10680 White Rock Rd, Rancho Cordova, CA 95670
Phone: (916) 844-0157
Email: CChappell@trcsolutions.com

C&S Attribution Study Report

March 6, 2019

Submitted To:

Southern California Edison
Emrah Ozkaya

Table of Contents

TABLE OF FIGURES

I. Executive Summary

The California investor owned utilities (IOUs) have developed the codes and standards advocacy programs (C&S Program) as an integral part of the portfolio of statewide efficiency programs. Codes and standards are a significant means to establish and ensure market transformation. Codes and standards advocacy comprises a portfolio-level strategy that complements incentive and information offerings in several ways. The C&S program serves as a link from research and emerging technology programs, to incentive and information programs, to code readiness, and to adoption of specific requirements into codes and standards. Through the Advocacy subprograms, the IOUs work to transition to regulatory intervention to maximize portfolio energy savings. The codes and standard program likely represents the largest source of energy savings for the IOUs.

To calculate those savings, the code and standards program has adopted a codes and standards attribution model which provides a framework to evaluate the codes and standards advocacy process. The model includes five major components: Potential Savings, Gross Savings, Net Standard Savings, Net Program Savings and Allocated Savings. Chapter 3 provides greater detail of each component of the model and how it fits into the bigger picture of the codes and standards program. This executive summary provides a brief overview of the study methodology, summarizes the key issues uncovered during the analysis phase and details the recommendations that IOUs should consider.

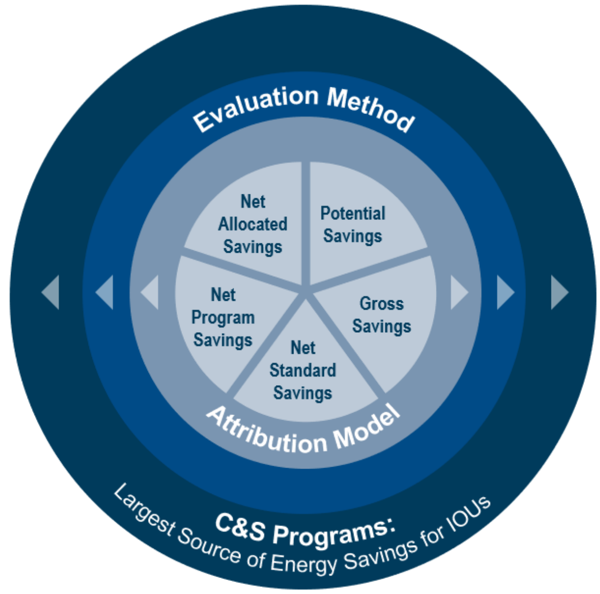


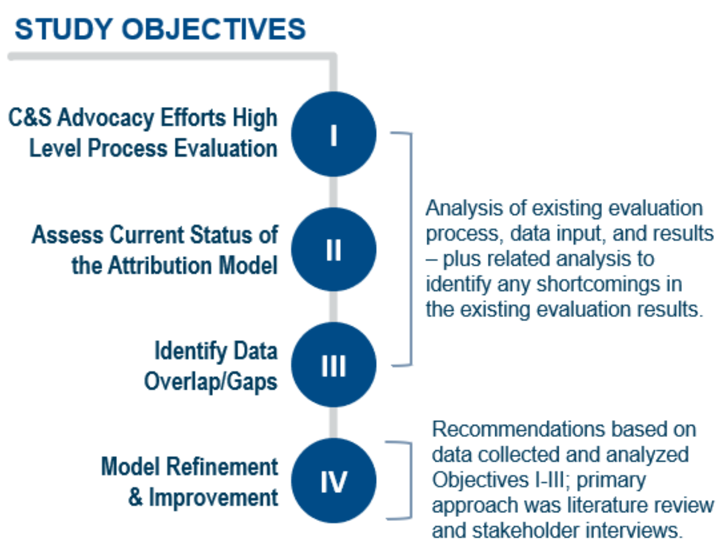
Figure 1: Model & Method for C&S Program

1.1 Study Methodology Overview

The purpose of this study is to conduct a high-level process evaluation of the overall IOU codes and standards advocacy efforts and revisit the attribution methodology to provide recommendations for improving the methodology and the evaluation approach. The goal of the research is to develop recommendations for updates to the model that best reflects codes and standards program contributions.

Problem Statement to guide our investigation and analysis:

"How can the CPUC codes and standards impact evaluation process increase transparency, streamline, and simplify existing methodologies, and improve accuracy of the final net savings allocation to each utility as it relates to their support in the advancement of codes and standards in California?"



1.2 Key Issues

To guide our investigation and analysis, we considered how the California Public Utilities Commission (CPUC) codes and standards impact evaluation process could increase transparency, streamline and simplify existing methodologies, and improve accuracy of the final net savings allocation to each utility as it relates to their support in the advancement of codes and standards in California. This framing of the problem allowed us to identify key issues related to transparency, streamlining, and accuracy within the overall model as well as each component of the model, as summarized in Figure 2.

ACCURACY	STREAMLINING	TRANSPARENCY
	POTENTIAL SAVINGS	
Lack of accuracy when choosing model baselines and when defining standards that involve multiple efficiency measures.	Opportunities to streamline when using CASE studies and other data sources to inform Unit Energy Savings verification.	Lack of transparency when establishing industry standard practice when there is no prior standard.
	GROSS SAVINGS	
Accuracy issues appear since the compliance rate is based on a binary approach which under-estimates compliance in terms of energy impact. The current gross savings approach does not capture actual market changes that happen after a new standard takes effect.	Opportunities to streamline the compliance assessment process should be developed to reduce cost and time it takes.	Lack of transparency in understanding how over-compliance is treated in the impact evaluations.
	NET STANDARD SAVINGS	
There are limited data sources for providing actual naturally occurring market adoption (NOMAD) information and the NOMAD curves developed through the Delphi process are not validated or cross-checked based on any market data. This leads to accuracy issues in this step of the model.	Alternative NOMAD approaches do exist and there may be opportunities to streamline the net standards savings component of the model.	Lack of transparency exists with how the NOMAD curves are derived. The NOMAD curves provided by different experts have very significant differences. The selection of the Delphi panel and how the panel member quantifies what are the naturally occurring market influences in not transparent.
	NET PROGRAM SAVINGS	
The entire spectrum of code advocacy activities are not captured, and the three existing attribution factors do not adequately reflect types of effort needed for standard adoption. This leads to accuracy issues in the model.	Opportunities to streamline exist. The attribution evaluation method and process can be improved by connecting the dots between program logic models and attribution factors. There is a need to reduce the heavy burden on program documentation.	There is a lack of transparency into how the impact evaluators consider program advocacy activities in determination of attribution factor scores and how the panel of experts are chosen and who is chosen.

Figure 2: Key Issues

1.3 Recommendations

> Our recommendations are to the CPUC codes and standards evaluation team.

The project team identified many codes and standards evaluation issues based on interviews with a broad range of stakeholders. These issues reveal the complexity of the codes and standards impact and suggest that existing codes and standards evaluation methods are not adequate to accurately capture energy savings generated by IOUs' C&S program. Resolving these issues may seem to lead to a more complex codes and standards attribution model and require more data to be collected for the impact analysis. The study revealed that these identified issues are interconnected to each other and need to be addressed in a coherent manner.

The CPUC codes and standards attribution model provides the framework of codes and standards program impact evaluation; and detailed evaluation methods for each model component determine how the codes and standards attribution model is implemented, including:

- Selecting the baseline used to calculate potential savings,
- Defining compliance and performing compliance adjustment to calculate gross savings,
- Defining and assessing NOMAD to calculate net standard savings, and
- Assessing C&S program activities to determine program attribution score.

The following recommendations are to the CPUC codes and standards evaluation team. In addressing individual evaluation issues, the project team developed two overarching themes of improvement for model components:

- Using energy-based impact assessment, and
- Collecting informative market data.

These themes allow different evaluation issues to be addressed coherently and, therefore, improve evaluation accuracy and streamline the evaluation process. We utilized the symbols below to illustrate how each recommendation worked towards addressing the issues of accuracy, transparency and streamlining.

A **Improve accuracy[1]:** Improved confidence in correctness of estimated savings

S **Streamlined model or process:** Identify opportunities to make the evaluation process simpler and more effective using less resources

T **Improve transparency:** Ability to see and comprehend how the estimated savings were determined

While the recommendations tended towards creating a more accurate model, there are areas of where streamlining and transparency also appear. It is important to note that there are trade-offs when taking a more accurate approach, but aligning around a more accurate and correct market approach we believe will ultimately result in a more streamlined model and reduce confusion of stakeholders.

The following sections present recommendations for improving the codes and standards attribution model, overarching evaluation methods, and individual model components. This study did not directly address various savings double counting concerns related to building and appliance standards due to limitations of project scope. However, the recommended codes and standards attribution model enhancement and overarching

[1] Accuracy is "the condition or quality of being true, correct, or exact; freedom from error or defect; precision or exactness; correctness."

evaluation methods provide a good framework to facilitate further development of solutions to resolve double counting issues.

1.1 Codes and Standards Attribution Model

The TRC team recommends that the CPUC codes and standards impact evaluation team recognize the importance of assessing overall market changes. The codes and standards advocacy program is a resource program aiming to reduce energy demand through mandatory[1] changes of the whole market, instead of selected population of the market, e.g. participants of incentive programs. Energy savings attributable to the codes and standards program should be based on actual energy use reduction of the whole market. Past CPUC codes and standards impact evaluations focused on verifying energy savings of standards, not market changes. In theory, energy savings are the same as the corresponding energy use reduction of the market. In practice, considering energy impact from the perspective of market changes provides a better understanding of actual energy savings achieved by standards in a complicated market.

Following this approach of considering overall market effects of standards, the TRC team recommends the following definitions of the codes and standards attribution model components:

- **Potential savings** are the standard's technical potential assuming idealized whole market change from baseline to the new standard;
- **Gross savings** represent actual market change, with consideration of how the average market condition deviated from the baseline before the new standard took effect and the how the average market condition deviates from the new standard after it took effect;
- **Net standard savings** correspond to the market change purely due to the standard, with exclusion of market changes introduced by all non-regulatory forces including those associated with voluntary standards and other utility efforts, including incentive programs; and
- **Net program savings** reflect the market changes attributable to utilities.

1.2 Overarching Evaluation Method

The TRC team recommends two overarching themes to be integrated into the codes and standards attribution model and analysis of evaluation issues related to each model components.

1.2.1 Perform Energy-based Assessment A S T

The assessment of market change, as emphasized in codes and standards attribution model improvement, should be based on energy performance of the overall market, instead of market share of qualified installation, i.e. those meeting or exceeding the standard.

> The TRC team recommends the assessment of market change should be based on energy performance of the overall market.

The latter cannot capture energy impact variations associated with different efficiency levels. Past codes and standards evaluation studies used energy-based evaluation for building compliance evaluation. Analysis in Chapter 5 shows that compliance evaluation based on compliance rate, which is market share of qualified installation, under-estimates compliance in terms of energy impact. Analysis in Chapter 6 indicates the NOMAD assessment has been based on market share of qualified installation and is incompatible with energy-based compliance assessment used for

[1] Note in this context mandatory is used in comparison to voluntary – such as incentive programs – and should not be confused with mandatory requirements in the building energy standards as compared to prescriptive or performance requirements.

building standards. NOMAD based on market share of qualified installation cannot adequately reflect the market trend of energy use in absence of the new standard. Therefore, it is necessary that energy-based evaluation be consistently used for all standards and for all model components.

The TRC team further recommends that energy-based evaluation be performed by assessing market average unit energy consumption (UEC) of related end-uses for installations covered by the standard. For example, evaluation of a Title 24 standard addressing lighting power density (LPD) and controls in warehouse buildings would assess the reduction of average indoor lighting UEC in warehouse buildings; evaluation of Title 24 whole building improvement for nonresidential new construction building would assess the reduction of average whole building UEC among all nonresidential new construction buildings.

UEC has been used to determine unit energy savings (UES), which is the UEC difference between the baseline and the new standard. Following the above recommendation, compliance evaluation would assess actual market average UEC and compare it to the UEC corresponding to the new standard. This compliance assessment method is the same as compliance margin approach described in Chapter 5, which indicates that compliance margin is a better indicator of compliance status than energy savings adjustment factor (ESAF) or compliance adjustment factor (CAF). Following the above recommendation, NOMAD evaluation would estimate the market average UEC trend in absence of the new standard. The difference between the baseline UEC and NOMAD UEC is the energy savings that would happen without the new standard and should be excluded from gross savings as NOMAD adjustment.

1.2.2 Collect Informative Market Data A T

The TRC team recommends that the CPUC codes and standards evaluation collect informative market data to support energy-based evaluation discussed in the prior section.

> The TRC team recommends the CPUC codes and standards evaluation collect informative market data to support energy-based evaluation.

Past CPUC codes and standards evaluations spent significant resources on data collection and analysis, not all of which provide informative data on energy use characteristics of the market to support energy-based evaluation. However, past codes and standards evaluations only used the collected data to estimate energy savings but did not provide UEC information. For appliance compliance evaluation, past codes and standards evaluations did not collect enough data to assess energy performance distribution among covered annual shipment to determine market average energy performance. NOMAD evaluation conducted by past codes and standards evaluations were based on input from industry experts, not market data.

To implement the above recommendation, the building compliance evaluators would continue to collect building performance data and use and collect the data to develop average UEC, which would then be used to determine average compliance margin to inform market compliance status. Compliance evaluation for appliance standards needs to collect appliance shipment data for different performance levels. For NOMAD assessment, the codes and standards evaluation team needs to collect market trend data to estimate efficiency improvement trends to inform how market average UEC for related end-use areas has been reduced overtime. Market trend data needs to be collected from past studies, e.g. California Commercial End-Use Survey (CEUS), Residential Appliance Saturation Study (RASS), and California Commercial Saturation Survey. Data collected at the national level and for other states could be used as supplemental information.

The CPUC evaluation team must anticipate that data gaps will exist when developing the completed market trend. However, the TRC team believes it is better to estimate NOMAD based on limited market data than not based on any market data. If future codes and standards evaluations collect market data to assess market average UEC on a periodic basis, there would be adequate data over time, along with data provided by other market studies, to provide more reliable market trend information.

1.3 Components of Codes and Standards Attribution Model

This section provides recommendations for each of the attribution model components. These recommendations address detailed evaluation issues following the framework of the enhanced codes and standards attribution model and overarching evaluation methods.

1.3.1 Potential Savings A S

Estimating baseline is important for determining potential savings. Current CPUC policy requires different approaches to determining baseline based on whether a prior standard exists. When a prior standard does not exist, an industry standard practice (ISP) baseline is required. There are complexities in selecting appropriate baseline options that carry through the rest of the attribution model.

The TRC team has the following recommendations:

- **When there is not a prior standard and ISP is used as the baseline per CPUC policy, ISP should be defined as market average UEC.** In addition, in net standard savings calculation, only incremental NOMAD, which is the additional NOMAD beyond that included in the market average, should be used for NOMAD adjustment.
- For Title 24 building standards with multiple sub-measures applied to multiple building spaces, **UES should be defined as the average UEC reduction for related building types and annual installation should be defined as the construction floor area of related building types**.

1.3.2 Gross Savings A S T

Past codes and standards evaluation studies spent significant resources on compliance assessment, yet still ultimately had small sample sizes used to determine statewide compliance. This approach does not provide robust compliance results to meet the C&S Program Attribution evaluation objectives, and therefore other alternative should be considered. Additionally, over-compliance must be understood and addressed correctly. Code compliance of individual buildings can exceed minimum code requirements due to normal market conditions or other parameters that are not a direct result of the code requirements.

Past approaches have not provided robust compliance results to meet the C&S Program Attribution evaluation objectives, and therefore other alternatives should be considered.

The TRC team has the following recommendations:

- **Use compliance margin as the metric to determine compliance status of Title 24 standards** and compliance adjustment can be performed using the method provided in Section 5.2.
- **Incorporate compliance adjustment for compliance of prior standard, when it exists and used as baseline.** This recommendation ensures that there is no savings overlap and gap between two consecutive standards.
- **Include over-compliance in calculating average compliance, i.e. not cap measured compliance to exclude over-compliance.** Some of the over-compliance might be associated with NOMAD and the corresponding savings would be removed through NOMAD adjustment. Also, by incorporating compliance adjustment for prior standard, savings associated with over-compliance of prior standard, would be excluded from savings of the new standard.

1.3.3 Net Standards Savings

Market conditions are intertwined with baseline and compliance affecting potential and gross savings as described above. Quantification of naturally occurring market influences is very challenging. Further, there are limited data sources for providing actual NOMAD information and the current evaluation approach lacks transparency on how the final NOMAD curves are derived. The current approach looks at each measure in isolation, the market is too complex and ambiguous for this to be a useful approach. The resulting NOMAD curves are difficult to test for accuracy since the process and results are not replicable.

NOMAD should be reduction of energy, not adoption of a particular product or technology.

The TRC team has identified opportunities for improving the net standards savings evaluation approach by developing long-term energy-based NOMAD trends by technology areas to better reflect codes and standards impact over time. NOMAD should be reduction of energy, not adoption of a particular product or technology. Energy-based NOMAD indicates how UEC reduction changes over time for a technology or end-use area.

