 percent level of confidence.[1] The larger the standard error, the less

[1] These values apply to a one-tailed t-test. Staff Earning Adjustment Mechanisms Panel, "Prepared Testimony of Staff Earning Adjustment Mechanisms Panel," August 2017, page 48.

explanatory power the model has, and the more factors beyond those captured in the model explain the deviation between the actual observations and the model estimates.

However, setting an EAM target based on the standard error of the regression is somewhat arbitrary for several reasons:

- By itself, the standard error contains no information regarding the costs and benefits of achieving a certain target. Thus, without an analysis of the costs and benefits, setting a target based on the standard error does not guarantee that a certain level of performance will result in net benefits for customers.
- There may be considerable uncertainty as to whether the utility was responsible for the deviation from the baseline. This is especially true with confidence levels of 60 percent (0.25 standard error), but even at higher confidence levels, it is always possible that a deviation between the model's estimated value and the actual value was caused by a factor not explained in the model and not caused by the utility.

Program-based EAMs also face important challenges. In general, they require more regulatory oversight to define the programs and initiatives and leave less room for utility innovation. They also pose challenges regarding the measurement and verification of performance, particularly programs with less history and experience, such as electric vehicle or distributed storage initiatives.

Action-based EAMs provide the utility with even less flexibility to innovate, and they might not lead to direct benefits to customers. Further, it can be difficult for regulators to determine just which actions are most likely to lead to desired outcomes.

Conclusions and Recommendations

The table below provides a summary of our primary conclusions regarding the advantages, challenges, and best applications of each type of EAM.

TYPE OF EAM	ADVANTAGES	CHALLENGES	BEST APPLICATIONS
Outcome-Based EAMs	• Directly tied to policy goals • Flexible and promote utility innovation • Metrics are typically easily measured • Do not require extensive program planning processes or oversight	• Establishing counterfactual baselines can be contentious • Determining appropriate targets can be challenging • It may be difficult to determine whether the outcome was the result of utility action or due to other factors • Costs associated with utility efforts may not be readily known, hindering ability to ensure EAM is cost-effective	• Well-suited for measuring specific, high-level outcomes such as peak demand or customer load factors.
Program-Based EAMs	• Provide clear and specific regulatory guidance • Programs must pass cost-effectiveness tests, so there is greater certainty regarding net benefits • Well-established programs provide a wealth of data useful for determining EAM targets and measuring performance	• Requires substantial resources to develop a program • Setting baselines and targets may be difficult, particularly where there is little existing experience to draw from • Most measurement and verification practices have been developed for energy efficiency; other utility programs may require development of additional methodologies	• Most appropriate where programmatic processes, regulatory oversight, and verification protocols already exist or are worthwhile for their own sake.
Action-Based EAMs	• Relatively simple to develop and administer • Performance easily measured • Costs can be estimated and approved in advance	• Little indication of whether the action helps achieve broader policy goals • Cost-effectiveness uncertain where benefits are not quantifiable	• Best when the ultimate policy goal may be beyond the utility's control, but where utility action is deemed important for achieving the policy objectives

While developing EAMs, commissions should carefully weigh the advantages, disadvantages, and applicability of each type of EAM and avoid too much emphasis on any one type. There should be a balance in the types of EAMs used and, importantly, the EAMs should be chosen based on the ability of the EAM to achieve the goals the commission wishes to achieve. In general:

- Outcome-based EAMs should be used when a specific outcome is desired and it is in customers' interest to allow the utility to pursue the outcome with innovative approaches that are not determined or overseen by the regulators. Outcome-based EAMs should recognize the challenges associated with verifying the extent of a utility's influence on the desired outcome. They should be designed to provide appropriate levels of incentives that are tied to desired outcomes where confidence is high that utility actions will produce those outcomes.

- Program-based EAMs should be used to support utility initiatives that include pre-determined program designs, regulatory oversight, and verification protocols. The EAM baselines, targets, and incentives can be based on the assumptions, analyses, and forecasts supporting the programs themselves. Verification of the EAM can take advantage of the EM&V activities of the program itself (e.g., EM&V that determines the capacity savings of specific demand response programs), rather than trying to verify the effect of utility actions on the desired outcome (e.g., annual peak demand reduction on the entire system).
- Action-based EAMs are most appropriate during transition periods to encourage utilities to take specific steps towards a desired vision of the utility of the future.

Finally, to the extent possible, EAMs should be designed to reflect net benefits. This ensures that EAMs are designed cost-effectively and that utilities are compensated for the net benefit that their actions are creating.

1. INTRODUCTION

One of the central tenets of the New York Reforming the Energy Vision (REV) process is that utility financial incentives must evolve to reflect the changing expectations and demands on electric utilities. Under the REV framework, utilities will be required to work closely with third parties, distributed energy resources (DER) providers, and customers to support adoption of DERs. The New York Public Service Commission (NY PSC) has recognized that:

> "[T]hese new expectations run counter to conventional methods of operation and, importantly, also run counter to the implicit financial incentives that are embedded in the cost-of-service ratemaking model. If cost-of-service calculations are to remain the basis of utility rates for the foreseeable future, then creating new earning adjustment opportunities are both a fair and a necessary means of promoting change."[2]

To provide the incentives necessary to achieve its policy goals, New York has begun to establish performance incentive mechanisms, referred to as "Earnings Adjustment Mechanisms" (EAMs) for its distribution utilities.

Types of EAMs

EAMs vary according to how closely they are tied to specific utility actions. Broadly speaking, an EAM can be characterized as one of three different types:

1. Outcome-based: An outcome-based EAM follows the premise that regulators should define a policy objective—such as a megawatt reduction in system peak demand—but should not specify explicit actions that the utility must take in order to achieve the outcome.

2. Program-based: These EAMs measure how well the utility achieves an outcome through a particular program or initiative, especially those that have been approved by regulators. For example, an EAM might measure a megawatt demand reduction achieved through an energy efficiency or demand response program.

3. Action-based: Action-based EAMs are focused on a more detailed level of specific utility actions, but do not measure the broader impacts of that action. For example, an action-based EAM might measure the extent to which a utility facilitates the interconnection of distributed resources, but would not measure the megawatt demand reduction achieved by those resources.

Much discussion has centered around outcome-based EAMs in New York and elsewhere, given that they have the potential to spur innovation while achieving broad policy goals. Because of that, we have

[2] State of New York Public Service Commission, "Order Adopting a Ratemaking and Utility Revenue Model Policy Framework," page 59.

concentrated our analysis on this type of EAM, whereas we address program-based and action-based EAMs in a qualitative manner.

While the appeal of outcome-based EAMs is undeniable, target-setting and measurement can be challenging, because it can be difficult to determine whether an outcome was achieved due to utility actions. On the other side of the spectrum, action-based EAMs are relatively simple to implement because they measure specific utility actions, but they provide little assurance that broader policy goals will be achieved. In between these two options, program-based EAMs more directly tie utility actions to policy goals, but also have challenges related to measurement of progress towards those policy goals. For example, a utility may receive an incentive for achieving a certain number of customer deployments of energy efficient heating and cooling systems, but measuring the actual impact on reduced energy use or reduced peak demand attributable to these systems can be difficult.

The purpose of this report is to describe the various types of EAMs and to provide guidance on some of the key issues in designing EAMs. Specifically, this report will seek to answer a variety of questions, including:

- What are the key advantages and disadvantages of outcome-based EAMs?
- What are the key advantages and disadvantages of program-based EAMs?
- How can regulators and utilities address the challenges associated with measuring performance (e.g., determining baselines and targets)?

While exploring these questions, we keep in mind the variety of recommendations that were made by New York Department of Public Service (DPS) Staff regarding EAMs, including that:

- EAMs need not be limited to activities under direct utility control. Instead, they may be designed to measure outcomes that utilities can influence.
- EAMs should generally be designed to span multiple years to allow time for utility actions to take effect.
- Financial incentives for new EAMs should be positive only, rather than symmetrical.[3]
- Financial rewards should be determined for each individual utility on a case-by-case basis.[4]

[3] With the exception of customer engagement and interconnection.