The TRC team has the following recommendations based on Chapter 6 analysis:

- **Perform NOMAD based on market trend of energy use,** i.e. UEC reduction in absence of the new standard, not based on adoption rate of installations meeting or exceeding the efficiency level determined by the standard.
- **Estimate long-term market trend of UEC** for specific end-use areas and whole building without any standard improvement as NOMAD for related standards.
- **Do not exclude market effects caused by utility incentive programs in NOMAD assessment.** These market efforts should be attributed to IOUs during net program savings calculation. Without excluding these market effects, net standard savings purely reflect energy use reduction due to the adoption of the new standard beyond the business-as-usual condition.
- **Collect market data to develop efficiency improvement trends to inform NOMAD limits.** Evaluators should utilize existing resources to collect more market data to develop efficiency improvement trends to inform NOMAD curves. Evaluators could base market adoption trends on a review of previous evaluation and market studies, planned market and potential studies as well as detailed energy efficiency program data available through the CPUC.
- **Create a more transparent process in the development of the Delphi panel.** The robustness of the final NOMAD results are dependent on the strength and credibility of the Delphi panel. Recruitment of appropriate panelists is critical to a successful outcome. All panel members bring inherent biases and opinions that need to be made as explicit as possible to assist with the interpretation of results. The CPUC Codes and Standards evaluation team needs to provide transparency throughout the process in the development of the Delphi panel and the evaluation plan needs to include updates to the IOU Statewide codes and standards team on the composition of the final Delphi panel(s) for review and input.

1.3.4 Net Program Savings

The evaluators approach to assessing program advocacy activities in determining attribution factor scores lacks transparency and accuracy. Within this step in the model, the critical component is capturing the spectrum of IOU advocacy activities into the current attribution factors. The attribution factors may not reflect the full range and types of activities needed to achieve adoption and therefore may be limiting. In prior evaluations, no guidelines were provided on how effectiveness of program activities were determined and how ultimately the factors were scored and weighted.

Critical component is capturing the spectrum of IOU advocacy activities into the current attribution factors.

The TRC team proposes a set of possible resolutions based on the analysis in Chapter 7:

- **Use the codes and standards program logic model to establish a list of program activity areas and determine attribution score associated with each activity area.** Program attribution evaluation needs to conduct detailed analysis of codes and standards program activities and this recommendation simply provides a better organized evaluation process without requiring additional data collection or analysis (see Appendix 9.5). This recommendation enables the attribution evaluation to be transparent and providing effective feedbacks to codes and standards program staff.
- **Use additional attribution factors or develop clear evaluation methods** using the existing attribution factors to determine attribution for the following two important program activity areas:
  1. Strategic preparation, including short-term and long-term standard development planning, coordination, data collection, and solution development, and
  2. Stakeholder engagement and advocacy.
- **Recruitment of panelists for program attribution evaluation should be conducted through a transparent process** including using selection criteria vetted by the CEC, IOUs, and other key stakeholders involved in standard development and adoption.

2. Study Overview & Methodology

The California investor-owned utilities (IOUs) have developed the C&S advocacy programs as an integral part of the portfolio of statewide efficiency programs. Codes and standards are a significant means to establish and ensure market transformation. Codes and standards advocacy comprises a portfolio level strategy that complements incentive and information offerings in several ways. The C&S program serves as a link from research and emerging technology programs, to incentive and information programs, to code readiness to adoption of specific requirements into codes and standards. Since IOU incentive and rebate programs typically capture only a small percentage of the market, transition to regulatory intervention is essential to maximize portfolio energy savings. Through the Advocacy subprograms, the IOUs strive to strengthen and expand building and appliance codes and standards as market experience reveals greater efficiency opportunities.

The California Public Utilities Commission, Energy Division (CPUC-ED) impact evaluations (previously conducted by Cadmus and DNV GL) verified the estimated savings of the statewide Codes and Standards Building Codes Advocacy and Appliance Standards Advocacy subprograms. The purpose of the impact evaluations was to validate or correct the IOUs estimates of energy savings attributable to their advocacy of codes and standards. Evaluation activities and resources were focused on the codes and standards expected to produce the largest energy savings and/or with the greatest relative uncertainty associated with specific parameters.

Following the California Evaluation Protocols, the California Public Utility Commission (CPUC) evaluators created a codes and standards program attribution framework and model. As summarized in the five-step modeling illustrated in Figure 6 below, the evaluation team estimated potential savings that would result if all construction projects met the Title 24 code. Next, the team adjusted for the observed energy savings based on primary research to determine gross savings. The team then estimated net savings by adjusting for naturally occurring market adoption of energy-efficient units. The team determined programs net savings using an attribution adjustment to account for the Program's effect on standards adoption. The final evaluation step allocated net program savings to each IOU.

2.1 Study Overview

The purpose of this study is to conduct a high-level process evaluation of the overall IOU codes and standards advocacy efforts and revisit the attribution methodology to provide recommendations for improving the methodology and the evaluation approach. The goal of the research is to develop recommendations for updates to the model that best reflects C&S program contributions. A key outcome of this work is to identify if and how the model and evaluation process can be improved or simplified for timeliness and effectiveness, while also properly accounting for realized C&S program savings.

The four study objectives constitute two related, yet distinct parts;

- High-level process evaluation of the overall codes and standards advocacy efforts (Objective I)
- Review of the five-step attribution model (Objectives II-IV).

Objective I. C&S Advocacy Efforts High Level Process Evaluation: Provide a solid understanding of codes and standards program activities. Through interviews with a broad range of stakeholders involved in the process, TRC developed a deeper understanding of successes and barriers to the IOUs activity, effective roles and activity in the codes and standards advocacy process and appropriate attribution of those activities. The C&S program covers Title 24, Title 20, and Federal Appliance Standards. The output of this objective identified areas of codes and standards program activities that need to be further reflected in the Attribution Model or by the evaluation process.

Objective II. Assess Current Status of the Attribution Model: Performed in-depth examination of the existing evaluation method/model. We prioritized the Attribution Model factors and reviewed all the factors. The output of this objective supported the analysis of the next two objectives.

Objective III. Identify Data Overlap/Gaps: Examined the data input used for the Impact Evaluation and Attribution Model to determine what additional data should be collected to improve the model and evaluation. The output of this objective supported the model modification analysis objective.

Objective IV. Model Refinement & Improvement: Developed recommendations for model and evaluation refinement to improve the C&S Program savings estimates in support of system planning.

TRC developed the following Problem Statement to guide our investigation and analysis:

"How can the CPUC codes and standards impact evaluation process increase transparency, streamline, and simplify existing methodologies, and improve accuracy of the final net savings allocation to each utility as it relates to their support in the advancement of codes and standards in California?"

Figure 3 provides the research topics and questions which this study aims to address.

Topic	Research Questions
T24 code adoption process and IOU involvement	How important is IOU involvement to the identification of new codes and standards opportunities? ♦ What has worked well? ♦ What has not worked well? What aspects of the code adoption process demand more attention from the IOUs? What are the implications of zero net energy (ZNE) code in the codes and standards development process in CA? Who uses the results of the IOU work (Including energy saving calculations)?
Attribution model	What are the areas of strength of the attribution model? What are known issues of the attribution model? How accurate are Delphi panels compared to alternative approaches? How should model improvements be prioritized? Are evaluation time and resources being used appropriately to address the issues? Where is there potential for bias and what is the result of this bias?
NOMAD alternatives	Are there alternatives to estimating the naturally occurring market adoption (NOMAD)? What are the advantages and disadvantages of alternative approaches? Does the NOMAD properly account for T24, T20, and federal standards advocacy collective benefits? Does the NOMAD adjustment address California and federal standards in a consistent fashion? ♦ Should it? Could on-going market data collection efforts support an alternative approach?
Data sources	What data sources are used? Are there any other data sources that should be considered for use or be developed? Are there overlaps in data collection efforts with other evaluations that should be coordinated? Is there existing codes and standards program data, e.g. CASE study data, or incentive program data, that can be better utilized to inform the model and improve the evaluation? Are there long-term data collection efforts that could/should be used?
Logic model comparison	How does the CPUC impact evaluation process reflect and follow the program logic model?
Right-size model and evaluation	Can and should the model be simplified? Can and should the model be expanded? What is the right-size for the evaluation effort (cost vs. rigor vs. confidence)? How do the results of a whole building (top down), measure-level (bottom up), and hybrid approach compare in terms of accuracy and cost?
Improving accuracy	How should accuracy be defined? Is the current model producing results that can be readily used for system planning purposes? If not, why not? Would any proposed alterations to the current model yield results that would be more accurate for system planning purposes

Figure 3: Research Topics and Questions

2.2 Study Methodology

The first three objectives involve analysis of the existing evaluation process, data input, and results, as well as related analysis to identify any shortcomings in the existing evaluation results. Objective 4 recommendations are based on data collected and analyzed for the other objectives. Our primary data collection approach included a literature review and stakeholder interviews.

2.2.1 Literature Review

The first step in our data collection was to review literature associated with the Statewide Investor Owned Utility (IOU) C&S Program advocacy and Attribution Model. The literature review provides information on the history and evolution of the Attribution Model and explores alternative approaches to model components and processes. The findings help inform and direct further investigation through the stakeholder interviews. As part of the literature review, we documented changes that have been made to the Attribution Model to improve transparency, accuracy, and streamlining. Additionally, we researched the Bass diffusion model in the context of building and appliance energy measures and innovations, as well as investigated alternative approaches to the NOMAD analysis methodology. The Literature Review documentation is provided in Section 9.3.

2.2.2 Interviews

Our primary data collection approach also included interviews with a range of key players and experts in the energy efficiency standards development process. TRC interviewed 30 key stakeholders involved in the codes and standards process to gain a holistic view of the IOU's involvement in the codes and standards advocacy process, as shown in Figure 4.

Type of Interview	Count	Description of Role
IOU Codes and Standards Program Staff	10	Individuals working directly in the codes and standards advocacy programs within the IOUs.
IOU Codes and Standards Program Support	2	Consultants, CASE Authors, etc. working in support of the codes and standards advocacy programs for the IOUs.
IOU Efficiency Program Staff	6	IOU staff that are working directly in the incentive programs for energy efficiency.
CEC Staff	5	California Energy Commission staff working on codes and standards.
CEC Support	2	California Energy Commission consultants.
T24 Stakeholders	2	Advocacy and industry groups that support T24.
T20/Fed Appliance Stakeholders	2	Advocacy and industry groups that support T20 and Fed Appliances.
CPUC Staff	1	CPUC staff directly engaged in the IOU codes and standards advocacy activities.

Figure 4: Interviewee List

A total of 26 high-level interviews were completed with individuals that were not familiar with the attribution model. TRC conducted eight in-depth interviews with individuals familiar with the attribution process and model to explore specific topics of the attribution model gathering insights from the different perspectives of these stakeholders. These interviews supported TRC in moving into the next step of analysis.

After the interviews were completed, TRC cleaned and organized all interviews into final transcript files. Those files were then uploaded into a qualitative coding software tool, NVIVO, to support the analysis of what each interviewee said in comparison to the objectives of the study. The interviews were coded identifying strengths, weaknesses, attribution model recommendations, opportunities and so forth, to easily categorize interviewee perspectives. This allowed the TRC team to investigate and develop themes out of analysis. The NVIVO analysis tool was then used for investigating these overarching themes via text queries, word clouds and word counts. The process is illustrated in Figure 5.

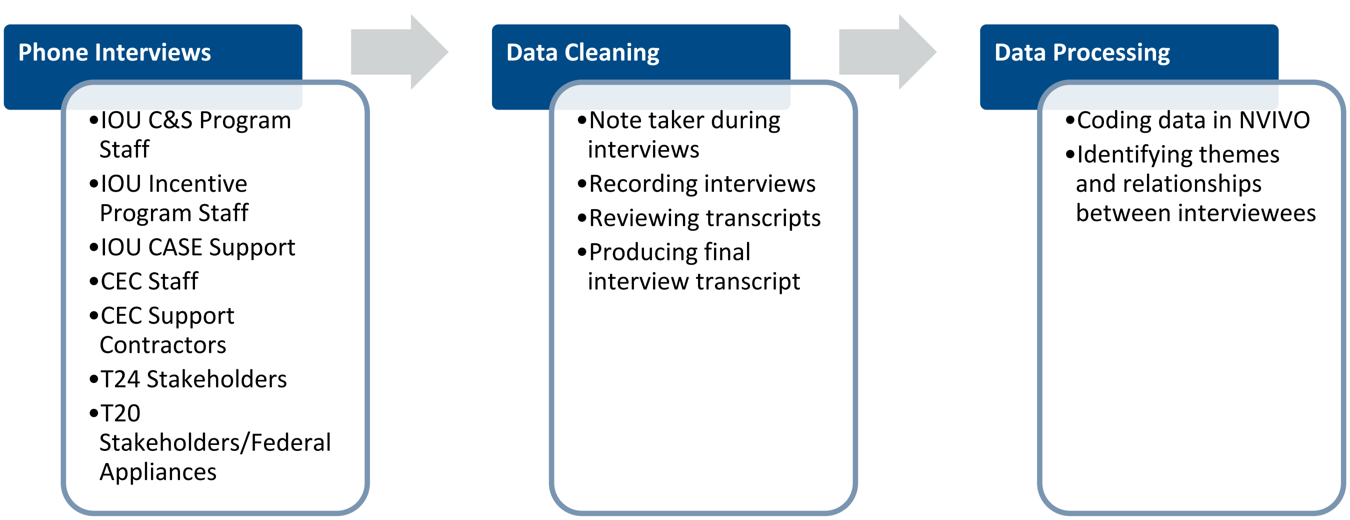


Figure 5: Qualitative Coding Process

After high-level coding was complete, each interview was coded by the attribution model component to understand specifically what the interviewees were saying about each model component. This information was further layered into the lens of transparency, accuracy and streamlining to reflect back on the initial problem statement of the research. Appendix 9.4 shares these insights from the interviews.

2.2.3 Analysis

TRC conducted an in-depth review of the evaluation attribution model framework for determining and evaluating attributable energy impacts of IOUs' efforts in supporting building code and appliance standards development. The team analyzed the data to develop a complete understanding and description of the existing evaluation approach and model. Analysis included both a bottom up (in-depth interviews and investigation into evaluation steps), and top down investigation (high-level overview of process and model), looking at the details of each step in the model and prior evaluation approach as well as considering the overall objectives and key outcomes of both.

The analysis approach and outcomes are:

- Perform in-depth examination of the existing model and evaluation method, with a focus on NOMAD and attribution.
- Investigate evaluation alternatives to the Delphi approach. The Delphi panel is an evaluation technique involving an interaction between the researcher and a group of identified experts on a specified topic. Delphi panels have been used to gain a consensus regarding future trends and projections using a systematic process of information gathering. Specifically, the codes and standards evaluation team identifies a Delphi panel to estimate how the market would have acted without codes and standards participation at play.
- Examine the data input used for evaluation. Determine if there is additional data that should be collected to improve evaluation, and if existing data can be better utilized to improve the evaluation.
- Provide a good understanding of codes and standards program activities and existing codes and standards evaluation model, identify areas of codes and standards program activities that need to be further reflected in attribution model or by evaluation process.
- Develop recommendation, especially address the question on providing proper codes and standards savings to support system planning.

3. Overview Codes and Standards Attribution Model

The code and standards attribution model provides a framework to evaluate the C&S Program process and determine how each IOU has contributed to advancing codes and standards in California. The model in Figure 6 below, illustrates the methodology devised by the CPUC impact evaluators during the 2006-2008 codes and standards evaluation (KEMA, Inc.; The Cadmus Group, Inc.; Itron, Inc.; Nexus Market Research, Inc.; ENRG, Inc., 2010). Since then, evaluators have conducted two more evaluation studies (2010-2012 and 2013-2015), evolving the framework minimally over time.

The codes and standards attribution framework was originally developed by the IOUs (Mahone, Hall, Megdal, Keating, & Ridge, 2005), in collaboration with the CEC and CPUC, to establish the codes and standards program as a resource program. In 2005, the CPUC adopted the framework proposed by the IOUs and impact evaluators further developed detailed evaluation methods according to changing codes and standards landscape. The attribution model provides estimated savings attributable to IOUs for their codes and standards development and advocacy efforts. As a resource program, the C&S program needs to provide reliable estimates of energy demand reduction to support the state's long energy resource planning.

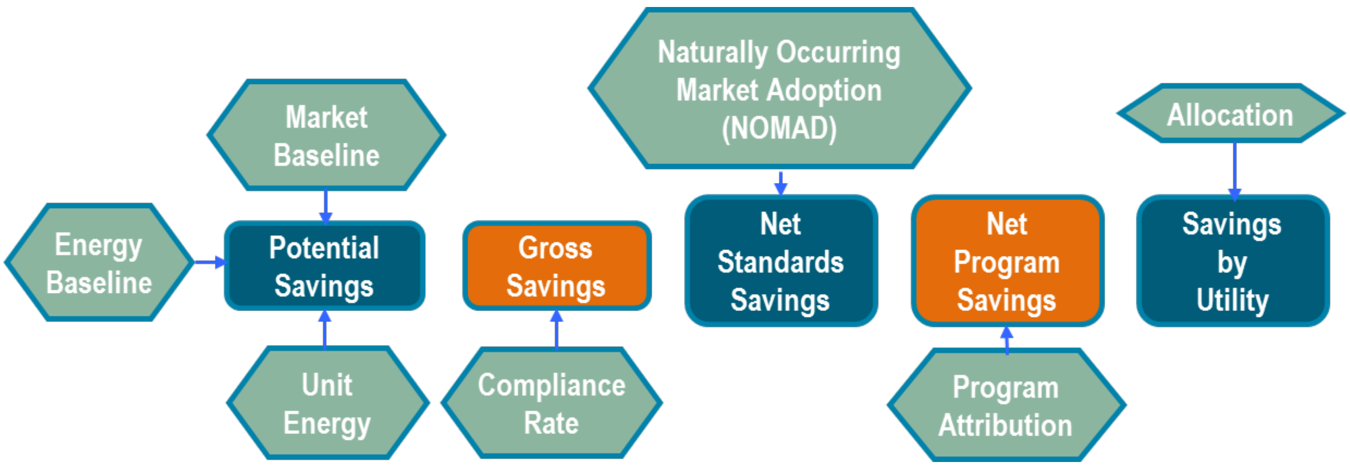


Figure 6: Codes and Standards Attribution Model

As shown in Figure 6, the model includes four major components of calculation to provide Potential Savings, Gross Savings, Net Standard Savings, and Net Program Savings, respectively. In general, these steps of the model are based on the same fundamentals of impact evaluations for incentive programs. However, because C&S Program affects the whole market instead of incentive program participants with relatively small populations, the codes and standards attribution model needs to accommodate more complicated market conditions. To address the detailed codes and standards evaluation issues, it is important to have a complete understanding of the purpose of each calculation step. We summarize each component below and provide in-depth descriptions in the following chapters.

3.1 Summary of Attribution Model Components

The code and standards attribution model provides a framework to evaluate the codes and standards advocacy process. The model includes five major components: Potential Savings, Gross Savings, Net Standard Savings, Net Program Savings and Allocated Savings. To assess the current status of the attribution model, the team reviewed key issues and concerns with the model and prior evaluation processes which have led, in part, to the need for this study. Following each summary, we provide issues that were identified early in the study, and while these issues may not have proven to be the highest priority in the final recommendations and are not exhaustive, they highlight the need for improvements. Below we provide a summary of each of the components. In following sections, we provide a more in-depth review of the first four components.

3.1.1 Potential Savings: Technical Potential

The first step of the model is determining potential savings. Potential savings are calculated as the product of annual unit of installation and unit energy savings (UES). UES are the difference in unit energy consumption (UEC) between the baseline and the new standard. According to this definition, potential savings reflect the energy use reduction from improving all annual installations covered by the standard from the baseline to the new standard level. In other words, potential savings represent the technical potential of the standard without considering actual market conditions, such as existing market adoption rate of the efficiency measure required by the new standard, actual market adoption rate of efficiency measure after the new standard becomes effective and expected market adoption rate of the efficiency measure if the new standard were not adopted. These market conditions are often very complicated.

Estimating baseline is important for determining potential savings. Current CPUC policy requires different approaches to determining baseline based on whether a prior standard exists. A prior baseline standard is straightforward. When a prior standard does not exist, an industry standard practice (ISP) baseline is required. Determination of an accurate baseline creates uncertainty and challenges since there is no clear definition of ISP or specific guidelines on how to determine it. Market conditions are important for ISP, but properly determining market conditions is often very complicated. There are complexities in selecting appropriate baseline options that carry through the rest of the attribution model.

As will be discussed in later sections, certain baseline options include embedded market conditions. The study will examine if they can be used as a baseline within the framework of codes and standards attribution model and, if so, what adjustment is needed.

3.1.2 Gross Savings: Actual Market Change

Gross savings reflect actual market improvement after the new standard becomes effective. Comparing to potential savings, the gross savings calculations needs to consider two actual market conditions: before the new standard becomes effective, how market conditions differ from the baseline, and, after the new standard becomes effective, how market conditions differ from the performance level set by the new standard. The latter is determined by the compliance of the new standard; the former is determined by the compliance of the prior standard, if it exists. If a prior standard does not exist, the baseline needs to reflect actual market condition before the new standard taking effect. Further analysis of this issues is provided in the next chapter. Gross savings are obtained by applying a compliance adjustment to the potential savings, which is to incorporate actual market conditions by using compliance rates. It is important that compliance adjustments are applied to both the new and prior standards in order to correctly capture the actual market improvement associated with the new standard.

Estimating compliance is important in assessing gross savings. Past codes and standards evaluation studies spent significant resources on compliance assessment, yet still ultimately had small sample sizes used to determine statewide compliance. For example, the evaluation approach used permit data for newly constructed projects permitted under the most recent code, which most likely biased the sample to smaller nonresidential buildings. The approach does not provide robust compliance results to meet the C&S Program attribution evaluation objectives, and therefore other alternative should be considered.

Additionally, over-compliance must be understood and addressed correctly. Code compliance of individual buildings can exceed minimum code requirements due to normal market conditions or other parameters that are not a direct result of the code requirements. Efficiency improvement occurs for various and multiple reasons depending on the product; not specifically to improve efficiency. Over-compliance at the building level should not be capped. Some over-compliance may represent normal market conditions.

3.1.3 Net Standards Savings: Net Impact of the Standard

Net standard savings represent the impact due to the new standard. They are obtained by excluding savings associated with naturally occurring market adoption (NOMAD). This calculation step is similar to net savings calculation for incentive programs, which removes free riders, i.e. program participants who would install the efficiency measure without incentives, from gross savings. NOMAD attempts to represent the technology adoption or building system design improvement in the market without the new standard.

Accurately estimating NOMAD is an important component. Market conditions are intertwined with baseline and compliance affecting the potential and gross savings calculations above; causing challenges in the quantifying naturally occurring market influences. Further, there are limited data sources for providing actual NOMAD information and the current evaluation approach lacks transparency on how the final NOMAD curves are derived. The approach currently looks at each measure in isolation, the "market" is too complex and ambiguous for this to be a useful approach. The resulting NOMAD curves are difficult to test for accuracy since the process and results are not replicable. Because the approach cannot be tested, and the resulting estimates cannot be systematically replicated it is difficult to determine if the process is sound and the final estimates are valid.

3.1.4 Net Program Savings: Attributable Savings to Codes and Standards Program

This step of calculation provides the savings attributable to the C&S program by applying an attribution score to the net standard savings. The attribution score, represented as a percentage, is determined by taking into consideration all the development and advocacy activities that the IOUs provided to achieve the adoption of a specific standard.

Within this step in the model, an important component is capturing the spectrum of IOU advocacy activities into the current attribution factors. The evaluators approach to assessing program advocacy activities in determining attribution factor scores lacks transparency. It is unclear how the attribution assessment might overlap with the net standards savings estimates. The attribution factors may not reflect the fully range and type of activity needed to achieve adoption and therefore may be limiting. In prior evaluations, no guidelines were provided on how effectiveness of program activities were determined.

3.1.5 Allocated Savings: Savings for Each IOU

Lastly, the final step allocates net program savings to each IOU. According to the latest CPUC decision (CPUC, 2018), net program savings should be allocated according by each IOU's C&S program budget spent on codes and standards advocacy. This step was not investigated in this research project.

3.2 Summary of Interviews

The interviews provided key insights into the overall advocacy process, strengths, and challenges of the process as well as opportunities for improving the attribution model. Interviewees stated continually that IOUs are critical to the codes and standards development and adoption process. Their role is a combination of stakeholder engagement and technical analysis. Stakeholder engagement is critical to the entire codes and standards advocacy process. T24 and T20 CASE reports are an integral and successful element to overall codes and standards advancement.

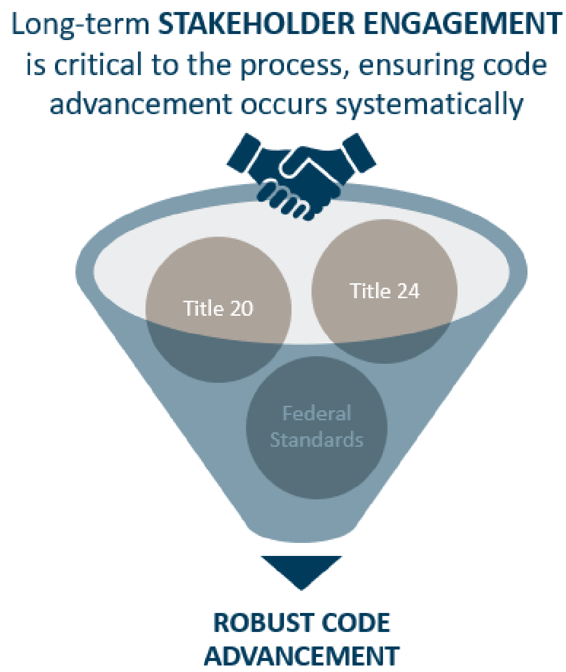


Figure 7: Stakeholder Engagement

3.2.1 Strengths

The interviewees stated that IOUs continue to improve the coordination effort and at each code cycle continuous improvement is being integrated into the process. CEC and CPUC staff and other advocates often stated that IOU codes and standards team deliverables are consistent and of high quality. Their continued role to act as technical experts in federal rulemaking and provide in-depth reviews sharing insights to the advocacy groups is important to the process. One example that was called out included that the integration and use of T24stakeholders.com has become an effective stakeholder management and coordination tool.

3.2.2 Challenges

While a number of strengths are apparent, there are challenges that the IOUs face in the advocacy process. Fundamentally, the success of the IOU advocacy efforts relies on action by the CEC or DOE. Based on interviewees' feedback, the CEC lacks funding, and funds may be spent unnecessarily making it challenging to go as deep or expansive as could be possible with the advocacy activities. For Title 20 updates in particular, CASE reports can sit on the shelf with the associated code changes never moving towards adoption. IOU success is often dependent on adoption of code measures based on CASE reports. If adopted, these do translate into savings. However, if the IOU CASE team pushes too far, the measure may get dropped based on stakeholder pushback. However, IOU codes and standards staff and consultant interviewees commented that it is difficult to predict the measures that will get a lot of stakeholder pushback, and therefore hard to judge what is considered too far. With the changing energy efficiency landscape, it is getting hard to change the code and show energy savings with efficient buildings. For example, federal preemption issues can be a roadblock for hitting zero net energy goals. Additionally, the state metrics are moving from energy savings to carbon reduction goals. Guidance from both the CPUC and CEC is needed for documenting greenhouse gas emissions reductions and associated energy savings.

Information and data limitations do exist. The IOUs could always use more cost information for cost effectiveness analysis. Cost-data inputs in CASE reports do not always align with the market truths; and there is a need to engage Title 20 and Title 24 stakeholders early on to verify the results. IOUs need incremental cost data, but only manufacturers know the incremental cost. If IOUs don't have information from the manufacturers, they have to take a conservative approach. One interviewee stated, that IOUs can do better at finding the right data, identifying a more accurate sample set for compliance studies, and associated project sites for compliance analysis. Having the accurate data to point is critical for pushing measures along. Title 20 and Title 24 advocates stated that certain trade associations are being left out of the engagement process. There is an opportunity if IOUs can encourage people to participate in the stakeholder meetings as soon as possible.

3.2.3 Informing the Evaluation and the Attribution Model

Through the interview process, the interviewees identified items that should be considered in investigating the overall attribution model as well as each component of the attribution model. Overall, interviewees felt that the attribution model should properly reflect the level of effort needed and expended on specific critical advocacy activities that are necessary to ensure code advancement.

The comprehensive IOU codes and standards team advocacy activities should be accounted for in the attribution model. The IOU codes and standards team is improving the process for developing code change savings reports to more holistically and accurately measure and reflect all IOU involvement.

Specifically, interviewees felt that the impact evaluators should consider:

- General activities to align all stakeholders on a process,
- The type of outreach activities conducted by the IOUs,
- Advocacy activity work that is beyond the CASE work,
- IOU codes and standards team strategic tactics to keep key elements of codes or standards in play, and
- IOU codes and standards teams use of counterproposals as a strategy to gain support and push forward specific codes or standards.

Relative to the attribution model itself, overall accuracy seems to be the main issue for interviewees as it relates to each component of the model. However, many stakeholders reflected on what the priorities are if it is accuracy, there could be increased cost and issues with streamlining.

4. Potential Savings

This section investigates the calculation method for potential savings. Potential savings are defined as the product of UES and annual unit of installation (or annual installation). UES is the difference between the baseline per unit energy consumption and maximum per unit energy consumption allowed by the new standard. Potential savings incorporates estimates of the energy baseline, UES (i.e. per measure savings), and the market baseline. To determine UES, the evaluation team collects field data and generates energy simulations for Title 24 measures and uses field data collection and rated energy performance for Title 20 appliance standards. Based on a CPUC decision, the energy and market baselines are set at the previous code or standard, if one exists, or ISP if one does not exist (Energy Efficiency Policy Manual R.09-11.014 Version 5, 2013, p. 6).

As the first step of the codes and standards attribution model, the potential savings calculation sets the scope of the savings calculation. The choice of baseline, unit, and corresponding UES determines how compliance and NOMAD assessment methods are completed. We investigated related issues.

4.1 Key Issues

Themes appeared during the key stakeholder interview regarding issues with accuracy, streamlining and transparency. Throughout this chapter we discuss key issues related to accuracy, streamlining and transparency for potential savings assessment.

ACCURACY	STREAMLINING	TRANSPARENCY
Lack of accuracy: • When choosing ISP as baseline. • When defining UES for standards involving multiple efficiency measures.	Opportunities to streamline: • Using CASE studies and other data sources to inform UES verification.	Lack of transparency: • When establishing ISP when there is no prior standard.

Based on interviewee input and our investigation, TRC identified the following issues related to potential savings assessment:

- What ISP should be used as baseline if there is no prior standard? Should it be the practice with the highest market share, the one reflecting the average market practice, or other options? Would model accuracy be affected by these baseline choices?
- How to define the unit for standards involving multiple efficiency measures? It is common for a Title 24 standard to include multiple efficiency measures applied to different building spaces, yet with some overlap among them. Additionally, there is often overlap among Title 24 requirements and T20 appliance standards.
- How can CASE studies and other data sources be effectively used to inform UES verification? These data sources provide useful input that can significantly reduce evaluation resources needed for UES verification. However, if savings provided by these data sources are not based on the same baseline as codes and standards attribution model, can they still be used?

4.2 Analysis

4.2.1 Baseline Choice

Current CPUC policy requires that the prior standard, if it exists, should be used as the baseline. If there is not a prior standard, ISP should be used as the baseline. This policy was originally designed for incentive programs. Savings assessment for incentive programs are mostly concerned about the impact on program participants, which represents a small fraction of the market. In contrast, codes and standards savings assessment is to assess energy impacts on the whole market by a standard. If the baseline does not represent the overall market conditions before the new standard takes effect, gross and net standards calculation methods need to adjust accordingly. This section analyzes related issues to develop methods to improve evaluation methods within the framework of CPUC's baseline policy.

Prior Standard as Baseline

When the prior standard is used as baseline, the embedded assumption in potential savings calculation is that the average efficiency of the market is at the same level as the prior standard, had the new standard not become effective. In other words, annual installations are in exact compliance with the prior standard, i.e. without partial compliance or over compliance, before the new standard takes effect. This assumption is made to obtain technical potential of the new standard and certainly does not reflect actual market conditions before the new standard takes effect. The next step in the Attribution Model, gross savings calculation, captures actual market conditions by adjusting potential savings based on standard compliance. As will be further explained in the next chapter, compliance of the prior standard needs to be incorporated in compliance adjustment in order to correctly reflect the actual market improvement after the new standard takes effect.

Using the prior standard as baseline also implies potential savings reflect market improvement with zero initial adoption of the new standard, or zero initial NOMAD. In reality, there is always some market adoption of the efficiency measures before they are adopted as the new standard. Potential savings are not accounting for any early market adopters in the calculation. In net standard savings calculation, energy savings associated with all NOMAD need to be excluded through a NOMAD adjustment to provide net impact of the standard.

In summary, using the prior standard as baseline is consistent with the methodology of existing codes and standards attribution model in that the resulting potential savings reflect a technical potential based on perfect standard compliance and zero NOMAD. To ensure model accuracy, compliance of both the prior and new standards needs to be incorporated in gross savings calculation and all NOMAD needs to be excluded in net standard savings calculation.

Industry Standard Practice as Baseline

When there is no prior standard, CPUC policy requires that ISP be used as a baseline. However, there is no clear definition of ISP or specific guidelines on how to determine it. The following three options of ISP are analyzed:

1. Typical practice: the practice or appliance model with the highest market share.
2. Market average: the market average performance in terms of UEC. Market average means the average weighted by annual installation.
3. Other ISP options

Typical Practice as ISP

Typical practice only represents part of the market. Other practices have less market share than the typical practice, but collectively, other practices can have a significant or even larger market share than the typical practices. Also, some practices may have smaller market share but generate larger energy impact than others.

The problem of using typical practice as baseline is that the corresponding UES does not reflect improvement potential of other practices. Applying this UES to all installations produces erroneous technical potential. It should be noted that incentive programs can use typical practice as baseline, because it only serves as a reference to determine savings from program participants and has no implications on the rest of the market.

Market Average as ISP

When market average is used as ISP to serve as baseline, the resulting UES properly represents the average improvement of the whole market from the prior condition to the new standard level. It should be noted that market average referred in this report represents the average of different energy performance levels weighted by their annual installation quantity.

When the baseline is based on market average, it properly reflects the overall market status in terms of energy consumption without the new standard. Therefore, gross savings calculations, which will be discussed in detail in the next chapter, only need to consider compliance of the new standard to capture the actual change of the market. This is different from when the prior standard is used as baseline. In that case, compliance of the prior standard also needs to be included in order to properly capture the overall market condition before the new standard takes effect.