[4] State of New York Public Service Commission, "Order Adopting a Ratemaking and Utility Revenue Model Policy Framework," page 55.

2. OUTCOME-BASED MECHANISMS

2.1. Description

Performance-based regulation is sometimes described as "hands-off" regulation, because theoretically it should reduce the extent to which regulators scrutinize utility management decisions. Outcome-based incentives are consistent with the hands-off regulatory approach, as they follow the premise that regulators should define a policy objective and reward the utility for achieving that objective, but should not specify or measure specific actions that the utility must take in order to achieve the outcome.

Outcome-based EAMs tend to address high-level outcomes that relate to the efficiency of the electricity grid. Examples of outcome-based EAMs include incentives related to:

- Peak demand;
- Customer energy intensity;
- Customer load factors;
- Substation utilization;
- Carbon emissions; and
- Transmission and distribution losses.

2.2. Primary Advantages

The primary advantages of outcome-based EAMs are their direct ties to policy goals and the flexibility they offer utilities. They focus on simple outcomes that regulators have determined should be pursued for the benefit of customers. By simply specifying the desired end result, outcome-based EAMs provide the utility with flexibility to achieve the outcome in the manner it sees fit, thereby engaging utility management and encouraging utility innovation. This flexibility also allows utilities to leverage actively engaged customers, DER developers, and other third parties in pursuit of the goal.

Similarly, outcome-based EAMs do not require regulators and other stakeholders to determine how best to achieve the desired outcomes. They do not require extensive program planning processes or regulatory approval and oversight of programs or initiatives. Importantly, outcome-based EAMs do not require extensive and potentially expensive efforts to evaluate, measure, or verify the results of specific programs or initiatives. Instead, they focus on important metrics that are relatively easy to measure, such as peak demand, customer energy consumption, or carbon emissions.

2.3. Primary Challenges

While outcome-based EAMs are simple in concept, implementing them can involve some complexities and challenges. The outcome itself may be simple to measure (such as MW of peak demand), but it only has meaning when compared to a counterfactual baseline. However, counterfactual baselines are often difficult to establish.

It can also be challenging to identify the best target, or set of targets, to be used for the purpose of determining financial incentives. Targets that are set too high might eliminate the potential for a utility to earn reasonable financial incentives, thereby defeating the purpose of the EAM. Targets that are set too low might unduly reward utilities and burden ratepayers without providing value.

Another challenge is that the cost of achieving a target might not be well known, hindering regulators' ability to ensure that the EAM will result in net benefits to customers. If a utility can pass the costs of achieving a target on to customers (either in the next rate case or immediately), then the utility will have a strong incentive to spend whatever it needs to in order to achieve the target. Further, there may be opportunities for the utility to be over-rewarded for actions it takes if these actions affect its ability to meet two different EAM targets.

These key challenges are addressed further below. Examples are provided in Section 2.4.

Defining the Baseline

To accurately assess the extent to which a desired outcome is achieved, it is necessary to first establish a counterfactual baseline of what would have occurred in the absence of the EAM. For example, when using a peak reduction EAM, it is first necessary to forecast the utility's peak demand under a business-as-usual case. This is typically done using econometric models that account for many factors that influence peak demand, such as weather, economic variables (e.g., economic output), adoption of federal energy efficiency standards, and the number of customers.

While econometric models are useful for making forecasts and determining business-as-usual scenarios, it is important to recognize their limitations. There are many different ways to design an econometric forecast, and many methodological choices and input assumptions can heavily influence the model and the forecast. It is critical to ensure that the model accurately accounts for the primary factors that might influence the baseline. This is necessary to provide a reasonable amount of confidence in the forecast.

Further, econometric models are based on historical data, but there is no guarantee that the future will look like the past. For instance, if central air conditioning is becoming more widely adopted, then hot summer afternoons may result in greater spikes in peak demand than they did ten years ago. While inclusion of a variable measuring air conditioner adoption might be helpful, there may not be sufficient data to include such a variable in the model.

Even if the adequate, desired data exist, a model can only provide predictions. A model should therefore be assessed based on the accuracy of its predictions. There are many methods for determining the accuracy of a model. One of the most common measures of accuracy for electric load forecasting models is the Mean Absolute Percentage Error (MAPE).[5] The MAPE is calculated as:

[5] Rafał Weron, "Electricity Price Forecasting: A Review of the State-of-the-Art with a Look into the Future," *International Journal of Forecasting* 30, no. 4 (October 1, 2014): 1030–81, https://doi.org/10.1016/j.ijforecast.2014.08.008.

$$MAPE = \frac{1}{n}\sum\frac{|Actual - Forecast|}{|Actual|} * 100$$

where n = the number of periods for which forecasted values are predicted and actual values exist.

The MAPE describes the average absolute percentage difference between the predictions of a model and the actual values. If the MAPE is large—say, 10 percent—then the predictions from the model should be viewed cautiously. Target-setting (discussed in the next section) should be conducted with the MAPE in mind. A target should not be set within the MAPE of a model; otherwise the target is within the average percent error of the model and incentives could easily be rewarded for an outcome that is not the direct result of utility efforts.

Finally, it is important to avoid overfitting when developing a model. Overfitting occurs when too many variables are included in a regression relative to the number of observations. An overfitted model often has poor predictive capabilities (for example, the MAPE tends to be larger) because they are highly sensitive to minor fluctuations in the data included in the regression.[6] When developing a model to forecast a baseline, it is important to carefully consider the variables and include only those that are conceptually relevant to the dependent variable. Comparing the predictive performance of different model specifications can assist in choosing the appropriate model to use for the forecasted baseline.

Setting a Target

Once a baseline has been developed and accepted, targets must be developed relative to the baseline. Targets are an important element of the EAM because they dictate the extent to which the utility will achieve the desired outcome. In addition, they determine the magnitude of the financial incentive that the utility may earn. Targets that are too easy to achieve will not provide much benefit to customers, while targets that are too difficult will undermine the utility's interest in achieving the desired outcome.

Setting a target for an outcome-based EAM can be somewhat arbitrary. For example, with a peak reduction EAM, what level of reduction warrants a utility incentive?

- Should the target be based on a reduction at a specific percentage level, e.g. five percent?
- Should it instead be based on one of various statistical measurements, such as the standard deviation or standard error?
- Should it be based on the level of the outcome that is estimated to be particularly beneficial?

These are questions that regulators and other stakeholders need to address to determine appropriate target levels. As a general principle, any financial incentive for the utility should be less than, and ideally

[6] Babyak, M. A. (2004). What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosomatic Medicine*, 66(3), 411-421.

a reasonable portion of, the net benefit associated with achieving a particular target. This will ensure that customers receive benefits from utility achievement of the EAMs.

Several New York utilities—including NYSEG and Rochester Gas & Electric (RG&E)—have proposed using targets based on the standard error. The standard error of the regression provides a measure of the distance between the fitted regression line and the actual values observed in the data,[7] thereby providing an indication of the accuracy of the regression model.[8] For example, 1.0 standard error is associated with an 83 percent level of confidence, while a standard error of 0.25 is associated with a 60 percent level of confidence. (The use of standard errors is demonstrated graphically in the example below for the customer energy intensity EAM.)

In the context of the EAMs proposed by NYSEG and RG&E, the standard errors represent bands that are used to create targets for the peak demand and energy intensity EAMs. If the actual values of future energy intensity and peak demand fall outside of the standard error bands, it can be said with a certain degree of confidence[9] that a factor external to the model (such as utility effort) truly affected the outcome.

In the context of the New York EAMs, the external factor assumed to cause this deviation from the forecast—i.e., the deviation from the business-as-usual case without the EAM—is utility effort to achieve the target. However, this assumption runs the risk of attributing too much of the deviation to utility actions. There may be other external factors that can affect an outcome—factors that are not captured in the model or are not the result of utility actions. Failure to account for these in the development of the baseline and targets can lead to attributing a shift in the outcome to utility action when some other factor may also have been in play.

Managing Costs

As noted above, the costs associated with meeting a target are generally not well known before-hand, since the manner in which a target is to be met is not approved in advance. This can create difficulty in ensuring that the EAM is cost-effective.