Market average is based on energy performance of all installations, including those already meeting or exceeding the new standard. In other words, NOMAD is included in the baseline when the ISP is based on market average. Effectively, a NOMAD adjustment is performed, even though it is not supposed to happen at this stage of the codes and standards attribution model savings calculation. This is in conflict with the overall codes and standards attribution method, which assumes zero NOMAD in potential savings calculation and only considers (or excludes) NOMAD efforts in calculating net standard savings. This is a convoluted issue and needs to be investigated analytically.

The analysis below shows how energy savings associated with different market sectors are captured in potential savings calculations when market average is used as ISP to serve as baseline. Each market sector represents a subset of the annual installation with a particular energy performance level before the new standard takes effect. Market sector 1 through n represent those that did not meet the new standard before it takes effect. For simplicity, one market sector is assumed to meet the new standard before it takes effect and this market sector is denoted as “NOMAD” in the analysis below. UEC_i represents unit energy consumption, UES_i represents unit energy savings, and $Install_i$ represents annual installation for different market sectors.

$$Install_{ALl} = Install_1 + Install_2 + \cdots + Install_n + Install_{NOMAD}$$

$$Potential\ Savings = UES_{Avg} \cdot Install_{All} = \left(UEC_{Avg} - UEC_{Std}\right) \cdot Install_{All}$$

$$= \left(\frac{UEC_1 \cdot Install_1 + UEC_2 \cdot Install_2 + \cdots + UEC_n \cdot Install_n + UEC_{NOMAD} \cdot Install_{NOMAD}}{Install_{All}} - UEC_{Std}\right) \cdot Install_{All}$$

$$= UEC_1 \cdot Install_1 + UEC_2 \cdot Install_2 + \cdots + UEC_n \cdot Install_n + UEC_{NOMAD} \cdot Install_{NOMAD} - UEC_{Std} \cdot Install_{All}$$

$$= (UEC_1 - UEC_{Std}) \cdot Install_1 + (UEC_2 - UEC_{Std}) \cdot Install_2 + \cdots + (UEC_n - UEC_{Std}) \cdot Install_n$$

$$+ (UEC_{NOMAD} - UEC_{Std}) \cdot Install_{NOMAD}$$

$$= UES_1 \cdot Install_1 + UES_2 \cdot Install_2 + \cdots UES_n \cdot Install_n + \cancel{UES_{NOMAD} \cdot INS_{NOMAD}}$$

In the last step of the above analysis, the last term is crossed out (set to zero) because UES for “NOMAD” market sector (UES_{NOMAD}) is zero.

The above analysis shows that, using a market average as baseline, the resulting potential savings does not include potential savings associated with “NOMAD” market sector. Effectively, a NOMAD adjustment is performed during the potential savings calculation. This NOMAD adjustment excludes energy savings associated with the level of NOMAD right before the new standard take effects. In net savings calculation, these energy savings should not be excluded again. NOMAD increases over time; NOMAD adjustment for net standard savings calculation should only exclude savings associated with additional NOMAD after the new standard becomes effective.

Other ISP options

Any other ISP options that do not reflect market average cannot represent the overall market condition. If these ISP options are used as baseline, they will have a similar effect as using typical practice as baseline, i.e. not all performance levels would be properly represented. Therefore, in general, they are unlikely to be suitable as baseline for the existing codes and standards attribution model.

4.2.2 UES for Complicated Standard Requirements

For Title 24 nonresidential building standards, unit of installation is usually defined as the unit floor area, or square feet (sqft). Many building standards have multiple sub-measures applicable to building spaces, which may or may not have some overlap with each other. For example, the 2013 Title 24 Nonresidential Lighting Alteration standard includes requirements on lighting power density (LPD) reduction and several lighting controls. These requirements depend on building space type. Some building space types are only subject to LPD reduction requirement, some building spaces types are only subject to certain lighting control requirements, and some building space types are subject to both LPD reduction and lighting control requirements. The situation is illustrated in Figure 8. For these standards, it is hard to specify the unit of installation to represent the overall impact.

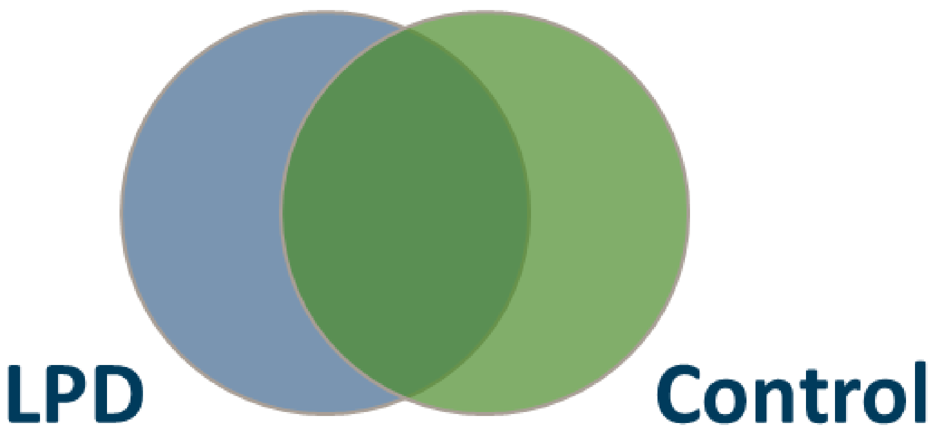


Figure 8: Overlap of Lighting Alterations LPD and Controls Requirements

For the example above, the 2013-2015 CPUC codes and standards evaluation estimated the annual installation and UES for each sub-measure, summed up annual installation of sub-measures to be the annual installation of the standard, and calculated UES for the standard as the average UES of all sub-measures weighted by annual installation. This approach is correct in mathematical sense. However, besides providing a misleadingly large annual installation which is much larger than actual annual lighting alteration floor area, the resulting unit of installation and UES have no practical meaning. This makes the compliance and NOMAD assessments hard to perform. It is important that units of installation are clearly defined to allow compliance and NOMAD adjustments be properly performed.

4.2.3 UES Verification

The previous CPUC impact evaluation studies used multiple data sources to verify UES including:

- CASE reports;
- DEER and publicly available IOU workpapers;
- Industry statistics published by product-manufacturing trade organizations;
- DOE Technical Support Documents for Building Codes; and

- Other CPUC evaluation activities or obtained specifically to support this evaluation project.

However, these data sources do not necessarily use the same baseline and UES definition. Previous codes and standards evaluation studies did not address these possible differences and their effects on evaluation results when using this data sources.

4.3 Possible Improvements

4.3.1 Baseline

Using a prior standard as baseline is consistent with CPUC policy and consistent with the existing codes and standards attribution model. However, compliance of a prior standard needs to be incorporated in the gross savings calculation to accurately capture the actual changes introduced by the new standard.

When there is no prior standard, ISP is to be used as baseline according to CPUC policy. The ISP should be based on market average in terms of UEC in order to correctly reflect potential savings of the whole market. Using market average as baseline implies initial NOMAD, right before the new standard becomes effective, is included in the potential savings and the resulting gross savings. In net savings calculation, this initial NOMAD should be removed from NOMAD in future years when the new standard is effective. It should also be noted that market average represents an energy performance level, rather than a particular building design practice or appliance model.

4.3.2 UES for Complicated Standard Requirements

One possible approach for calculating UES for complicated standard requirements is to consider the impact of these requirements on the corresponding end-use area of all related building spaces. For example, for the nonresidential lighting alteration standard discussed in section 4.2.2, UES of the standard will be defined as lighting energy use reduction for all annual lighting alteration floor area. Each sub-measure contributes to the total UES by affecting a subset of the total annual lighting alteration floor area. In this way, the unit of installation reflects the whole market sector (e.g. all office buildings, all retail buildings, or all nonresidential buildings) affected by the standard and UES reflects average per unit energy use reduction for an end use area (e.g. interior lighting or HVAC) of the whole market sector. Compliance and NOMAD would be assessed from this perspective achieving the UES reduction for the whole market sector, without the need for addressing compliance and NOMAD of specific sub-measures. Both compliance and NOMAD assessment can be greatly streamlined with this approach. Evaluation results can be more useful to provide a better understanding of overall impact of the standard.

4.3.3 UES Verification

Future CASE studies may consider following the same baseline rules as the codes and standards evaluation in order to make UES estimates easily used by future codes and standards evaluation studies. The CPUC evaluation should consider directly using UES estimates provided by CASE studies in order to streamline codes and standards evaluation and save resources for activities that can provide additional useful data. If other data sources are used to validate UES, CPUC evaluators need to be aware of discrepancies caused by differences in baseline.

5. Gross Savings

In this chapter, we describe and investigate the gross savings evaluation methodology. In particular, we examine approaches used to define compliance, assess compliance, and perform compliance adjustment.

5.1 Key Issues

Throughout this chapter we discuss key issues related to accuracy, streamlining and transparency for gross savings assessment.

ACCURACY	STREAMLINING	TRANSPARENCY
Lack of accuracy: • The compliance rate based on binary approach under-estimates compliance in terms of energy impact. • Does not capture actual market changes that happen after a new standard takes effect.	Opportunities to streamline: • Improvements to the compliance assessment process should be developed to reduce cost and time it takes.	Lack of transparency: • In understanding how over-compliance is treated in the impact evaluations.

As explained in the discussion on overall codes and standards attribution model, gross savings need to reflect actual energy use reduction of annual installations after the new standard takes effect. While potential savings represent marked improvement based on perfect compliance of the new and prior standards (if it exists), gross savings reflect actual market changes based on actual compliance of the prior standard before the new standard becomes effective and actual compliance of the new standard after the new standard becomes effective. The codes and standards attribution model shows that gross savings are obtained by applying the compliance adjustment to potential savings. Past codes and standards impact evaluation studies conducted compliance evaluation and performed compliance adjustment accordingly. Several questions have been brought up by stakeholders regarding gross savings evaluation:

1. Do existing definitions of compliance properly reflect compliance status in terms of energy impact? There are multiple ways to define compliance and past codes and standards evaluations used different compliance definitions. The 2010-12 codes and standards evaluation provided very high compliance values in terms of compliance adjustment factor (CAF), which has caused some confusion among stakeholders. We will investigate existing definitions of compliance.
2. Does the existing compliance adjustment method correctly capture the actual market changes that happen after the new standard takes effect? Past codes and standards evaluation studies only considered compliance of the standard. This approach implicitly assumes compliance rate of a prior standard, if it exists, is 100%, which is not consistent with findings of those evaluation studies themselves. Therefore, existing compliance adjustment methods do not fully capture market change after the new standard is effective. We will explore compliance adjustment methods to properly capture market changes.
3. How should savings associated with over-compliance be treated in gross savings calculation? Past codes and standards evaluations used different methods to treat over-compliance. The last CPUC codes and standards evaluation study, 2013-15 codes and standards evaluation study, limited compliance to 100% for building samples having measured compliance rate over 100% without providing detailed reasons. We will investigate possible issues related to over-compliance.

4. How to improve compliance evaluation process? Past codes and standards evaluation studies spent significant resources on compliance assessment, especially for Title 24 compliance assessment. Yet these studies obtained compliance results from very limited number of samples, which cannot adequately represent overall market compliance status. We will examine existing compliance evaluation methods and explore possible improvements.

The next section provides analysis of these four issues and the following section discusses potential improvement options regarding these four issues.

5.2 Analysis

5.2.1 Definition of Compliance

The concept of compliance is simple; it indicates if standard requirements are met. However, there are multiple ways to interpret the exact means of compliance, including:

- Required efficiency measures are installed,
- Required efficiency measures are installed and work properly per standard requirements,
- The appliance or building construction meet the performance requirements set by the standard, and
- Compliance procedures are properly and thoroughly followed, even though these compliance procedures do not directly generate any energy savings.

For codes and standards impact evaluations, compliance is used to assess market status in terms of meeting energy performances set by appliance and building standards, instead of meeting compliance procedures. Ideally, compliance for individual installation needs to indicate if its annual energy consumption is lower or higher than what is allowed by the standard and by how much. Correspondingly, compliance of the overall market needs to indicate if annual appliance shipments or building construction use less or more energy than the total energy consumption according to the standard. It is challenging to evaluate compliance in terms of energy consumption. The definitions of compliance and method of compliance assessment have evolved from the CPUC 2006-08 codes and standards evaluation to 2013-15 codes and standards evaluation. There are still two types of compliance definition being used: compliance rate and CAF or energy savings adjustment factor (ESAF). We will analyze these two types of compliance definition. In addition, we will also discuss another definition of compliance: compliance margin (CM), which is commonly used by building efficiency incentive programs but less often used by the C&S Program. This helps to provide a comprehensive view of compliance from different perspectives and understanding of linkages among these definitions, so codes and standards evaluation results can be better understood and utilized by the broad energy efficiency community.

Compliance Rate

Compliance rate is defined as percentage of annual appliance sales or building construction that meet the new standard. For individual samples of appliances and building construction projects, compliance is measured using a binary approach, i.e. the sample is considered to either meet or not meet the new standard. Partial compliance is considered as non-compliance and over-compliance is not differentiated from those exactly meet the new standard. Compliance rate of the overall annual installation is calculated as the ratio of samples meeting the new standard to the total sample quantity.

This definition of compliance is easy to understand and matches with the general understanding of compliance rate. However, this compliance definition cannot capture different levels of compliance in terms of annual energy consumption as compared to standard requirements. Market variation in compliance can be separated into three categories, as shown in Figure 9: below the baseline, meet or exceed the baseline but below the new standard, meet and exceeding the new standard. With the binary compliance measurement method, those with efficiency between the baseline and the new standard are considered to perform at the baseline level and those with efficiency exceeding the new standard are considered to perform at the new standard level. Therefore, energy efficiency for these two groups of installations is under-estimated. On the other hand, installations with efficiency below the baseline are considered to perform at the baseline line (no negative savings are considered in compliance rate approach). However, for most standards, even with the existence of non-compliance, market share of below-baseline installation is small. The combined effect is that overall performance of the annual installation is under-estimated. In other words, the compliance rate based on binary approach under-estimates compliance in terms of energy impact.

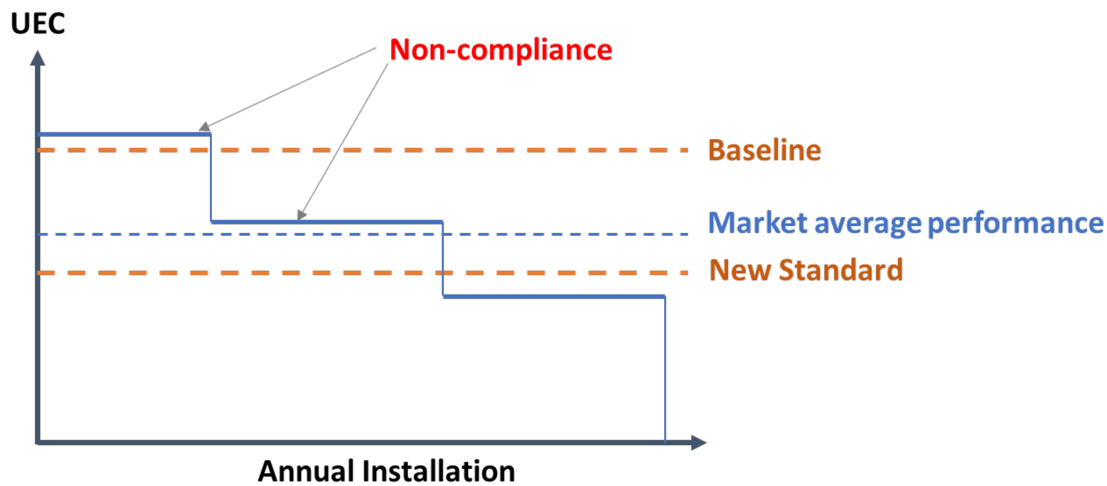


Figure 9: Compliance Variation

This definition of compliance was applied to both appliance and building standards in the 2006-08 codes and standards evaluation study and continued to be used in 2010-12 and 2013-15 codes and standards evaluation studies for appliance standards and certain building standards, which were not covered by a performance compliance method. Past codes and standards evaluation studies did not use energy-based compliance for appliance standards, because it is hard to obtain shipment data for products in different efficiency categories. According to the above discussion, compliance of appliance standards has likely been under-estimated. However, this simple definition of compliance may be part of the reason that the IOUs have been less concerned about compliance evaluation results of appliance standards than those of building standards.

Compliance Adjustment Factor / Energy Savings Adjustment Factor

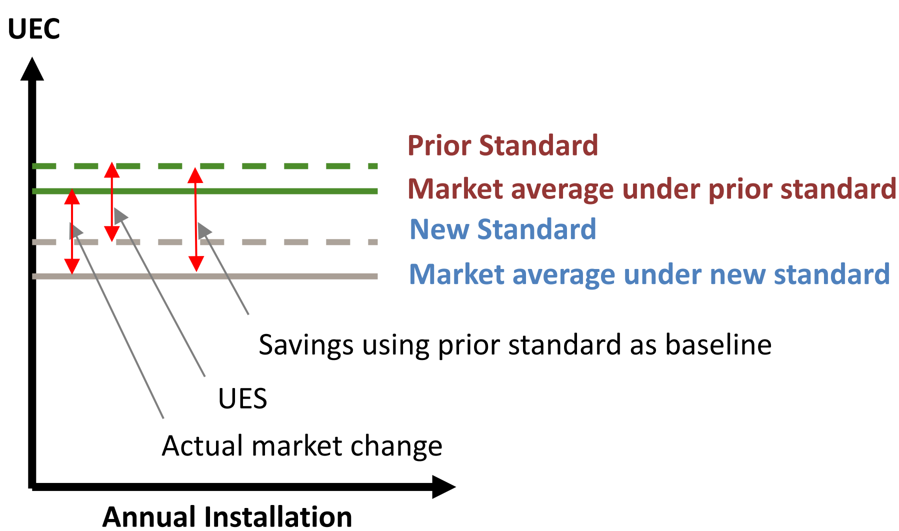
In order to properly capture energy impacts in compliance evaluation, the 2010-12 CPUC codes and standards evaluation introduced the concept of CAF and used it to assess compliance of Title 24 building standards covered by a performance compliance method, including new construction whole building compliance and lighting alterations compliance. According to the CPUC 2010-12 codes and standards evaluation:

- "CAF is defined as the ratio of actual energy savings based on actual building construction to the expected energy savings based on the new Title 24 requirements".

The CPUC 2013-15 codes and standards evaluation renamed the CAF as ESAF. ESAF/CAF method can capture different levels of compliance to properly reflect actual energy savings. Perfect compliance has an ESAF/CAF value of 100%, partial compliance has an ESAF/CAF value between 0% and 100%, over-compliance has an ESAF/CAF value above 100%, and performance below baseline leads to a negative ESAF/CAF value. It should be noted the ESAF/CAF method has only been applied to new Title 24 standards that have a prior Title 24 standard to be used as the baseline for energy savings calculation.

ESAF or CAF assess the amount of savings achieved in the market, but do not directly indicate the level of compliance of the new standard, at least not in a way conforming with common understanding of compliance. Assessing ESAF or CAF requires the consideration of the difference between two standards and how much the actual installation performs better than the prior standard. In contrast, the binary method, used to assess compliance rate, compares efficiency of the sample only to the new standard without any consideration of prior standard or other baseline options. Therefore, it has the advantage of providing a clear indication of compliance of the new standard, even if it is not accurate in terms of energy impact.

Compliance assessment results based on ESAF/CAF method can be hard to understand and, in some cases, misleading. When the expected energy savings of a new standard over the prior standard is small, over-compliance can lead to a high ESAF/CAF value, which may be hard to understand without knowing the details of ESAF/CAF calculation method and specific Title 24 requirements. For some end uses, e.g. nonresidential indoor lighting, it is possible that market average performance is better than Title 24 requirements because over-compliance outweighs non-compliance. After a new Title 24 standard takes effect, which drives up the performance of the overall market, the new average market performance could be better than the new standard, as shown in Figure 10 . The overall market improvement, indicated by the difference between two averages, may be equivalent to the expected energy savings, or UES, determined by the two versions of Title 24 standard. However, the ESAF/CAF value can be significantly higher than 100%, because the actual savings calculated by using the prior code as baseline are higher than those by using market average as baseline. This example shows that ESAF/CAF can lead to confusing or even misleading understanding of compliance status. It also shows the importance that compliance of prior standard needs to be included in assessing actual savings of a standard. We point out that this example is not intended to explain the reasons for high ESAF/CAF values obtained in past CPUC codes and standards evaluations, even though the effect explained here can be a contributing factor.



Savings using prior standard as baseline > UES → ESAF >100%

Actual market change ≈ UES

Figure 10: Market Conditions Relative to Standards

Compliance Margin (CM)

The energy efficiency industry, especially new construction incentive programs, has been using compliance margin as a Title 24 compliance measurement metric before the C&S Program attribution model was established. Compliance margin is calculated as the amount of building energy use below Title 24 requirements, presented as the percentage of Title 24 energy budget. When energy use of a building project is greater than Title 24 requirements, the resulting compliance margin is negative. So, negative compliance margin indicates non-compliance. Compliance margin properly captures different levels of energy performance and is obtained based on comparison to the new Title 24, not the difference between the new and prior Title 24. Therefore, compliance margin is a good indicator of energy-based compliance.

5.2.2 Compliance Adjustment Method

As discussed previously, gross savings should reflect actual energy savings achieved in the market, i.e. to reflect changes of market conditions after the new standard takes effect. To do so, gross savings calculation needs to consider compliance of the new standard in the year for which savings are calculated, as well as market status had the new standard not become effective. The former is captured by the existing compliance adjustment methods, which will be reviewed first. The latter needs to be further investigated since past CPUC codes and standards evaluation had not addressed this issue. The investigation needs to consider implications for different baseline options.

Existing Compliance Adjustment Method

Existing compliance adjustment method for the new standard depends on the type of compliance definition being used.

Adjustment using compliance rate (CR):

$$Gross\ Savings\ with\ Compliance\ Adjustment\ for\ the\ New\ Standard$$
$$= CR_{New\_Std} \cdot Potential\ Savings_{New\_Std} = CR_{New\_Std} \cdot UES_{New\_Std} \cdot Annual\ Installation$$

Because compliance rate is obtained based on a binary approach, which does not account additional savings achieved by over-compliance, the maximum possible value for CR is 100%. Therefore, gross savings never exceed potential savings.

Adjustment using ESAF/CAF:

$$Gross\ Savings\ with\ Compliance\ Adjustment\ for\ the\ New\ Standard$$
$$= ESAF_{New\_Std} \cdot Potential\ Savings_{New\_Std} = ESAF_{New\_Std} \cdot UES_{New\_Std} \cdot Annual\ Installation$$

These two compliance adjustment methods have different implications. Compliance rate is measured based on a binary approach to determine compliance. A sample is considered to either meet the new standard to generate the UES defined in the potential savings assessment or not meet the new standard to have zero savings. Compliance rate indicates the percentage of annual unit installations that would generate savings. Clearly, this simple compliance measurement method cannot properly reflect the diversified compliance scenarios in the market. Compliance rate indicates the percentage of total annual installations that meet the new standard. The installations meeting the new standard are considered to achieve the UES defined in the potential savings assessment. Therefore, the adjustment is essentially applied to annual installation. In contrast, ESAF/CAF specifies the level of savings achieved, presented as percentage of UES. The level of savings is applied to total annual installations. Therefore, ESAF/CAF adjustment method implies an actual average UES is achieved, which are the product of ESAF and UES calculated in potential savings assessment.

Compliance Adjustment for Prior Standard

codes and standards.

For many of the adopted standards, especially Title 24 building standards, a prior standard exists and is used as the baseline for potential savings calculation. For these standards, market status without the new standard is determined by compliance of the prior standard. If compliance of the prior standard was less than 100% right before the new standard took effect, it means the potential savings of the prior standard was not fully achieved. The new standard will drive efficiency improvement to enable the remaining potential savings of the prior standard being realized and then further achieve some or all of the potential savings of the new standard. Compliance adjustment needs to capture both components of energy impact. On the other hand, if compliance of the prior standard was more than 100% right before the new standard took effect, it means the market was improved beyond the prior standard. This additional improvement, represented in terms of the potential savings of the prior standard, should be excluded in calculating gross savings of the new standard.

Compliance Adjustment when there is no Prior Standard

If there is not a prior standard, baseline for the potential savings calculation should be based on ISP. The last chapter discussed possible options of ISP and indicated that market average is the only feasible option. Because market average, calculated as average UEC weighted by annual installation, properly reflect overall market status in energy performance before the new standard takes effect, no adjustment is needed to correct the potential difference between the baseline and actual market condition. In this case, compliance adjustment only needs to consider compliance of the new standard as has been done by past CPUC codes and standards evaluation.

5.2.3 Treatment of Over-compliance

Data collected by the CPUC 2010-12 and 2013-15 codes and standards evaluation studies indicated average compliance rate, measured by CAF or ESAF, was higher than 100% for nonresidential lighting system and nonresidential new construction whole building performance. Both the IOUs and CPUC posed the following questions regarding if savings from over-compliance should be included in gross savings and attributed to IOUs' C&S Program were posed by both.

- Is over-compliance due to NOMAD, instead of compliance of the new standard?
- Do over-compliance savings overlap with savings generated by incentive programs promoting above-code measures?
- Do over-compliance savings overlap with savings for the next code cycle?

Compliance and NOMAD are two intertwined market condition adjustments. NOMAD always exists and, in fact, is a pre-requisite for standard adoption, because it demonstrates measure feasibility. Therefore, some of the over-compliance is reflected in NOMAD. On the other hand, as a new standard is adopted and becomes a minimum requirement for the covered products and practices, innovative market leaders will likely move toward increased efficiency or other improvements in order to maintain their market leadership. Such improvement is driven by the adoption and implementation of the new standard and, therefore, should not be considered natural market conditions.

Additionally, code compliance of individual buildings can exceed minimum code requirements due to normal market conditions or other parameters that are not a direct result of the code requirements. For example, installed products, including glazing, HVAC and water heating equipment may have efficiency performance criteria that slightly exceed minimum criteria. This efficiency improvement occurs for various and multiple reasons depending on the product; not specifically to improve efficiency. When assessed at a whole building energy use level, the difference between the actual product and minimum requirements will show an improvement relative to code. For nonresidential lighting compliance, particularly in small buildings or spaces, the actual installed LPD (in Watts/sf) will not exactly match the required LPD, showing a lower energy use (in Watts).

Another aspect to consider is the effect of whole building compliance on measure adoption. Past codes and standards evaluation results suggested that builders of commercial buildings will install above-code high-efficiency lighting and reduce envelope requirements to meet the overall building performance compliance target. As a result, it is likely that over-compliance exists for Title 24 nonresidential indoor lighting requirements.

Gross savings capture all improvements, due to both compliance and normal market conditions, after the new standard becomes effective. Efficiency improvements associated with normal market conditions are then excluded through a NOMAD adjustment during the next step of the codes and standards attribution model when net standard savings are calculated.

Codes and standards compliance measurements have been based on non-participants of incentive programs. Therefore, the corresponding gross savings reflect improvement of non-participants and do not include additional improvements driven by incentive programs. Incentive programs claim net savings, i.e. excluding savings from those who would have over-compliance even without incentives. Therefore, there is no overlap between over-compliance captured by codes and standards evaluation and net over-code savings achieved by incentive programs.

5.2.4 Compliance Assessment Approach

The CPUC Evaluation compliance estimation approach and process are time consuming, costly, and complex, yet may not be robust and therefore not adequate for the C&S Program attribution evaluation objectives. The compliance rate used to estimate gross savings needs to reflect the actual market. There are "real-world" compliance issues that should be considered when estimating a compliance rate:

- Compliance usually increases over time,
- Compliance with building standards vary by jurisdiction,
- Larger buildings take longer to get through the construction and permitting process, and
- Code compliance, and the impact of code requirements, is reflected by the energy use.

Overall compliance rates vary over time. For T24 building standards, the compliance rate tends to be cyclical with the code update cycle. There tends to be lower compliance when a new code update first takes effect. This happens because there is a normal learning curve that must occur. In addition to the technical changes to the code language itself, such as is provided by the IOU CASE reports, there are typically changes to the compliance forms and software. These changes may result in limited resources, primarily for code officials, to focus on process changes to show compliance, at the expense of the code details. The learning curve and adjustment period result in oversight and omissions that may negatively impact compliance rates but are not specifically related to intentional non-compliance. After the code has been in effect for several years, compliance is higher as the market and code officials better understand the requirements and have established procedures for implementing and verifying code requirements. Since the T24 standards update cycle is typically on a three-year cycle, the California building industry experiences a three-year ebb and flow code compliance cycle. For the gross savings adjustment needed for the Program Attribution model, it is important to understand that taking a snapshot of market code compliance may not reflect the overall, average code compliance rate that is relevant for estimating actual, long-term savings. For codes that are revised approximately every three years, snapshots at different times in the cycle are needed to capture real changes in the market.

Large, complex nonresidential buildings can take years to complete construction and get through the permitting process. These projects have more resources to commit to proper code compliance and permitting, and therefore are prone to better (and more complete) compliance than smaller, simple fast-track buildings. Due to the timeframe of most new construction evaluation studies, both codes and standards and new construction incentive programs, the larger, complex projects are under-represented by the evaluation sample. The underrepresentation skews the study results.

Compliance rates inevitably vary by jurisdiction, and even within a jurisdiction based on the specific permit applications. This is particularly true for nonresidential permits in smaller jurisdictions. These jurisdictions will see a handful of permits over the course of the year that may vary greatly in their scope. Aggregation of these permits over the year provides the best representation of overall code compliance within the jurisdiction. Variation of compliance across the jurisdictions can result from variations in staffing levels, staff knowledge and staff turnover. There may be some consistency in jurisdiction performance, but these variables can impact jurisdictions at various times. A statewide assessment of compliance, not targeted to jurisdictions within the IOU service territories, will expedite the gross savings approach and provide overall code compliance.

For T20 and federal appliance standards, there is also lower compliance in the market for a new code, but for different reasons. Appliance standards regulate minimum efficiency energy thresholds on the manufacturer of a range of appliances and consumer products. Typically, when an appliance standard is adopted the implementation or effective date of the standard is several years in the future to allow manufacturers to retool, as necessary, to have products available by the effective date. There is a normal transition period where older products are still in the market. As the older products flow through the supply chain and are purchased, the newer products begin to show up in the market. If compliance data collection efforts begin too soon, the existing products may not have made it through the market, giving a wrong signal of non-compliance.

Evaluation Compliance Rate Estimation Approach

The overall evaluation approach for estimating statewide compliance rate is provided in the Appendix 9.1. The approach describes the detailed steps that were taken to attempt to establish a reasonable compliance rate. While this is a good theoretical approach to establishing a reasonable compliance rate, the TRC team concludes that not enough sample sites were included in the data collection to ensure robust compliance measurement".

The evaluation approach uses a complex sampling process. The sample frame excluded jurisdictions not served by an IOU, because the evaluation was "only interested in estimating impacts in IOU service territories." This is an unnecessary step for the purpose of establishing code compliance needed to estimate gross savings. Further, this approach implies that the compliance rate needed for this analysis is influenced by the IOUs, which is not related to the codes and standards advocacy efforts. Using this approach suggests that the IOUs are responsible for code compliance specific to these code change proposals.

The evaluation approach used permit data for newly constructed projects permitted under the most recent code, which most likely biased the sample to smaller nonresidential buildings. While the evaluation assumed a nine-month delay between code implementation date and permit date, due to extended construction timelines for larger nonresidential buildings it is likely that more complex projects are not included in the sample. This results in smaller projects, with potentially different design and construction practices and priorities than larger more complex buildings. This is an on-going issue with site data collection for new nonresidential construction studies.

For both residential and nonresidential gross savings, the evaluator had difficulties reaching the target number of buildings to conduct sites visits. Their approach required that they recruit building departments, secure building department permission to review building permit data, visit building department offices to review documentation, contact the building owner, developer or contractor to secure permission, schedule a site visit, and conduct a site visit. Coordination of all of these steps took considerable time. The staged approach of working through specific building departments was not successful in achieving the target sample, and the evaluator directly identified residential builders to find additional site. Ultimately, the targeted number of site visits was not met, and the end result was that the study used a small sample size to determine statewide compliance rate.

The evaluator initially focused on stratifying the nonresidential lighting alteration and new construction populations by building type and climate region because they expected the effects of Title 24 and compliance with Title 24 codes to vary by both factors. Because they faced considerable challenges in recruiting for site visits, these stratifications made it impossible to complete the evaluation of the recommended number of buildings within the evaluation timeline. Therefore, the evaluator revisited the sampling plan to decrease the sample size to remain on schedule. The evaluator incorporated results of their analysis of 2008 Title 24 code compliance margins to update the sample design. Ultimately their sample was stratified only by building type.

Title 20 and Federal Appliance standards compliance estimation is more straight-forward. For a state-regulated or federally-regulated product to be compliant with the California appliance efficiency regulations, its manufacturer must demonstrate that the product meets the performance requirements of the regulations. The evaluators based the definition of compliance on the share of the market sampled that met the efficiency requirements. However, the evaluator did not consider levels of efficiency or energy impact in their compliance assessment. Additionally, near-term compliance can be incorrectly low if shelf surveys[1] are conducted to soon after the standard takes effect, and enough time has not passed to allow for older products to move off the shelf.

5.3 Possible Improvements

5.3.1 Redefining Compliance

CPUC codes and standards impact evaluations have been using two types of compliance definition: compliance rate and ESAF/CAF. Compliance margin is commonly used in the building construction industry to indicate Title 24 compliance of individual projects. Compliance rate is not an energy-based compliance measurement and tends to under-estimate energy performance of the market. ESAF/CAF and compliance margin are energy-based compliance measurements. Compliance rate and compliance margin directly indicate compliance of a standard; while ESAF/CAF indicate energy savings and may lead to misunderstanding of compliance status. Using more than one compliance definition for codes and standards evaluation provide somewhat confusing messages of compliance status.

It is preferable that one energy-based compliance definition is used for all appliance and building standards and this compliance definition can provide a clear indication of overall compliance of a standard and enable actual energy savings calculations. Such a definition of compliance would provide clear understandings of compliance status to the broad energy efficiency practitioners and related market actors.

Among the three compliance definitions investigated by this study, compliance margin meets the above criteria best. It is energy based, provides a clear indication of compliance status, and is widely accepted by the building construction industry. The next section explains how to use compliance margin to perform compliance adjustment. In theory, compliance margin can also be applied to appliance standards. For example, it is normal to state that an air conditioner model is 30% more efficient than minimal standard requirement. This is equivalent to having a compliance margin of 30%, if 30% more efficient is based on the measurement energy consumption as compared to the minimal standard requirement.