Avoiding Double-Recovery

Another challenge of outcome-based EAMs is setting target levels across EAMs in a way that avoids double-recovery of incentives. Many different utility actions may contribute to achieving a certain outcome. If the utility is rewarded for some of these actions through other EAMs, it creates the potential for double-recovery.

---

[7] Jack Schwager, *Futures: Fundamental Analysis*, 1st ed. (Wiley, 1995).

[8] Note that the standard error of a forecast is related, but not identical, to the standard error of the regression. The standard error of the forecast takes into account both the standard error of the regression, as well as the standard error of the independent variables at which the forecast is computed. The standard error of the forecast is always at least as large as the standard error of the regression.

[9] The degree of confidence is determined by the standard error level used to create the band.

For example, if the utility has one EAM for peak demand reduction and one for customer energy intensity, reducing customers' peak demands can help the utility achieve the targets for both EAMs and thus earn incentives for both. Or if the utility has an outcome-based peak reduction EAM and a program-based demand response EAM, then customer demand response savings will help the utility achieve the targets of and earn incentives for both of those EAMs. In order to avoid this, targets and benefit calculations will need to be coordinated across EAMs so that multiple utility activities are not double-counting the same benefits, and that the utility is not overly-rewarded.

2.4. Examples

In this section, we explore the development of outcome-based EAMs using two examples from New York: Peak Demand Reduction and Residential Customer Energy Intensity. While doing so, we refer to and discuss the challenges discussed above.

Peak Demand Reduction

In its 2017 rate case, Niagara Mohawk proposed creating an EAM for peak reduction.[10] The utility proposed to set the baseline as the weather-normalized non-coincident system peak load in 2016 (6,846 MW).[11] With its baseline set, it next proposed to tie its Peak Reduction EAM compensation to reductions in system peak load below the 6,846 MW baseline. The Department of Public Service Staff's response correctly points out the risks associated with Niagara Mohawk's proposed baseline and targets:

> "Using the peak load from 2016 would create a counterfactual baseline against which future metric performance would be measured, and is arbitrary. Furthermore, from a target-setting perspective, 2016 represents the single-highest coincident peak load, and the second-highest non-coincident peak load, within the last 11 years, and may therefore result in setting targets which require little utility effort to achieve."[12]

This response highlights the oftentimes arbitrariness and complexities of determining baselines. It also highlights the caution that needs to be taken while establishing targets and the difficulty involved in disentangling utility efforts while determining targets and appropriate financial incentives.

Customer Energy Intensity

The New York Public Service Commission and the utilities have been exploring the development of customer energy intensity EAMs. These EAMs are often broken out by commercial energy intensity and residential energy intensity. This section focuses on the Residential Customer Energy Intensity EAM as proposed by NYSEG.

[10] Niagara Mohawk Power Corporation, "Direct Testimony of the Electric Customer Panel," April 28, 2017.

[11] Staff Earning Adjustment Mechanisms Panel, "Prepared Testimony of Staff Earning Adjustment Mechanisms Panel," August 2017, page 13.

[12] *Ibid*, page 18.

Reviewing Historical Data

Figures 2 and 3 below show historical residential customer counts and energy use in megawatt hours (MWh) in NYSEG's service territory since 1980. Figure 2 shows a steep increase in the number of customers until the mid-1990s, at which point the customer count begins to slow, and then plateau. Figure 3 depicts the seasonal volatility of energy use. It also depicts a steady long-term increase in energy use, a trend that is at least partly due to the increase in customer count.

Figure 2. Historical Monthly Residential Customer Counts

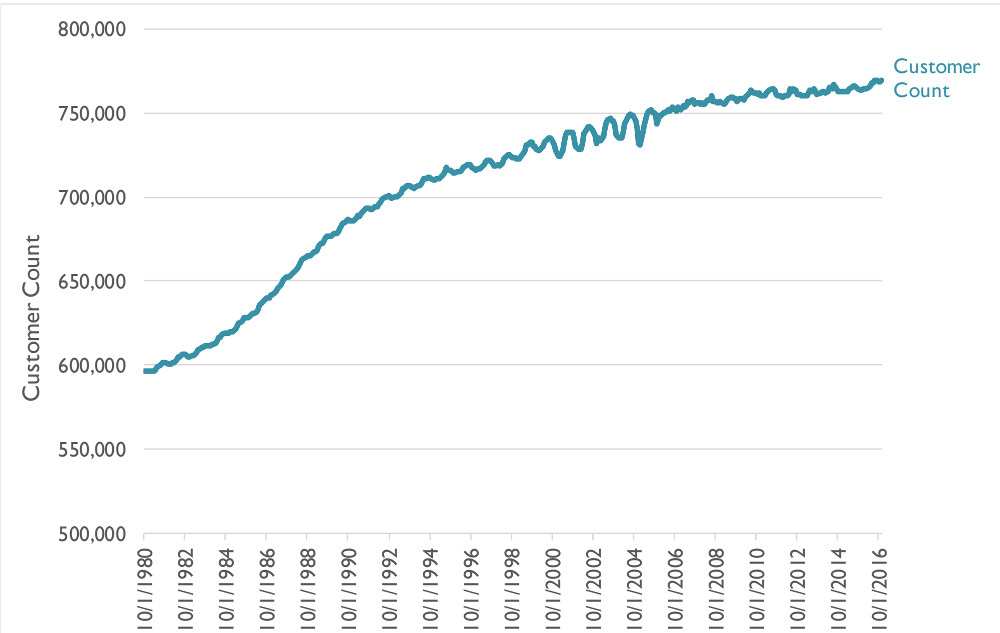
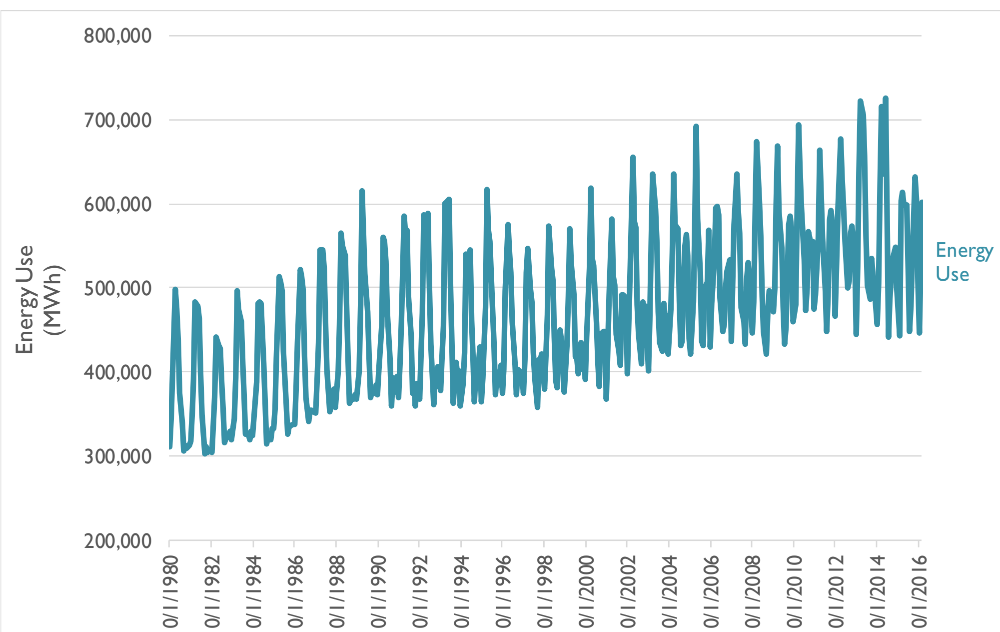


Figure 3. Historical Monthly Residential Energy Use (MWh)



Taking the annual average smooths out the volatility of the data series, particularly the seasonal, cyclical volatility that exists in the energy use data. However, there remains a wide range of values in the average energy use data (Figure 5), which may present a challenge to developing the counterfactual baseline.

Figure 4. Average Annual Customer Count

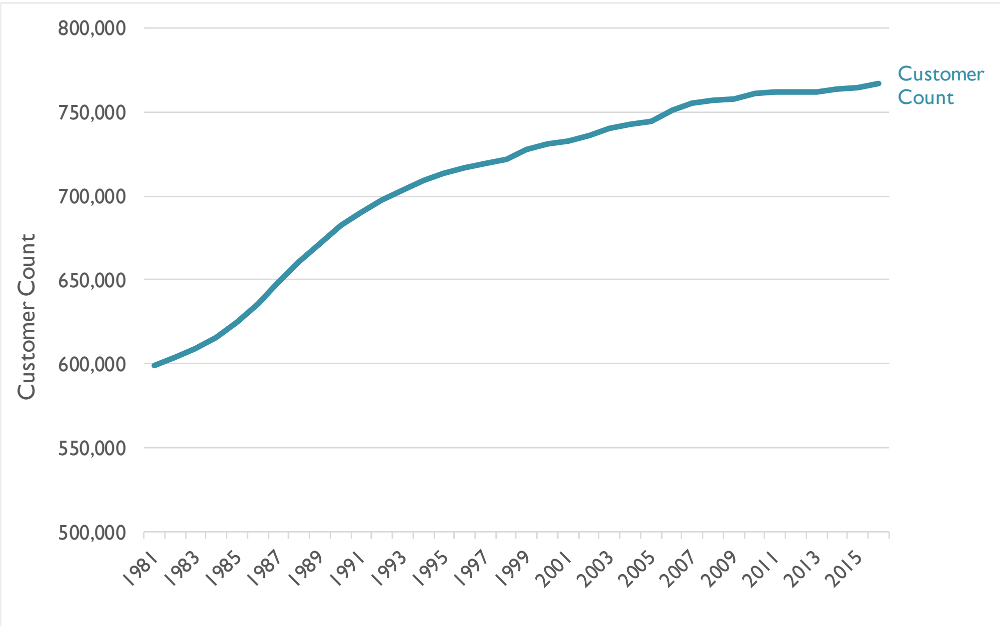
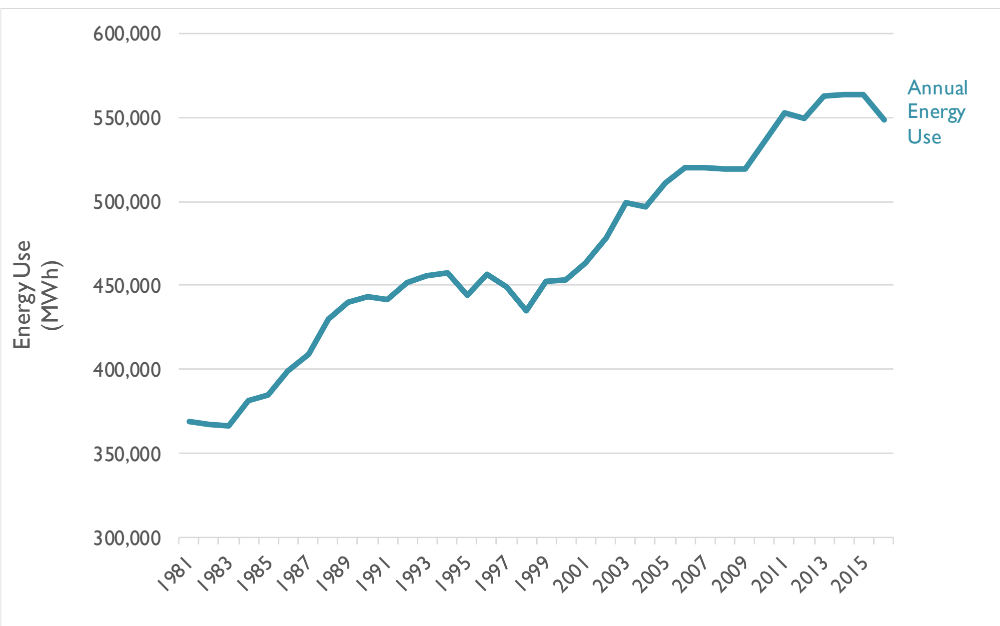


Figure 5. Average Annual Energy Use (MWh)



Creating a Regression Model

The next step is to identify which variable to use as the dependent variable in the regression analyses. The regression model, discussed in more detail below, forms the basis for forecasting the counterfactual baseline.

NYSEG proposed treating customer count and energy use as dependent variables in two separate regressions. However, it is worth considering whether energy use and customer count should be modeled together as a single variable: energy intensity. Such a consideration requires exploring the energy intensity data, expressed in MWh per customer and derived by combining the customer count and energy use data.

Figures 6 and 7 show the monthly and annual energy intensities. The monthly data series exhibits similar seasonal volatility as the energy use data series. As indicated in Figure 7, much of the volatility is smoothed out in the annual series. However, similar to the monthly energy use data series, there is still significant variation in the annual data series.

Figure 6. Monthly Historical Energy Intensity (MWh/Customer)

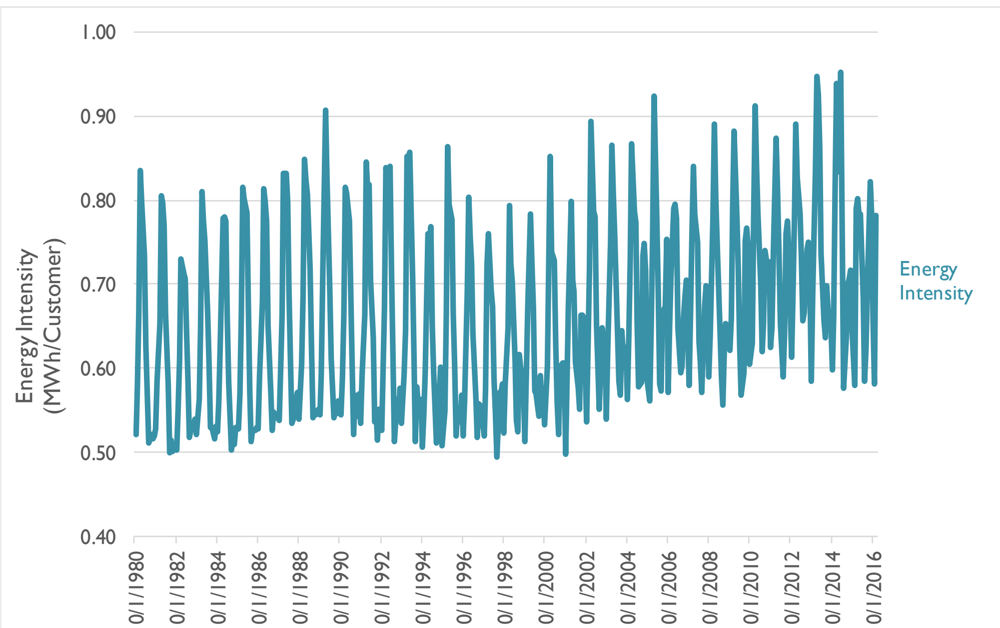
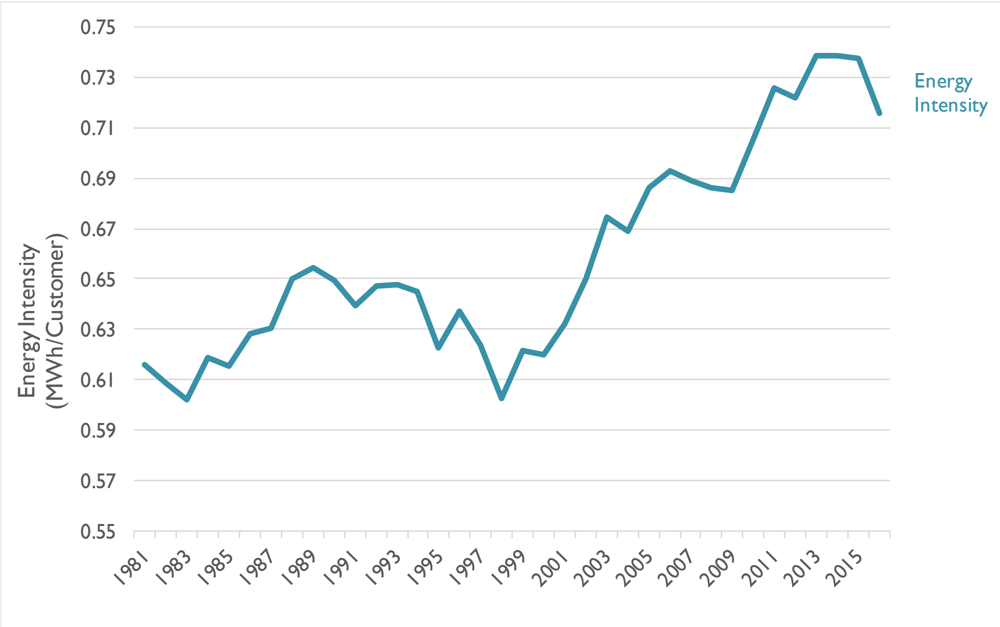
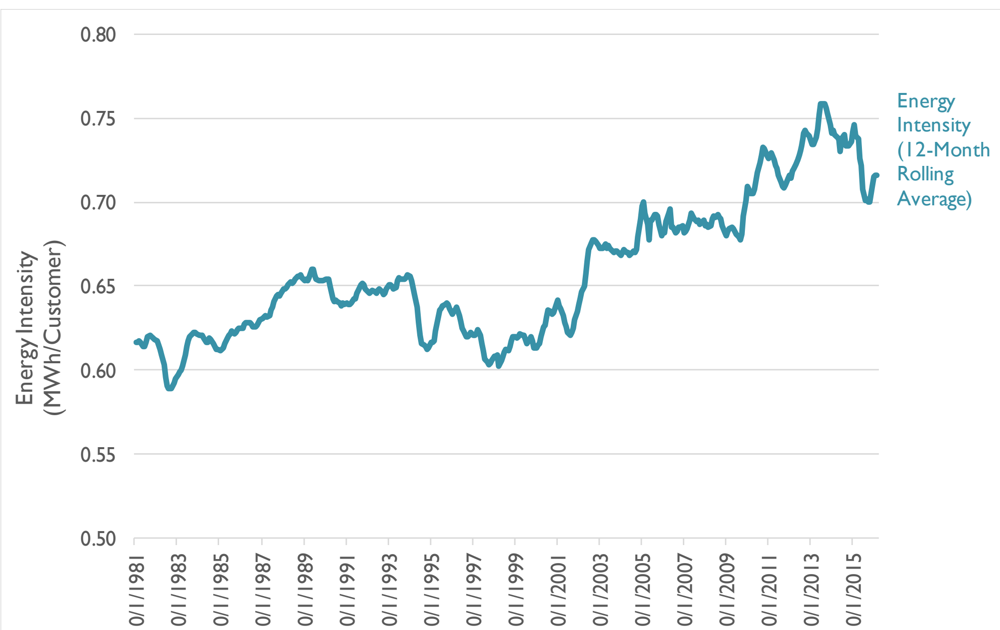