There is no barrier to using compliance margin for building standards. For appliance standards, it may take a while for related industry practitioners to understand and get used to this concept. However, any energy-based compliance definition would be new to the appliance standards community; using compliance margin is not expected to cause more confusion, if not much less, than other compliance definitions.

5.3.2 Adopt a New Compliance Adjustment Method

Based on the previous discussion, gross savings calculations should include compliance adjustment for both the new and prior standards:

Gross Savings = Gross savings with compliance adjustment for the new standard
+ Adjustment for compliance of prior standard

The following two sections explain how these two adjustments can be performed.

[1] shelf surveys are visits to retail stores to record comprehensive inventories of the product of interest including availability and pricing.

Gross savings with compliance adjustment for the new standard

This is the first component of gross savings calculation. This section provides an alternative compliance adjustment method, based on compliance margin (CM), instead of compliance rate and ESAF/CAF. The adjustment can be performed using the method represented below. In the following formula, "actual" is used to denote actual market average UEC, "Old_Std" is used to denote UEC defined by the prior standard, and "New_Std" is used to denote UEC defined by the new standard.

$$CM = \frac{UEC_{New\_Std} - UEC_{Actual}}{UEC_{New\_Std}}$$

$$UEC_{Actual} = (1 - CM_{New\_Std}) \cdot UEC_{New\_Std}$$

$$UES_{Actual} = UEC_{Old\_Std} - UEC_{Actual}$$

$$= UEC_{Old\_Std} - \left(1 - CM_{New\_Std}\right) \cdot UEC_{New\_Std}$$

$$= (UEC_{Old\_Std} - UEC_{New\_Std}) + CM_{New\_Std} \cdot UEC_{New\_Std}$$

$$= UES_{New\_Std} + CM_{New\_Std} \cdot UEC_{New\_Std}$$

In the above, we used $UES_{New\_Std} = UEC_{Old\_Std} - UEC_{New\_Std}$

$$Gross\ Savings\ = UES_{Actual}\ \cdot Annual\ Installation$$

$$= (UES_{New\_Std} + CM \cdot UEC_{New\_Std}) \cdot Annual\ Installation$$

$$= UES_{New\_Std} \cdot Annual\ Installation + CM_{New\_Std} \cdot UEC_{New\_Std} \cdot Annual\ Installation$$

$$= Potential\ Savings + CM_{New\_Std} \cdot UEC_{New\_Std} \cdot Annual\ Installation$$

The second component in the above equation represents the compliance adjustment. The adjustment is negative, if overall market CM is less than zero, positive, , if overall market CM is larger than zero.

Adjustment for Compliance of Prior Standard

This is the second component of gross savings calculation. Compliance adjustment for prior standard is to capture unrealized or over-achieved potential savings for the prior standard. The adjustment method depends on the type of compliance definition being used.

Adjustment using compliance rate (CR):

$$Gross\ Savings\ from\ Compliance\ Adjustment\ for\ the\ Old\ Standard$$

$$= (1 - CR_{Old\_Std}) \cdot Potential\ Savings_{Old\_Std} = (1 - CR_{Old\_Std}) \cdot UES_{Old\_Std} \cdot Annual\ Installation$$

Adjustment using ESAF/CAF:

$$Gross\ Savings\ from\ Compliance\ Adjustment\ for\ the\ Old\ Standard$$

$$= (1 - ESAF_{Old\_Std}) \cdot Potential\ Savings_{Old\_Std} = (1 - ESAF_{Old\_Std}) \cdot UES_{Old\_Std} \cdot Annual\ Installation$$

Adjustment using compliance margin (CM):

$$Gross\ Savings\ from\ Compliance\ Adjustment\ for\ the\ Old\ Standard$$

$$= -Annual\ Installation \cdot CM_{Old\_Std} \cdot UEC_{Old\_Std}$$

5.3.3 Acknowledge and Integrate Over-compliance

Based on previous analysis on related issues, we recommend that over-compliance be included in the gross savings calculation. Doing so would not lead to savings over-estimation for several reasons. First, as indicated previously, codes and standards compliance evaluation does not include samples from incentive program participants and, therefore, any selected samples with over-compliance are not due to incentive programs.

Second, over-compliance savings associated with NOMAD will be excluded in the net standards calculation. Third, with the recommended compliance adjustment for prior standards in gross savings calculation, saving from over-compliance will be excluded from gross savings of the succeeding standard. This proposed improvement to gross savings calculation ensures savings from different standard revisions are properly captured without any overlap and gap between any two succeeding standards. With this approach, gross savings show continuous market improvement as layers of energy use reduction over time. A layered approach should be used to guide the impact assessment to consider impact of succeeding standards Even though compliance of individual standards may not be assessed with high accuracy due to resource limitations and practical difficulties. Correspondingly, gross savings of individual standards may not be perfectly accurate, but the long-term market improvement can be correctly captured. Therefore, including over-compliance savings does not double count savings. Savings overlap and gap between standards could happen when layering effects of standard improvement are not reflected in codes and standards saving assessment.

5.3.4 Establish a More Robust Compliance Assessment Approach

Determining a robust and accurate compliance rate, especially for nonresidential buildings, is very costly and difficult. There are probably other ways to determine compliance than the way the CPUC evaluators have done so far. The cost and timing that is required and in-depth knowledge that is required to determine compliance is high.

Given the complexity and practical challenges, a large amount of resources is needed for accurate compliance evaluation. codes and standards compliance evaluation can be conducted with more strategic planning.

For selected high-impact standards, comprehensive and well-designed compliance studies are needed to provide accurate and comprehensive compliance data to support codes and standards impact assessment as well as other incentive programs. Codes and standards compliance evaluation needs to leverage data from other programs and activities or be integrated with market studies or evaluation studies for other programs, e.g. new construction related incentive programs.

For standards with less impact or when resources are limited, codes and standards evaluation can streamline compliance evaluation by using default compliance values:

- Use default values based on input from the CEC and industry experts.
- Use a default assumption on how compliance improves over time, based on input from the CEC and industry experts.
- Use consistent compliance assumptions for incentive programs and C&S Programs.
- Incorporate the compliance adjustment method presented in section 5.3.2 to ensure savings from different cycles are counted as layer of improvement to eliminate savings overlap and gap.

CPUC evaluation should look at how to integrate new construction studies with codes and standards studies so as not to duplicate the work, but rather leverage the studies to get more robust data.

Using prevailing compliance rates and code, they should match up with the other programs with the new construction programs (evaluation has used this approach in the past for residential compliance).

There needs to be a comprehensive, long-term compliance study that establishes compliance rates used as pre and post conditions for new construction and C&S Programs.

6. Net Standards Savings

This section describes the investigation of the net standards savings calculation with a focus on evaluation methodologies for naturally occurring market adoption (NOMAD).

6.1 Key Issues

Throughout this chapter we discuss key issues related to accuracy, streamlining and transparency for net savings assessment.

ACCURACY	STREAMLINING	TRANSPARENCY
Lack of accuracy: • NOMAD curves developed through the Delphi process were not validated or cross-checked based on any market data. • Limited data sources for providing actual NOMAD information. • Quantification of naturally occurring market influences.	Opportunities to streamline: • Alternative methods to simplify NOMAD evaluation and possibly improve evaluation accuracy.	Lack of transparency: • NOMAD curves provided by different experts have very significant differences. Reconciliation of the differences is not clear. • How Delphi Panel experts are selected. • How do IOUs' energy efficiency efforts influence NOMAD

Net standard savings represent the energy impact due to the standard. They are obtained by excluding market improvement that would happen without the standard or would happen within natural market adoption (NOMAD), from the total market improvement since the standard becomes effective, which is addressed in the gross savings calculation. However, quantifying all naturally occurring market influences is very difficult, because NOMAD represents counterfactual, not actual, market trends.

The 2006 California Energy Efficiency Evaluation Protocol Codes & Standards NOMAD calculation section includes the following guidance for developing NOMAD impacts:

"…Naturally occurring adoption rates for premium energy efficient products typically occur in an "S" shape pattern that never reaches 100 percent penetration as long as there are alternative technologies in the market. This step requires the evaluation contractor to establish expected adoption curves for each technology included in the impact assessment. The evaluation contractor will use a range of approaches to establish the estimated penetration curves, including conducting literature searches on the penetration rates of similar technologies with similar product characteristics, the use of expert opinions on the expected penetration rates in the absence of a requirement to use the technology, relevant market data and other approaches as deemed appropriate in the evaluation planning effort."

Past CPUC evaluation studies used the Bass diffusion model to estimate NOMAD and a Delphi expert panel approach to obtain Bass diffusion model parameters (details of this approach are provided in the next section). Interviewees expressed the following concerns on this NOMAD assessment approach:

- For many standards, NOMAD curves provided by different experts have very significant differences, even after outliers are removed. It is not convincing that the existing evaluation method can provide reliable NOMAD estimate.

- NOMAD curves developed through the Delphi process were not validated or cross-checked based on any market data to the extent we could validate with prior CPUC evaluators.
- The expert recruiting process is not transparent enough. It is unclear if experts recruited by CPUC evaluators have proper qualifications needed to perform NOMAD assessment according to California's market conditions.
- Can the Delphi method be improved to provide more transparency on how experts develop NOMAD curves?
- Are there alternative methods to the Bass diffusion model and Delphi method?

To address these concerns and develop possible improvement options, we investigated how should NOMAD be considered, how should NOMAD be modeled and how should NOMAD model parameters be assessed.

6.2 Analysis

6.2.1 How to Consider NOMAD

It is important to understand how the market operates in the context of what can be considered as NOMAD. The Net Standards Savings method should account for all the naturally occurring influences. Naturally occurring market conditions reflect the market absent additional (new) codes and standards updates.

Energy efficiency improvement in the market is influenced by multiple market mechanisms as shown in Figure 11. Pure market process (shown at the bottom of the figure) is due to interaction between demand and supply, e.g. consumers and appliance manufacturers. The pure market process can lead to efficiency improvement and cost reductions due to the demand for products and services with better performance and lower operational costs and competition on the supply side to provide more efficient products. The pure market process is further affected by different drivers (as shown in the middle layer of Figure 11), all of which influence prices (typically reductions) that have an iterative impact on market penetration and adoption including:

- Non-utility energy efficiency advocacy
- Utility non-codes and standards programs
- Voluntary standards
- Mandatory standards[1]
- Intervention on compliance of mandatory standards

Governmental policy directives, e.g. California Zero Net Energy (ZNE) goals and SB 350 double energy efficiency goals, certainly have significant impact on efficiency improvements. These policy directives are implemented by governmental agencies through programs and regulations. These policy directives influence the market through efficiency drivers.

[1] Note in this context mandatory is used in comparison to voluntary – such as incentive programs – and should not be confused with mandatory requirements in the building energy standards as compared to prescriptive or performance requirements.

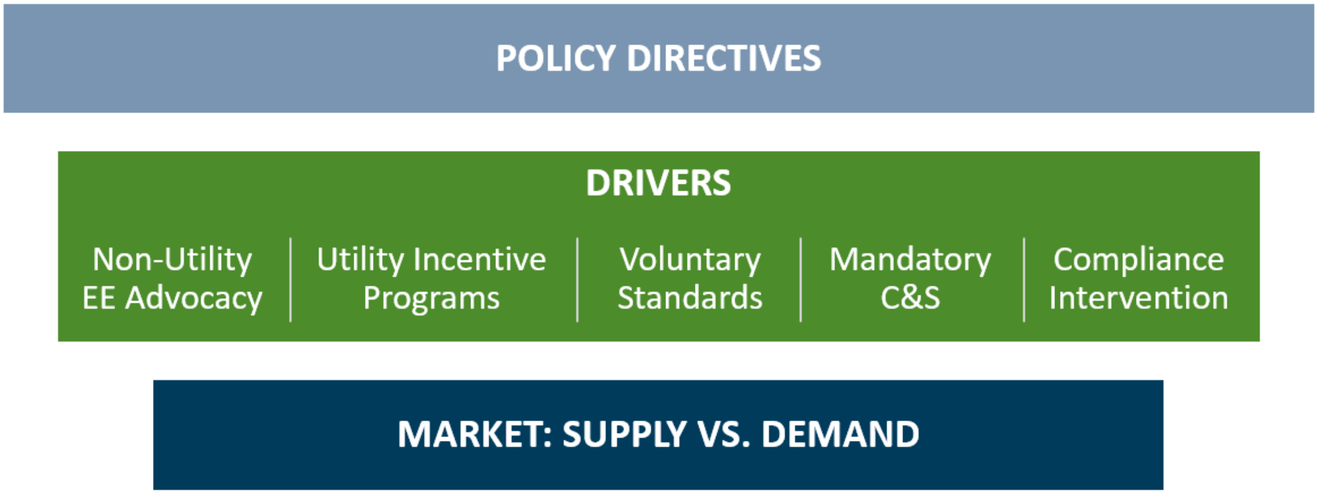


Figure 11: Overall Energy Efficiency Landscape

NOMAD evaluation needs to determine which market mechanisms presented in Figure 11 should be considered and how their influence is addressed. To facilitate the consideration, these market mechanisms are explained in more details.

Policy Directives

Policy directives from the CPUC and CEC influence the direction of the market, even if the market players may not be aware of them. Policy directives are provided in the California Long Term Energy Efficiency Strategic Plan (Strategic Plan). The Strategic Plan provides a comprehensive, integrated framework of the state's goals and strategies for saving energy, covering government, utility, and private sector actions, which directly or indirectly influences market movement. For example, The Strategic Plan establishes market transformation as a unifying objective. The market transformation objective aims to encourage suppliers and manufacturers to produce and sell efficiency products or services and to encourage consumers to buy these products or services. The Strategic Plan recognizes that transformation is an evolutionary process, and that markets for a given end-use or sector are transformed continuously. The Strategic Plan was developed in response to passage of the California Global Warming Solutions Act of 2006 (Assembly Bill 32, AB 32). In 2015, Senate Bill 350: Clean Energy and Pollution Reduction Act (SB 350) established clean energy, clean air, and greenhouse gas reduction goals as a key part of California's climate change strategy. SB 350 requires the state to double statewide energy efficiency savings in electricity and natural gas end uses by 2030. AB 1109 requires specific reductions in lighting energy use. AB 32, SB 350, AB1109 and other California legislation and policy including AB 2021 (establishment of statewide energy efficiency goals), the Low-Income Energy Efficiency statutes, and the Governor's Green Building Executive Order14 create a market environment where energy efficiency efforts are integrated into most aspects of the building industry. Many California jurisdictions have developed local Climate Action Plans to meet AB 32 goals.

EE Driver: Non-Utility EE Advocacy

Non-utility energy efficiency advocates influence the market in various ways, however, in the manner most analogous to the utility efforts, their activity typically runs in parallel to utilities. These include informational campaigns to educate the market on the benefits of energy efficiency products and practices. Examples include the DesignLights Consortium promoting high performance lighting products, ACEEE promoting multiple efficiency products and technologies such as water heating equipment, and regional energy organizations (REOs), such as the Northwest Energy Efficiency Alliance (NEEA), that promote energy efficiency through various activities across the country, as well as participate in standards development at the federal level. These organizations work collectively at the national level to promote energy efficiency improvements in appliances and buildings.

EE Driver: Utility Non-Codes and Standards Programs

The California IOUs have implemented energy efficiency rebate and incentive programs targeting consumers, building owners, builders, and designers since the late 1970s/early 1980s. They have developed and implemented programs that engage midstream and upstream market actors including retail, suppliers, distributors, and manufacturers. The utilities program portfolio includes non-resource programs including the Statewide Emerging Technologies program that supports the development and deployment of promising new technologies, practices, and tools.

EE Driver: Voluntary Standards

IOU C&S Programs have been actively involved in the development of voluntary energy efficiency standards and specifications (VSS). Some of the important VSS include ASHRAE[1] 90.1 Energy Standard for Buildings Except Low-Rise Residential Buildings, ASHRAE 90.2 Energy-Efficient Design of Low-Rise Residential Buildings, ASHRAE 189.1 Standards for the Design of High-Performance Green Buildings, International Energy Conservation Code (IECC), and Energy Star™ specifications for various appliances. These VSS play important roles in advancing energy efficiency in the market. For example:

- Voluntary test standards developed by the industry provide the basis for the development of many state and federal appliance standards;
- The California Energy Commission (CEC) references to the latest ASHRAE and IECC standards in updating Title 24 building standards;
- EnergyStar® is well-recognized by consumers for guiding their selection of efficient products.

EE Driver: Mandatory Standards

Mandatory standards include California's Title 24, Part 6 Building Standard, Title 20 Appliance Standards, and the federal appliance standards.

EE Driver: Intervention on Compliance of Mandatory Standards

Compliance intervention activities occur through the IOUs compliance improvement through Energy Code Ace and other compliance intervention activities. IOU compliance intervention targets market actors throughout the entire compliance chain including product manufacturers and suppliers, building designers and contractors, energy consultants and building officials. Specific activities include providing education, training, outreach, and technical support, tools, and other resources to improve compliance with both the building and appliance energy standards. The program also carries out strategic activities that support or shape future codes and standards.