Figure 7. Historical Energy Intensity (MWh/Customer)



A final consideration for the dependent variable is a 12-month rolling average of monthly energy intensity values. Figure 8 presents the results from the 12-month rolling average. Using a rolling average smooths out the volatility of the monthly energy intensity presented in Figure 6, but has greater volatility than the annual energy intensity presented in Figure 7.

Figure 8. Energy Intensity (MWh/Customer) – 12-Month Rolling Average



We developed a regression model using historical data to illustrate some of the issues that arise when determining baselines and setting targets. Most of the design and assumptions in our model are based on the regression model proposed by NYSEG.[13]

For our energy intensity regression model, we chose the annual energy intensity (Figure 7) as the dependent variable (measured as energy usage per capita). The annual energy intensity effectively removes the immense seasonal volatility that exists in the monthly values. It simplifies the regression such that we only have to run a single model for both variables of interest (customer count and energy use).

The independent variables included in our regression—based off the regressions proposed by NYSEG—are: consumption days, heating degree days (HDD), cooling degree days (CDD), residential electricity prices, disposable personal income, and residential energy efficiency savings. We have excluded forecasted residential customers and the population of the service territory, because energy intensity is defined as total energy divided by total customers, and thus these variables do not contribute to the predictive capability of the model. In other words, forecasted customers should not be an independent variable because it is already a part of the dependent variable.

Determining the Baseline

Finally, we use the coefficients from the regression and the forecasted values for the independent variables to create forecasted energy intensity values through 2020.[14] To do so, we multiply the coefficient for each independent variable by the forecasted value for that year. The values that are calculated by multiplying the variable coefficients by the forecasted values are aggregated to arrive at a forecast for a given year. For example, the equation below represents the generic calculation of forecasted values:

$$Y_t = B_0 + B_1 * X_{1,t} + B_2 * X_{2,t} + B_3 * X_{3,t}$$

In this equation, Y_t is the forecasted value for a future year t, B_0 represents the intercept from the regression, $B_{1,2,3}$ are the coefficients for three variables, and $X_{1,2,3,t}$ are the values of the independent variables in future year t.

The forecasted values, Y_t, are used as the baseline from which the target levels will be set.

We have weather-normalized the energy intensity series by setting the values of the heating degree days and cooling degree days variables to the average of the values in the historical data. This shows the energy intensity data series as it exists outside the influence of weather.

[13] See: Electric_Forecast_DriversR1.xlsm, provided in response to DPS-2 in Cases 16-M-0429 and 15-E-0283, et al. on July 19, 2017.

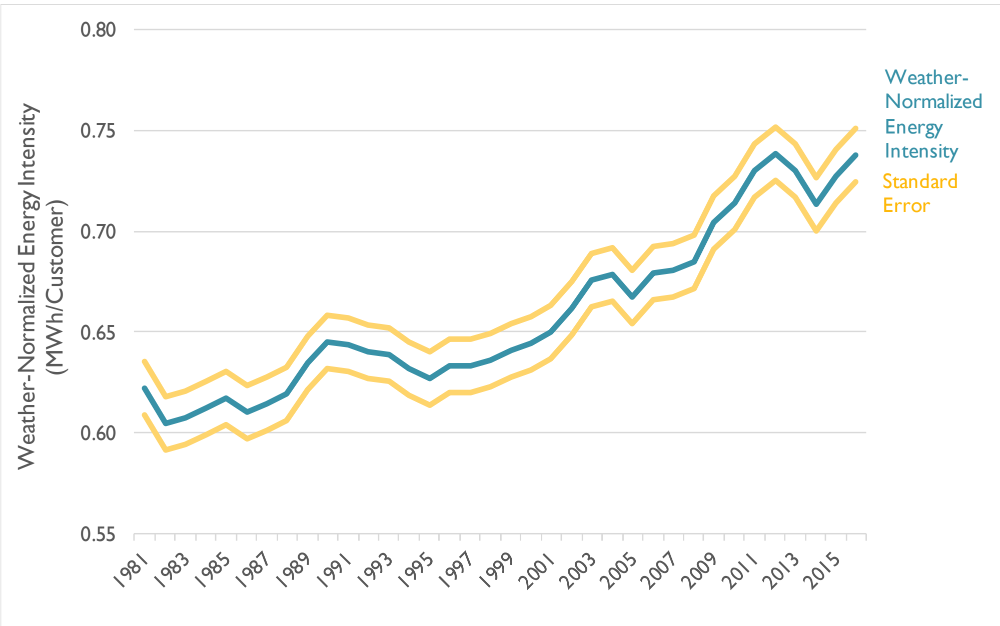
[14] We did not have forecasted CDD and HDD data, and so instead used the average monthly values from the historical data as the forecasted values.

Setting a Target

A critical statistic from a regression equation is the standard error. As mentioned earlier, the standard error of the regression measures how much the fitted line deviates from the actual values. Therefore, it is useful in setting estimated bounds on a regression model and for setting targets.

The regression we ran had a standard error of 0.01 MWh/customer. Figure 9 plots the weather-normalized annual energy intensity series along with two bands representing the series plus and minus one standard error.

Figure 9. Annual Energy Intensity with Standard Errors

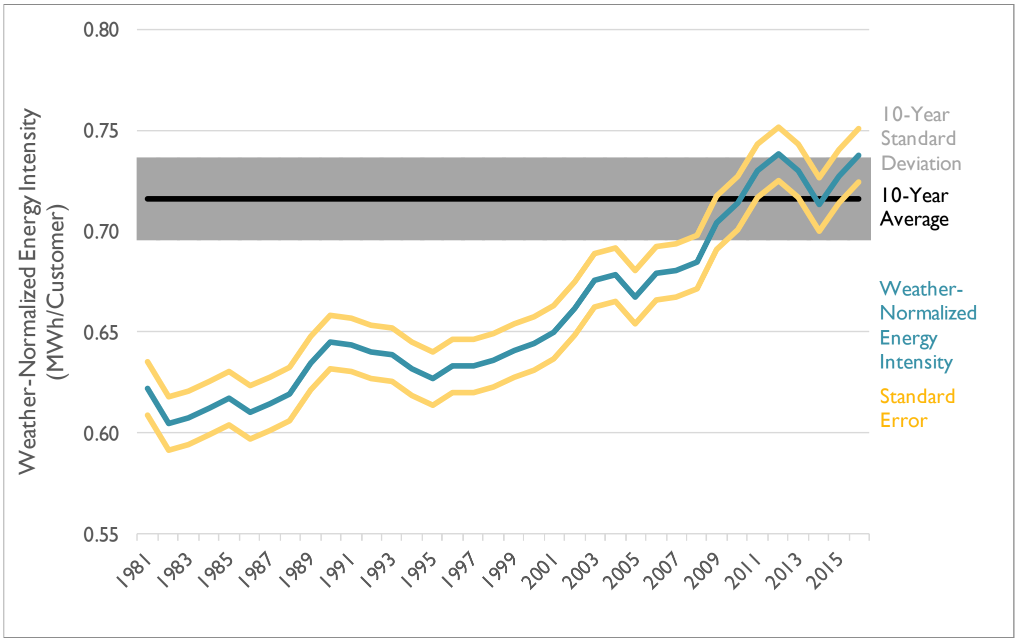


An alternative to using the standard error is to use the standard deviation of the historical values of interest (e.g., energy intensity) for target setting. Use of the standard deviation is particularly common where a regression model is not used to determine a baseline. Approximately two-thirds of the data will fall within one standard deviation of the mean, assuming a normal distribution. Thus, the standard deviation provides some indication of what values can be expected in the future, assuming no substantial changes in other factors. When used in target-setting, a standard deviation might be calculated based on the past ten years of data. The standard deviation could then be used as a deadband, where the utility would only be rewarded (or penalized) for performance that falls above or below the standard deviation. However, this approach is only effective if all other variables that impact the utility's ability to achieve the target remain relatively stable.

Figure 10 again shows the annual energy intensity series plotted with a band representing the series plus and minus one standard error, as well as one standard deviation of the most recent ten years of values from the series. We chose to use only the past ten years of data to calculate the standard deviation, since there is a clear upward trend in the longer-term data.

The figure shows that, while the standard deviation is useful for describing the volatility of a data series for the last ten years, by itself it does not capture trends in the data, and thus is of limited usefulness in predicting future values. This suggests that the standard error of the regression is generally a better mechanism for setting EAM targets than the standard deviation, particularly when the underlying data show clear trends.[15] An alternative method would be to apply the standard deviation to the forecasted baseline.[16]

Figure 10. Annual Energy Intensity with Standard Errors and Standard Deviation



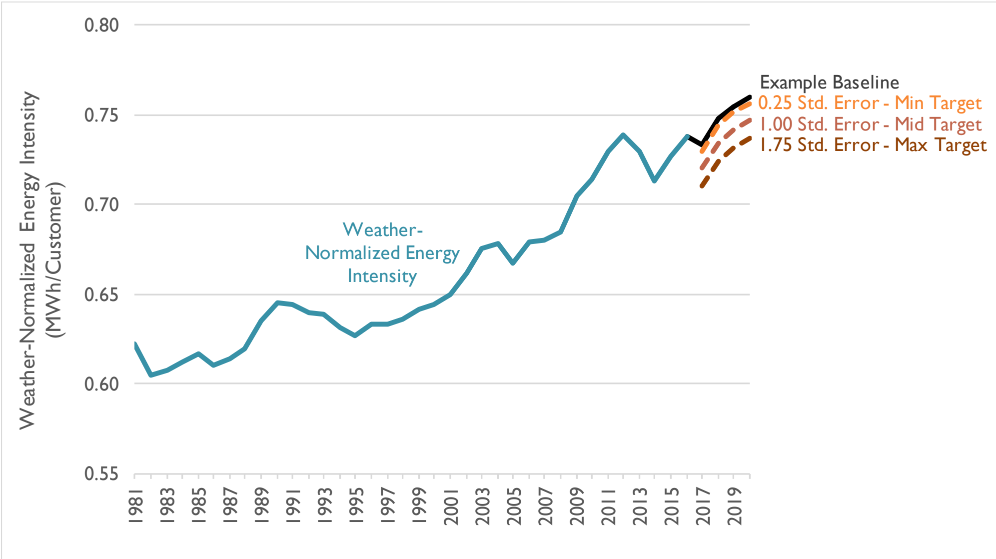
In Figure 11 we present three potential targets: the forecasted energy intensity minus 0.25 standard error, the forecasted energy intensity minus 1.0 standard error, and the forecasted energy intensity minus 1.75 standard error. These targets are based off NYSEG's proposed minimum, mid, and maximum target levels, respectively. They represent the bands beyond which we can say with 59.5, 83, and 94.5 percent confidence, respectively, that factors beyond those captured in the model explain the deviation from the forecast.[17] Utility action would be one of the external factors, perhaps the most likely external factor, explaining this deviation.

[15] See footnote 8 above. Technically, the standard error for a forecast should incorporate both the standard error of the model and the standard error of the independent variables at the point where the forecast is calculated. It is unclear whether NYSEG's proposed standard error is referring to the standard error of the regression or of the forecast, but our general conclusions are the same.

[16] If one has performed a regression analysis, there is little reason for using the standard deviation instead of the standard error, since the standard error incorporates the variation of both the dependent variables and the independent variables. However, if a baseline is developed through an alternative method, a standard error may be unavailable and thus appropriate to apply to the baseline.

[17] Staff Earning Adjustment Mechanisms Panel, "Prepared Testimony of Staff Earning Adjustment Mechanisms Panel," Cases 17-E-0238 & 17-G-0239, August 2017, page 48.

Figure 11. Annual Energy Intensity with a Baseline and Targets



It is notable that the actual value for energy intensity in 2014 falls below the 2017 targets set using 0.25 and 1.0 standard errors in 2017. This suggests that—due to the volatility that exists within the historical energy intensity data—there will be years in which energy intensity can vary significantly regardless of utility action.[18]

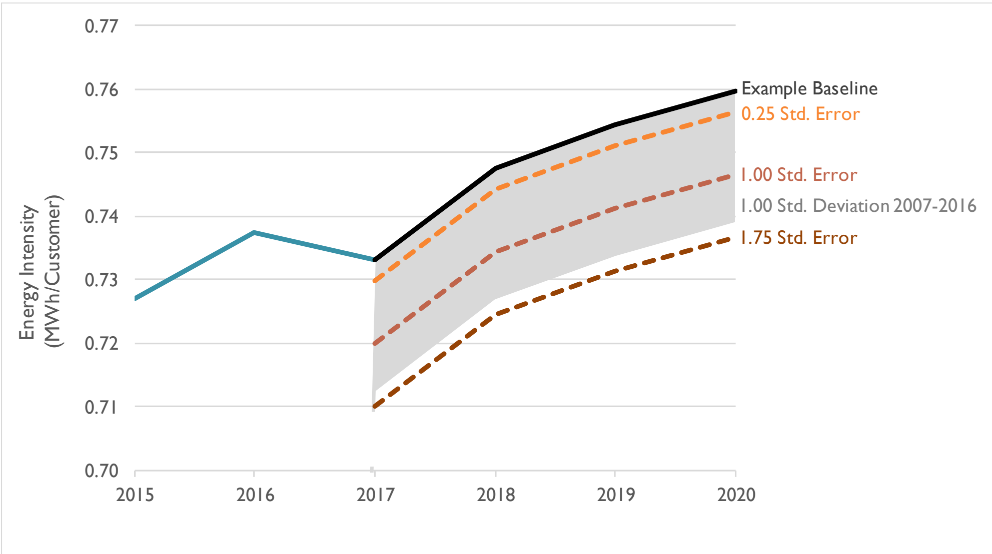
Further, our regression model has a MAPE of approximately 1.5 percent. To gauge whether one standard error represents a reasonable target, we can use the MAPE from the model. The average of the weather-normalized energy intensity data series is 0.66 MWh/customer. Multiplying this by 1.5 percent converts the MAPE to just under 0.01 MWh/customer, which is approximately equivalent to one standard error. Therefore, in this instance, one standard error is similar to the average model error. However, 0.25 standard error exists well within the range of the average model error, which indicates that a target set at 0.25 standard error could easily be achieved without utility action and without the EAM.