Additionally, the IOU compliance improvement and codes and standards Advocacy teams integrated compliance improvements, and subject matter experts within the CASE process during the last T24 update cycle. Not only was this valuable to get the compliance perspective in the advocacy and CASE process, but it allowed the compliance improvement team to get an early start on understanding the code update and developing appropriate resources and trainings. The early engagement allows the compliance improvement team to identify the biggest changes and prioritize engagement with specific market actors. The integration effort has been valuable for both teams.

[1] American Society of Heating, Refrigerating and Air-Conditioning Engineers

Collective Impact of EE Drivers

Among all EE drivers, mandatory standards and compliance intervention are associated with codes and standards. Efficiency improvements and price reductions due to pure market process and other EE drivers (non-utility EE advocacy, utility incentive and non-resource programs, and voluntary standards) is considered to be NOMAD. Policy directives affected both NOMAD-related EE drivers and codes and standards-related EE drivers. The former contributes to NOMAD. For NOMAD assessment, there should be considerations on how policy directives would be achieved. Can policy directives be achieved through NOMAD-related EE drivers? What are the roles in codes and standards in achieving policy goals? How were efforts and activities from NOMAD-related EE drivers and codes and standards-related EE drivers adjusted according to policy goals? How much of the policy goals would be achieved through voluntary market responses to NOMAD-related EE drivers, without mandatory standards? These questions may be considered during NOMAD assessment.

It should also be noted that some of the NOMAD-related EE drivers are related to utility EE efforts. Utility incentive programs obviously represent utilities' efforts. Promoting voluntary standards development has been part of the C&S Program's long-term efforts in its holistic codes and standards development plan. C&S program's efforts in this area promote voluntary market adoption and facilitate corresponding future mandatory standard development. The TRC White Paper on Voluntary Standards and Specifications, developed on behalf of SCE (Services, 2016) provides detailed analysis on related issues and can be used to inform possible evaluation method improvement specifically related to C&S program's efforts on voluntary standard development.

Past CPUC codes and standards evaluation studies included all NOMAD-related EE drivers for overall NOMAD assessment. CPUC evaluators also established a method to correct NOMAD by removing market effects introduced by utility incentive programs. For the 2006-2008 and 2010-2012 codes and standards Impact Evaluations, CPUC acknowledged that IOU incentive programs can have significant influence on the market and this should be excluded from NOMAD (DNV GL & CADMUS, 2014). CPUC evaluators adjusted NOMAD according to IOU incentive program influence; therefore, crediting the IOUs for pre-code market influence. However, it is unclear whether during the Delphi panel process (discussed in further detail below) the panelists excluded this market influence in their assessment of market conditions.

For considering net impact of a standard, correction of market effects introduced by incentive programs is not necessary, because these market effects, regardless who cause them, are not due to the standard. However, in the context of considering net standard savings attributable to IOUs, these market effects should be considered. In addition, market effects associated with utilities' voluntary standard development efforts should be considered. Methods for considering these market effects will be addressed in the next section on how to model NOMAD.

Interviewees suggested that codes and standards rulemaking and even proposed standards development can increase market adoption of corresponding efficiency measures, because these activities increase market awareness of the efficiency measure and some market actors, e.g. manufacturers and building designer, may improve their products and services ahead of the standard adoption to ensure they are ready for pending regulations and to maintain competitive edge. In other words, NOMAD could be boosted in a short period (could be several years depending on specific standards) before the adoption due to standard development and rulemaking efforts. To obtain actual NOMAD without the influence by standard development processes, NOMAD assessment could rely on market information and trends before the standard development efforts started.

Another issue to consider is the long-term effects of codes and standards development on NOMAD. The development of new products, technologies and construction practices are already influenced by having the code and on-going code improvement in play. Each standard adoption stops the existing NOMAD process, pushes the market to a higher level, and forces the NOMAD process to restart from that level. In fact, the intention of codes and standards development is to interrupt the NOMAD process and continuously speed NOMAD. The market would not advance as quickly without codes and standards development. Assessment of

NOMAD associated with individual standards over-estimates long-term naturally occurring market progress. From this perspective, what is needed is an estimated long-term NOMAD, representing average efficiency improvement of the market, without any efficiency development. The long-term effect with and without robust codes and standards development is clearly demonstrated by the well-known comparison of per capita energy use trend between California and the rest of the United States, shown in Figure 12.

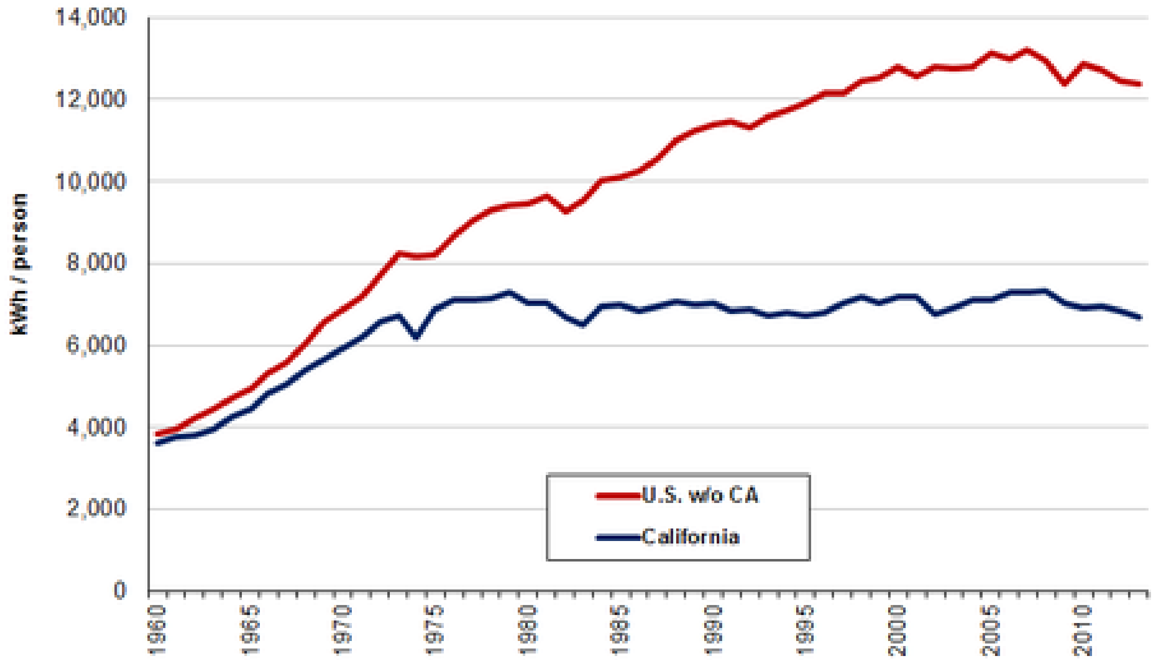


Figure 12: California and U.S. per Capita Energy Use

For codes and standards evaluation, it is desirable to estimate long-term NOMAD without any codes and standards influence. This might be too difficult to do for all standards. Alternatively, efficiency improvement of the market, without any standard development, in different technology or end-use areas, could be used to indicate NOMAD for associated standards.

IOU and non-IOU influences are illustrated graphically in Figure 13. Influences on efficiency improvement by the different market mechanisms, and their impact on determining NOMAD are summarized in Figure 14.

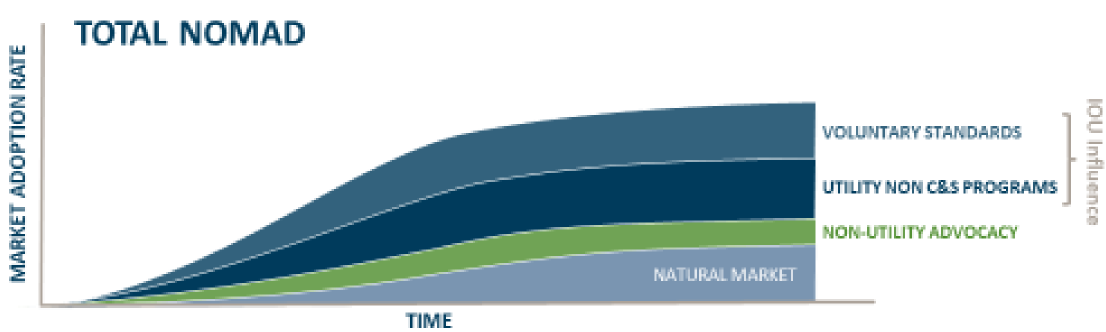


Figure 13: NOMAD Influences

	Market	Non-Utility EE advocacy	Utility Non Codes and Standards Programs	Voluntary Standards	Mandatory Codes and Standards	Compliance Intervention	Policy Directives
Types of activity	Respond to consumer signals, program opportunities, laws, government, and policy directives	Participate on standards setting committees and public review process. Educate market	Incentives and rebates Educate market	Initiate standard proposal, participate on standards setting committees and participate in public review process	Initiate standard proposal, participate on standards setting committees and participate in public review process	Educate market Develop tools to make compliance easier and less error prone	Set state and federal energy efficiency and carbon reduction goals and objectives
Types of impact on the market	reactive	Market awareness	Market awareness, early adoption of new technologies and products	Develop test standards which enable development of product performance standards.	Develop mandatory appliance and building performance standards	Title 24 compliance tools, training resources to understand how and when to comply with California's building and appliance energy efficiency standards.	Regulator mandates
Levels of impact	Market reacts to multiple influences. Market innovators will drive to the next level.	Increases consumer awareness (EnergyStar)	Impacts market innovators	Promote market transformation	Locks in energy savings	Improves compliance process and outcome	Establish goals and parameters for achieving energy efficiency and carbon reduction goals

Figure 14: Market Influences

6.2.2 How to Model NOMAD?

In this section we analyze the approach used by the past codes and standards evaluations (2006-08, 2010-12, 2013-15) to model NOMAD. The existing NOMAD model is based on the Bass Diffusion model. There were inconsistencies in applying the Delphi panel method within and across the evaluation years; for instance, expert panels consisted of different number of individuals from measure to measure, meaning that individual responses have greater weight in smaller panels.

The literature is sparse on examples or assessment of applying a Bass model or other diffusion of innovation theories for energy efficiency measures, especially for buildings. In general, most of the literature either explicitly or implicitly indicated that the Bass model is not ideal for non-binary choices (such as efficiency measure options) but is still the best option available, due to limited relevant alternatives. The most important inputs are the starting and end (maximum) market adoption and the leading adoption behavior. From Diffusion of Energy Efficient Technology in Commercial Buildings: An analysis of the Commercial Building Partnerships Program study, the assessment concluded that the Bass model may be more appropriate for whole building energy efficiency (a single efficiency metric) rather than looking at individual measures or combinations of measures with multiple options (non-binary) which becomes complex (Antonopoulos, 2013).

There are many other variations of market diffusion theories and models for different types of technologies and market processes. The differences among these market diffusion models are mostly due to differences in detailed market mechanisms regarding how information is propagated and how adoption decisions are made. Some of them might be better than others for modeling market adoption of energy efficiency measures. Due to resource limitation, C&S Program evaluation is unlikely to be able to collect enough market adoption data to investigate detailed diffusion mechanisms for energy efficiency technologies and determine which diffusion model could be a better alternative to the Bass diffusion model. It should be noted that past CPUC codes and standards evaluation studies used the mathematical format of the Bass diffusion model to estimate NOMAD trend but did not try to assess diffusion mechanisms for related energy efficiency measures. In other words, the two Bass diffusion model parameters, p value (population of innovators) and q value (population of imitators), are only used to construct an S-shape adoption curve. Other mathematic formula having an S-shape curve may also be used without impact on evaluation accuracy.

In NOMAD modeling, the meaning of adoption rate also needs to be further clarified. The common explanation of adoption rate is percentage of annual installation that would install the efficiency measure. There are usually multiple levels of efficiency, meeting or exceeding the standard, available in the market. Adoption of any of these efficiency options can be considered to be NOMAD, had the standard not been adopted. A market share-based adoption rate does not reflect which efficiency levels were adopted and, therefore, cannot fully represent the level of efficiency improvement, or energy impact, associated with NOMAD. For accurate energy impact assessment, NOMAD should be available to indicate how average market performance, in terms of UEC, changes over time, had the standard not been adopted. This would allow energy-based NOMAD correction. In the previous last section, we also suggested that NOMAD can be better reflected by using long-term market naturally occurring UEC. In addition, compliance analysis presented in the last chapter also indicated that compliance should reflect market average UEC as compared to standard requirements.

If the concept of NOMAD is expanded to capture naturally occurring UEC changes, the concept of diffusion process needs to be expanded from measure adoption rate to capture UEC reduction over time.

6.2.3 How to assess NOMAD Model Parameters?

Past CPUC codes and standards evaluations used a Delphi panel approach to estimate Bass diffusion curves for NOMAD assessment. In this approach, each member of an expert panel individually develops Bass curves that graphically represent the natural market adoption of the code or standard level of efficiency without the codes and standards process.

Existing NOMAD Approach

The existing NOMAD estimation used a modified Delphi approach to collect the information necessary to build the Bass curves for each code and standard. The CPUC evaluator identified panelists for each evaluated code and standard able to provide an informed estimate of naturally-occurring market adoption of that code or standard. The evaluators developed screening criteria for identifying and selecting Delphi panelists.

Once selected, these panelists were requested to provide their informed estimates via the online Market Adoption tool. After all panelist provided their initial estimates, each individual had the opportunity to modify his or her estimate after reviewing those offered by their peers or confirm their original estimate. The online tool allowed the panelists to provide the supporting reasons behind their estimate, in text format. The Market Adoption tool has several inputs for describing the shape of expected market adoption (Antonopoulos, 2013). The tool provides visualization of how the factors impact the shape of the adoption curve. The evaluators use the panelists' responses to develop convergence of the curve estimates.

Concerns Raised with Existing NOMAD Approach

Interviewees raised several issues and concerns with the existing Delphi panel approach. The IOU codes and standards standard team members we interviewed commented that although in the past the impact evaluators shared the Delphi panelists selection criteria, they did not share the results of the process in terms of the final panel members.

Assembling the Delphi panels is challenging because evaluators have a hard time finding experts. Many of the people who can provide the most valuable market insight believe the information is proprietary or they want to get paid for the information. Some individuals who agree to be panel members, are not necessarily familiar with California codes and standards activity. It is not sufficient to have respondents who are knowledgeable about specific technologies/measures but who have not been involved in the code-setting process. These people may be technology area experts and will provide insight and opinions on efficiency and technology trends, but unless they understand and have been involved with the process, the information isn't particularly helpful for establishing NOMAD. Additionally, there may be bias in having respondents over-value and over-emphasize their own involvement.

This Delphi panel method is relatively straightforward and simple, which streamlines the process, but it may not be completely appropriate or accurate for measures that allow for a range of efficiencies and technology choices, where the technologies compete for market share. The evaluation report does not fully document the factors panelists considers in in their responses and how they might consider competing technologies in assessing the market.

Since the resulting curves are not based on clear instructions or consistently available data, it is difficult to identify if the results give an accurate estimate of what would have occurred absent the code or standard. For example, panelists are not told to either include or exclude incentive program influences, therefore it is unclear whether and what each panelist considered. Because the approach cannot be tested, and the resulting estimates cannot be systematically replicated it is difficult to determine if the process is sound and the final estimates are valid. The resulting NOMAD curves lack transparency and are difficult to test for accuracy since the process and result are not replicable.

In reviewing the detailed results by measure it is difficult to understand the analysis process for determining the final result. For some measures and technologies, the panel gives wildly different feedback, which is hard to reconcile. If the Delphi panelists are market experts, one would expect to see consistent patterns in the NOMAD curves, but for many measures there is no consistency. As stated above, the concept of the Delphi panel is that there would be convergence of the results. If the Delphi panelists are technical experts, but don't necessarily understand the larger market context and the influence of codes and standards and other factors, they may not be properly accounting for all NOMAD factors.

Addressing Concerns Raised with Existing NOMAD Approach

For increased transparency, the evaluators should provide descriptions of how they are interpreting and averaging curves provided by each panel member. Particularly for some of the individual outlier curves, it would be helpful for the evaluators to provide the rationale given by the respondent, so that the curves can be better understood. It is unclear from the results section how the curves are used in the analysis to arrive at the final adjustment value.

Is there a better approach to estimate market effects associated with NOMAD? The Delphi approach is based on experts' judgment. Some stakeholders suggested that a data-driven estimation approach should be used. The CPUC 2009 Market Effects and Market Transformation study provided some insight into these two approaches (KEMA, 2009, pp. 90-101) . The following tables provide summaries of the strengths and weaknesses identified in the report, which overall align with the TRC team's assessment.

Structured Expert Judging: Structured expert judgment studies assemble panels of individuals with close working knowledge of the technology, infrastructure systems, markets, and political environments addressed by a given energy efficiency measure to estimate baseline market share and, in some cases, forecast market share with and without the program in place. Structured expert judgment processes employ a variety of specific techniques to ensure that the participating experts specify and take into account key known facts about the program, the technologies supported, and the development of other influence factors over time. The Delphi process is the most widely known method of this family of methods.