Finally, Figure 12 shows the same example baseline and targets as Figure 11 with the addition of a shaded region that represents the baseline less the standard deviation from the past ten years. The standard deviation (representing the volatility of the data series from the past ten years) is larger than the targets based on 0.25 and 1.0 standard errors. This should be a cautionary sign, as it suggests that if the degree of volatility from the past ten years continues in the future, the energy intensity in some future years should be expected to fall below these targets regardless of utility action. Though our analysis uses the standard error as the metric from which targets are set, one standard deviation is

[18] One way to avoid rewarding a utility for a year of low energy intensity resulting from exogenous, non-utility action is to set rewards based on longer-term energy intensity reductions measured over several years. By tying the reward to multiple years of reductions, a "fluke" year with a significant non-utility reduction in energy intensity will not be rewarded – only persistent, utility-influenced reductions will be rewarded.

often used to create a "deadband" in which a utility neither receives a penalty nor a reward.[19] If the underlying data are normally distributed, there is roughly a 64 percent chance that an observation will lie within the deadband, and therefore it is not recommended to create targets within the deadband.[20]

Figure 12. Example Baseline and Targets using Standard Error and Standard Deviation



The regression results presented in Figures 11 and 12 highlight the key challenges with determining EAM baselines and targets, including questions such as:

- Is the baseline likely to be a reasonable forecast of the business-as-usual scenario without the EAM?
- Should the targets be based on standard errors, and if so, how many standard errors?
- Does it make sense to provide a financial incentive, even the minimum level of financial incentive, for an outcome at 0.25 standard errors, when this suggests that there is roughly 60 percent confidence at most that the utility was responsible for this deviation from the baseline?

It is premature for us to provide definitive answers to these questions. Nonetheless, our preliminary analysis suggests that outcome-based metrics should be used with some caution. They could be appropriate where the standard error is sufficiently small compared to outcomes that could reasonably be achieved by utilities, such that confidence would be high that you could link cause and effect between those actions and the measured outcomes. Incentives could also be set such that they ramp non-linearly with the measured outcome. For example, they could be set at modest levels when

[19] Synapse Energy Economics. 2015. "Utility Performance Incentive Mechanisms: A Handbook for Regulators," page 38.

[20] Hanser, Philip, et al. 2013. "Review and Analysis of Service Quality Plan Structure in the Massachusetts Department of Public Utilities Investigation Regarding Service Quality Guidelines for Electric Distribution Companies and Local Gas Distribution Companies," page 10.

confidence was only moderate that utility action would lead to the outcome and be set at more aggressive levels once confidence passed a certain threshold.

Additionally, incentives provided for multiple years of performance can limit the risk that an incentive is awarded based on an outlier year or unexplained volatility (such as 2014 in Figure 11 above) in the measurement of the EAM, and can instead provide rewards based on longer-term and more durable performance. This comes with a tradeoff of delaying incentive payouts to utilities, which might decrease their attractiveness to management and exert less influence over decision making.

3. PROGRAM-BASED MECHANISMS

3.1. Description

Program-based EAMs are more prescriptive than outcome-based EAMs, as they measure utility performance related to a specific program or initiative (such as energy efficiency or demand response programs). Program-based EAMs have both a specific policy outcome (such as energy or peak demand savings), as well as specific approved means of achieving the outcome (such as through a certain mix of energy efficiency programs). Because of this, program-based EAMs leave less room for utility innovation, but may provide greater certainty that the outcomes are cost-effective and result from utility actions.

Examples of program-based EAMs include:

- Energy efficiency programs;
- Demand response programs;
- Net metering and other distributed generation initiatives;
- Electric vehicle initiatives; or
- Storage initiatives.

These programs and initiatives are typically reviewed and approved by regulators prior to being implemented. The programs are supported by a variety of elements, including:

- A specific funding source from customers;
- A demonstration that the program or initiative is cost-effective, including detailed estimates of costs and benefits;
- Program designs, proposed budgets, and proposed savings estimates;
- Mechanisms for evaluation, measurement, and verification of the impacts; and
- Processes to allow for stakeholder engagement throughout the development, review, and approval of the program or initiative.

3.2. Primary Advantages

A key advantage of program-based EAMs is that they are typically applied to programs or initiatives that are reviewed and approved by regulators, with significant stakeholder input. This provides an opportunity for regulators and others to provide guidance on how to achieve desired outcomes, as well as improves the likelihood that utility actions will lead to those outcomes. Program-based EAMs also typically require a more concrete demonstration that the programs are cost-effective, providing regulators with a better sense of whether the costs of EAMs are justified by sufficient net benefits for customers.

In addition, program-based EAMs are often applied to programs that already exist, either at the utility or elsewhere, providing a body of evidence to use for the design and implementation of the EAM. This greatly facilitates the development of counterfactual baselines and targets for EAMs. It can also help with the measurement and verification of how well the EAM targets were achieved.

Energy efficiency programs are one example of a good candidate for program-based EAMs. They are typically the result of considerable planning efforts, stakeholder input, and regulatory review. Energy efficiency programs are typically required to pass a relatively high standard for determining that they are cost effective. Furthermore, there is a wealth of experience and expertise related to the practices of measuring and verifying energy efficiency program savings. All of these elements help to create a set of utility programs that are very likely to lead to a desired outcome: increased end-use efficiency and reduced electricity system costs.

3.3. Primary Challenges

One of the challenges with program-based EAMs pertains to the development of the program itself. A well-designed program requires sufficient planning, analysis, regulatory review, and stakeholder input in order to achieve the desired outcome. For some programs that already exist and are delivering the desired outcomes, there may be little need for improving program design to support an EAM, while for other programs (e.g., programs to support innovative storage initiatives) there may be the need for significant analyses and regulatory input before they are ready to be supported with an EAM.

Setting the baselines for program-based EAMs can be challenging in some circumstances. Setting baselines requires reasonable estimates of the desired outcome in the absence of the program in question (e.g., energy savings from federal appliance standards or the natural rate of adoption of electric vehicles).

Target setting for program-based EAMs can also be challenging where there is little utility experience, in state or elsewhere, to draw from. Setting good targets requires reasonable estimates of how difficult an outcome is to achieve, what programmatic activities would be necessary to achieve the targets, the costs and benefits of implementing such programs, and the likely customer or third-party involvement in achieving the outcome.

Perhaps the most significant challenge with program-based EAMs pertains to the measurement and verification of performance. If utilities are to be provided with financial incentives for programmatic results, it is important that those results are measured and verified with a reasonable degree of certainty. Relative to outcome-based EAMs (e.g., a peak reduction EAM based upon peak demand), measuring the performance of program-based EAMs can be resource-intensive and contentious, and if not performed well can lead to uncertain results. Measuring performance for programs that utilities and regulators already have substantial experience with (e.g., energy efficiency programs) can be relatively straightforward using well-established techniques, while measuring performance for other types of programs (e.g., demand response programs) might involve greater challenges. Given the significance of this issue, we address it in more detail in the following section.

3.4. Measuring Performance

There is a rich history of evaluation, measurement, and verification (EM&V) techniques for assessing the efficacy of utility-sponsored energy efficiency programs. International protocols have been established to guide practices and promote consistency across countries.[21] There is also an entire industry of independent contractors who conduct independent EM&V studies to document the savings of utility energy efficiency programs. Regulators throughout the United States regularly rely upon such studies to review, approve, and justify spending on energy efficiency programs. These studies typically cost only three to five percent of the efficiency program budgets.

Energy efficiency programs employ a variety of performance measurement approaches, including econometric models, self-reports, random control trials, deemed savings, or advanced measurement and verification. While these approaches are well-established for energy efficiency programs, they are not as well established for other types of DERs. Further, even well-established EM&V approaches are subject to some level of uncertainty.[22] Questions associated with the impact evaluation of program-based EAMs may include:

- Did the customer actually reduce load due to the program, or did the customer reduce load for other reasons?
- Does the control group include customers that are similar to the customer reducing load? This can be particularly difficult for large industrial customers with unique load profiles.
- Does the measurement technique accurately capture DERs that operate in an atypical manner (such as storage)?
- Do laboratory or modeled estimates hold when extrapolated across measures that are deployed in a variety of scenarios and usage conditions?

Moving forward, more advanced methods for EM&V may be necessary for the evaluation of new types of programs and technologies and to improve the accuracy of EM&V for well-established programs. The

[21] International Performance Measurement and Verification Protocol, available at https://evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp

[22] In New York, the minimum level of certainty is set at 90%.

New York Public Service Commission acknowledged this need in its January 21, 2016, order in Case 14-M-0094, writing: "We restate here the Commission's interest in exploring how advances in technology may be used to challenge and enhance our traditional approaches, and minimize associated costs, to EM&V. We direct Staff, in consultation with the CEAC [Clean Energy Advisory Council], to conduct a review of the current evaluation guidelines… [and] determine what changes are necessary to meet the current and future needs of our clean energy programs…." The following section describes some of these new approaches in greater detail.

Advanced M&V

New information and communications technologies such as smart meters, smart thermostats and other connected devices, and non-intrusive load metering devices are changing the way that measurement and verification of energy efficiency programs is conducted.[23] These technologies extract granular energy consumption data and allow new data analytics software to store, track, and analyze the data in near real time using cloud-based software. This capability allows program administrators to implement automated M&V, which takes advantage of automated data processing to produce building energy profiles, estimate savings potential, or estimate whole-building energy savings in near real time. These new approaches for evaluating measures, projects, and programs based on emerging technologies are often referred to as "advanced M&V" or "M&V 2.0".

New York's 2016 EM&V guidelines state that "Program administrators and evaluators are encouraged to use advanced M&V techniques when appropriate and cost effective, to collect, aggregate and analyze data."[24] The advancement of data availability and data analytic practices will allow for faster, more granular, and more encompassing data analyses. Further, the practices associated with M&V 2.0 will also facilitate impact measurement for many other utility programs. However, different types of resources create different challenges, and the measurement and verification practices might need to evolve to address some of these. For example, customers with behind-the-meter storage may exhibit vastly different consumption patterns relative to customers without storage, making the development of a counterfactual baseline even more difficult to establish.

4. Action-Based Mechanisms

4.1. Description

Instead of targeting a broad policy outcome or a specific program, action-based EAMs directly measure utility actions. Action-based EAMs are particularly useful when a regulator wishes to encourage the

[23] Details of these ICTs are described in: DNV GL 2015, *The Changing EM&V Paradigm*; and, ACEEE 2015, *How Information and Communications Technologies Will Change the Evaluation, Measurement, and Verification of Energy Efficiency Programs*.

[24] New York Office of Clean Energy. 2016. "Evaluation, Measurement & Verification Guidance," page 8.

utility to take a certain action, but the outcome of that action is uncertain or largely outside of the utility's control.

For example, a jurisdiction may have a policy goal of adopting electric vehicles in order to reduce greenhouse gas emissions. An action-based EAM might target the installation of EV charging stations, but would not measure the extent to which customers actually adopt electric vehicles or the extent to which greenhouse gas emissions are reduced.

Other examples of action-based EAMs might include:

- Introduction of time-varying rates (but not the number of customers who adopt them);
- Timeliness for interconnecting distributed generation;
- Installation of energy storage facilities; and
- Provision of customer and third-party access to information (e.g., websites for customers to access their energy usage, or information for third parties regarding the distribution grid and distribution system planning).

4.2. Primary Advantages

Action-based EAMs can provide utilities with an incentive to take actions that may have hard-to-quantify net benefits but that clearly help to achieve certain outcomes or policy goals. For example, if regulators wish to promote customer empowerment and market animation, an action-based EAM could incentivize the development of a web portal providing relevant information to customers and third-party developers. In this example, the web portal might lead to better information for regulators and other stakeholders to help inform policy and planning decisions; better information for customers to help inform distributed energy resources installation and operational decisions; and better information for third-party vendors who might install or operate distributed energy resources. Identifying the net benefits of the web portal or the specific efficiency savings or DER generation resulting from such a web portal would be challenging and highly uncertain.

While the net benefits associated with action-based EAMs may be difficult to measure, the costs can generally be estimated and approved in advance, thus providing some certainty to regulators and other stakeholders.

Action-based EAMs also offer the benefit of being relatively easy to measure and verify. This typically requires confirmation that the specific action occurred (e.g., a web portal was created), and that the action generally achieves its intended objective (e.g., the web portal is well-designed, and will provide customers and third-parties with timely, useful information).

4.3. Primary Challenges

A significant challenge associated with action-based EAMs is ensuring that the action is sufficiently tied to an intended outcome or benefit. While the action itself is easily measured, there exists little to no focus on the extent to which the action helps achieve a desirable outcome, e.g., emissions reductions or peak load reductions. This challenge highlights the situations in which it may be appropriate to use action-based EAMs; they are useful primarily when developing incentives for utility efforts that do not contain quantifiable outcomes or for which the net benefits are ambiguous, but for which utility actions are deemed to be important in contributing to overall policy objectives.

Over the long-term, action-based incentives may be less desirable than program-based or outcome-based incentives because they are not as closely linked to desired outcomes. Action-based incentives may best be used to enable certain utility activities or investments, which will in turn support programs and initiatives that will achieve the desired outcomes. They may be best used as a transitional approach where they can lead to program- or outcome-based EAMs.

5. RECOMMENDATIONS

Each type of EAM comes with its own set of advantages and challenges. When developing an EAM, all the challenges need to be carefully considered and steps should be taken to ensure that EAMs are effective, provide sufficient incentives, appropriately reward utilities for their actions, and result in net benefits to customers.

While developing EAMs, commissions should consider a balanced approach that avoids placing too much weight on any one type of EAM. Each type of EAM should be used where it is most appropriate:

- Outcome-based EAMs should be used when a specific outcome is desired and it is in customers' interest to allow the utility to pursue the outcome with innovative approaches that are not determined or overseen by the regulators. Outcome-based EAMs should recognize the challenges associated with verifying the extent of a utility's influence on the desired outcome. They should be designed to provide appropriate levels of incentives that are tied to desired outcomes where confidence is high that utility actions will produce those outcomes.

- Program-based EAMs should be used to support utility initiatives that include pre-determined program designs, regulatory oversight, and verification protocols. The EAM baselines, targets, and incentives can be based on the assumptions, analyses, and forecasts supporting the programs themselves. Verification of the EAM can take advantage of the EM&V activities of the program itself (e.g., EM&V that determines the capacity savings of specific demand response programs), rather than trying to verify the effect of utility actions on the desired outcome (e.g., annual peak demand reduction on the entire system).

- Action-based EAMs are most appropriate during transition periods to encourage utilities to take specific steps towards a desired vision of the utility of the future.

All types of EAMs should be designed to reflect net benefits to customers, to the extent possible. This ensures that EAMs are designed cost-effectively and that utilities are rewarded for the net benefit that their actions create.