<table>
<tr><th>STRENGTHS</th><th>PRACTICAL CHALLENGES</th></tr>
<tr><td>Strengths for application in market effects evaluation. The principal strengths of expert judgment approaches applied to the assessment of energy efficiency programs are the following:
• Accuracy relative to more data-intensive methods. As mentioned above, academic marketing researchers generally rank expert judgment equal to or, in some cases, better than more data intensive modeling approaches for accuracy of results, based on retrospective review. While it is true that there are no “right” baseline estimates, the ability of the basic method to yield acceptable results in other forecasting contexts provides support for its use where it is otherwise appropriate.
• Development of actionable insights. The iterative process very often calls forth detailed insights into market structure and dynamics as the participants provide evidence to support their forecasts.
Limitations for application in market effects evaluation. The literature review uncovered many applications of structured expert judging approaches to assess the outcomes of market interventions. These studies addressed the full range of program types, as well as the full range of retrospective and prospective energy and economic benefits. Some types of expert judging methods, such as the Delphi technique and related approaches, are specifically designed to capture and structure opinions regarding future trends in the market acceptance of products and services.</td><td>Practical challenges. The key practical challenges to using expert judgment approaches in the assessment of energy efficiency programs are as follows:
• Recruitment of appropriate panels. Depending on the nature of the product or markets addressed, it may be difficult to identify and recruit an appropriate expert panel.
• Management of logistics and schedule. The logistics of a structured expert study are complicated, especially if the approach entails face-to-face meetings or multiple iterations. Combined with difficulties in recruiting the panel, the logistics of study administration can make it difficult to maintain a tight study schedule.
Promotion of transparency. Even when program staff suggest scenarios and provide fact packets, it is not practical to identify all of the important assumptions that judges bring to their forecasts. This lack of transparency can complicate the interpretation of results.</td></tr>
</table>

Historical Tracing: Case Study Method: This method involves the careful reconstruction of events leading to the outcome of interest, for example, the launch of a product or the passage of legislation, to develop a 'weight of evidence' conclusion regarding the specific influence or role of the program in question on the outcome. Historical tracing relies on logical devices typically found in historical studies, journalism, and legal argument. These include:

- Compiling, comparing, and weighing the merits of narratives of the same set of events provided by individuals with different points of view and interests in the outcome.
- Compiling detailed chronological narratives of the events in question to validate hypotheses regarding patterns of influence.
- Positing a number of alternative causal hypotheses and examining their consistency with the narrative fact pattern.
- Assessing the consistency of the observed fact pattern with linkages predicted by the program logic model.

STRENGTHS	PRACTICAL CHALLENGES
Strengths for application in market effects evaluation. The historical tracing approach is particularly useful for the following aspects of market effects evaluation: Validation of program theory logic models in the early stages of program or market development. In the early stages of program and/or market development, targeted market changes such as the commercialization of product designs, manufacturers' licensing of product designs, or acceptance of test procedures to measure energy efficiency simply are not susceptible to quantification or formal testing. In these cases, historical tracing, as exemplified by the NSPC case above, is the only method available to explore the relationship between market development and program activities. Analysis of the unique contributions of a given program in a market targeted by multiple programs. Generally, this type of analysis is interesting primarily in the later stages of program and market development. Historical tracing is the only practical approach to address the unique contribution of one program out of many active in a given market. Development of actionable insights. Historical tracing requires in-depth questioning of market actors regarding their motivations and inhibitions to promoting or adopting energy efficient products and services, as well their perceptions of the subject program. This information can be very useful in developing recommendations to improve program effectiveness. Limitations for application in market effects evaluation. By itself, historical tracing cannot be used to develop quantitative estimates of program effects. However, detailed information on market structure, market actor behavior, and self-reported program responses can be very useful to supplement and to serve as a sanity check on quantitative analyses program attribution.	**Practical challenges**. The key practical challenges to using historical tracing in the assessment of energy programs are as follows: Identification and recruitment of interviewees from all groups of key informants. The usefulness and credibility of historical tracing analysis depends on the analysts' ability to capture the views of all key groups, since the assessment of the strength of evidence lies primarily in checking the consistency of these views. Promotion of transparency. Since historical tracing does not rely on formal methods of hypothesis testing, it is up to the reader to assess how convincing the analyst's reconstruction of events is. To do this, the reader must be aware of the informants' role in the market, their economic interests, and the potential effects of the program on those interests. Reliability. Unlike statistical studies, there are no formal criteria in case studies for assessing reliability, that is: the likelihood that another researcher would arrive at the same conclusions if presented with the same data. In case studies, the researcher by definition exercises a great deal of discretion in selecting and shaping the information to be presented from the large trove of data he or she collects."

6.3 Possible Improvements

Our analysis indicates that NOMAD could be based on consideration of energy impact, rather than market share of efficiency measure adoption, so that it would be consistent with energy-based compliance evaluation to make the model more accurate. An energy-based NOMAD indicates how UEC reduction changes over time for a technology or end-use area. Such a NOMAD curve can be used for different standards belonging to the same technology areas or serving the same end-use area. For technologies with dual fuels impact, such as building envelope components, or fuel substitution HVAC strategies, UEC for both fuel types would need to be assessed, or a common kBtuh metric could be developed. A comprehensive strategy would need to be developed by the evaluation team to determine the best overall approach. This approach could allow the NOMAD evaluation to focus on collecting useful data through market studies to inform NOMAD assessment. This will encourage a more data-driven approach that can be tested for validity and reliability. The TRC team has not investigated in-depth the potential pool of data sources but suggests that evaluators could compile and review previous and planned studies, as well as incentive program data as described further below.

An alternative approach to the development of NOMAD curves by a Delphi expert panel approach is to greatly simplify the approach, acknowledging that the previous level of effort was not achieving the implied precision. The simplified approach of a pre-determined NOMAD by technology or end-use that would not require engaging a Delphi panel would acknowledge less precision, instead of inferring it exists though the use of a more complex process. The TRC team has not investigated what values should be used for deemed NOMAD but suggest that evaluators could base these values on a review of previous studies.

6.3.1 Data-driven NOMAD Curves

NOMAD curves need to be drawn more carefully based on market transformation activities and information. The contribution of other market mechanisms needs to be recognized for altering the market as discussed above. Data-driven NOMAD curves should be able to account for all market influences. Developing and utilizing long-term NOMAD trends in absence of codes and standards individual building codes or appliance standards improvements could be a better way to reflect true codes and standards impact over time. If this approach is accepted, NOMAD assessment can further be greatly streamlined by eliminating or reducing the need to generate new curves for individual codes and standards.

CPUC evaluators should utilize existing resources to collect more market data to develop efficiency improvement trends to inform NOMAD curves. Evaluators could base market adoption trends on a review of previous evaluation and market studies, planned market and potential studies as well as detailed energy efficiency program data available through the CPUC.

Many market studies exist that provide energy use per technology for various technologies including market characterizations, manufacturer interviews, investigation of participation rates, new construction incentive program data, etc. A great deal of data is collected but is not necessarily readily available. The evaluators would have to conduct a thorough secondary research exercise, looking at resources (listed below) not previously considered for codes and standards impacts.

- ***Market characterization data***. Market characterization data can be useful to show penetrations by technology or end use, as either percentage of square footage or sales. Many market studies include market actor interviews that may yield information on market trends.
- ***Manufacturer Interviews***. Manufacturer interviews in particular are a good source for establishing the trajectory of product and technology development. The evaluators could use a combination of mining existing resources and conducting complementary primary research to develop NOMAD curves and improve the accuracy over time.
- ***Market studies combined with program data.*** Market studies combined with program data could provide a perspective of how the market reacts to new energy efficiency opportunities and how standard practice changes. Incentive programs participation rate reflects the willingness of the market

to adopt technologies. For example, if only a small percentage of customers are willing to embrace lighting controls when provided with an incentive, an even smaller percentage would do so as part of naturally occurring market adoption.

- ***New construction incentive program data***. New construction incentive program data, in particular, can provide information by technologies and UEC.
- ***Potential study adoption models***. Potential study adoption models developed for incentive programs can be used in a similar manner to provide a perspective of market adoption trends for new technologies.

Additionally, CPUC evaluators should work with the Energy Commission Forecasting office to leverage data from the Integrated Energy Policy Report (IEPR) process. For this study, the TRC team contacted the Forecasting office to better understand their analysis model, required inputs and outputs. While we were not able to obtain the model, the CPUC should be able to obtain the model and coordinate with the Energy Commission to understand the granularity of the inputs and the useful of the outputs. With a better understanding of the IEPR data, the CPUC can determine its applicability to informing codes and standards savings.

The evaluation team could use this data-driven approach to develop its own in-depth assessments of the codes and standards development process and outcomes, similar to the IOU Codes Change Savings Reports (CCSRs), but from a presumably more objective perspective. Utilization of such resources will not provide a complete picture of the market, but taken together, or used to triangulate different sources, can provide limits on NOMAD curves. Utilizing these other data sources over time can develop technology and market trends that can be refined to provide a better picture of market conditions.

This method would involve systematic data collection to reconstruct events leading to the outcome, including compiling, comparing, and weighing the merits of narratives of the same set of events provided by individuals with different points of view and interests in the outcome and then assessing the consistency of the observed fact pattern with linkages predicted by the program logic model. The evaluators would then compare their findings and the IOU Codes and Standards Statewide CASE team CCSR to consider a number of alternative causal hypotheses. Alternatively, or in addition to, the evaluators could also summarize this information for the Delphi panelists along with NOMAD curves instructing the panelists to examine the curves for consistency with the narrative fact patterns.

6.3.2 Delphi Panel Information and Instructions

The Delphi panel should use a data-driven approach to develop the NOMAD curves so that they can be validated and tested for reasonableness. As described previously, the current approach cannot be tested, therefore it is difficult to determine if the process can provide accurate results.

The Delphi approach is a structured, interactive technique for obtaining expert group inputs, usually to develop forecasts. Cadmus developed a modified Delphi approach using a flexible, web-based data-collection application to interact with identified experts. The experts used the tool to develop NOMAD curves and were encouraged to provide a rationale for their response. The intent was for Cadmus to facilitate iterative reviews over a number of rounds with the intention of reaching consensus or stability. Based on the results provided in previous evaluation reports, it was difficult for the C&S Program team and other reviewers to understand how the consensus and stability were achieved.

In addition to the approach summarized above, the, evaluators should provide all panelist with sufficient market data so that they all use common information to determine the adoption curves. There needs to be a clear and systematic approach for instructing the panelists on what market influences to consider, and as importantly, what not to consider. In addition, the evaluators should collect better documentation from the experts on the rationale behind their selections either through increased documentation by the experts or through follow up interviews. Using an energy-based NOMAD, panelists could aggregate multiple measures to see how that impacts the improvement of energy use. Alternatively, panelists could be provided with various usage curves representing various market adoption scenarios and be asked to comment on which is most likely and why. The panelist could be asked specifically to acknowledge, consider, and explain the impact of multiple influences.

6.3.3 Delphi Panel Recruitment

The robustness of the final NOMAD results is dependent on the strength and credibility of the Delphi panel. Recruitment of appropriate panelists is critical to a successful outcome. Depending on the nature of the product or markets addressed, it may be difficult to identify and recruit an appropriate expert panel. All panel members bring inherent biases and opinions that need to be made as explicit as possible to assist with the interpretation of results. In their evaluation plan, Cadmus developed a detailed plan for selecting Delphi panel experts. Cadmus laid out the process and criteria for selection and participation in the Delphi panel. However, the final evaluation reports did not provide sufficient information to explain whether the plan was successfully implemented and how any deviations from the plan impacted the outcome of the Delphi panel activity. For example, the 2014 evaluation report appendix C provides details of the approach including that Cadmus classified panel candidates by four organization type categories of. However, for some measures there were less than four Delphi panel members based on the number of curves on the NOMAD graphs. No explanation was provided in the report on how this might impact the final NOMAD result.

The CPUC Codes and Standards evaluation team needs to provide transparency throughout the process in the development of the Delphi panel. The TRC Team recommends that the evaluation plan includes updates to the IOU Statewide codes and standards team on the composition of the final Delphi panel(s) for review and input.

6.3.4 Alternative Approaches

In addition to the possible solutions we present based on our analysis, we provide here information on additional alternatives and approaches that can be incorporated in whole or in part to a revised NOMAD approach. In presenting the DOE approach, we acknowledge that it might not work under the existing CPUC baseline framework, but it may still provide useful insight into the complicated process of determining NOMAD.

The 2005 Heschong Mahone Group white paper (Mahone, Hall, Megdal, Keating, & Ridge, 2005) proposed a few alternative methods that were not adopted in the existing evaluation approach:

- Consider three scenarios of market adoption: low, medium, and high. This method may require more resources to estimate and would result in a range of possible net standard energy savings rather than a single value, as is accomplished with the Delphi panel Bass curve.
- Look at new construction program data to estimate how technologies and energy efficient measures are being voluntarily adopted.
- Estimate what the market saturation of a technology or measure was before code adoption to inform the likelihood of market adoption without a code (e.g., if a technology has captured a large portion of the market, there is a greater probability that the market would have standardized that level of efficiency without a code).
- Using the Delphi panel method and weighing responses from experts based on their level of knowledge. However, this could introduce an additional layer of subjectivity or bias into the NOMAD estimate.

Department of Energy (DOE) Appliance Standard Approach

The DOE assesses energy impact as part of technical analysis for federal appliance standards development. The assessment provides annual national energy savings, which are equivalent to California's net appliance standard savings. The overall approach of DOE's savings assessment is to compare annual national energy consumption by the appliance of interest with and without the proposed standard options in future years.

Energy consumption of an individual appliance is determined by its efficiency rating and operational parameters. For estimating annual national energy consumption, DOE considers possible combinations of efficiency rating and operational parameters for existing and new appliances. DOE first identifies possible levels of energy efficiency rating and values of operational parameters and then assess their market distribution among installed appliance based on market study results or assumptions. DOE then uses a Monte Carlo method to develop a sample of appliances with their efficiency rating and values of operational parameters reflecting market distribution. These appliance samples, along with their efficiency rating and values of operational parameters, are used to estimate the distribution of energy consumption for the entire stock of installed appliances.

DOE assesses the annual national energy consumption for the baseline market condition, under which no new standard is adopted, and different standard adoption scenarios. The difference in annual national energy consumption between the baseline and a standard adoption scenario for a future year represents the annual national energy savings of the corresponding standard option for that year. It should be noted that the baseline market condition includes a forecast of future adoption trend of high-efficiency options and, therefore, NOMAD is included in the baseline. In this way, the baseline represents a business-as-usual case, reflecting market trend without standard improvement. DOE does not explicitly consider compliance of new standards. However, for each adoption scenario, DOE forecasts future market adoption of different efficiency options and, therefore, effectively includes compliance forecast for future years.

7. Net Program Savings

Net program savings represent savings attributable to the C&S Program. CPUC evaluators assess IOUs' attribution based on the roles and efforts of the C&S Program in standard development and advocacy. IOUs undertake development and advocacy activities throughout the entire code development process, acting as leaders, coordinators, and subject matter experts during the rulemaking and long-term planning activities. The impact evaluators assess program attribution based on documentation provided by the IOUs and further investigation on how the IOU C&S Program activities have influenced codes or standards advancement. Net program savings is calculated as the product of program attribution and net standard savings.

The existing attribution evaluation method assesses C&S program attribution in three attribution areas and relative weighting of these three areas, the final score is the weighted average of attribution scores in these three areas. The three attribution areas are based on CEC's standard adoption criteria as summarized below:

Factor 1: Development of Compliance Determination Methods

- Factor 1 investigates the 1a) development of a reliable test methods for energy consumption, the assessment of existing test methods to use for standards and the development of reliable methods for estimating performance, and 1b) investigates the development of a method for estimating energy savings including algorithms for calculating those energy savings, and development of compliance software or modules capable of accurately analysis.

Factor 2: Development of Technical Information

- Factor 2 investigates 2a) the definition of the standard, drafting standard language and presenting ideas or recommendations on the standard, 2b) calculates the energy and peak demand savings by market studies, engineering studies, and energy modeling calculations, and 2c) determines the costs and cost-effectiveness via cost research and cost-effectiveness analysis.

Factor 3: Feasibility of Meeting the Standard

- Factor 3 discuss the feasibility of meeting the standard by 3a) documenting the market readiness of the standard, 3b) documenting that the standard does not impose unreasonable and avoidable costs to end users and 3c) document no significant health and environmental externalities.

Past codes and standards evaluations used the following steps of attribution assessment:

1. Review of Evidence. To establish a final score, the evaluation team collects data on stakeholder activities from a diverse range of sources, including rulemaking dockets, Code Change Savings Reports (CCSRs) (written by the IOUs), and stakeholder interviews. CCSRs (previously called code change theory reports) are tools used by IOU CASE authors in detailing the activities, communication, and role the IOU played to advance a specific code or standard. The program logic model, created by the IOUs, is provided to the impact evaluators as a way to visualize how the IOUs plan to influence both appliance and building standards development. The logic model identifies barriers to standard development, describes program activities developed to address those barriers, provides measurable activity output to allow program efforts to be measured, and provides short-term and long-term outcome of program activities.
2. Convene a Panel. The evaluation team then selects codes and standards experts to form a panel to perform attribution assessment. The panel in the CPUC 2013-2015 codes and standards evaluation was made up of four panelists of which three had participated in the past codes and standards evaluation.
3. Assign Factor Scores and Weights. The expert panel collectively determines the program's contributions score for each standard based on a review of the collected evidence.
    - **Factor Scores.** The factor scores are determined when the panel meets to discuss each code or standard. The impact evaluators explain to the panel the development of the standard, including

the prescriptive or performance requirements, the key stakeholders, and the history of the development of the standard. Then, the panel is presented with evidence about the C&S Program contributions within each attribution factor area. The panel members discuss their thoughts on the three factors for each code or standard and consider the inputs of all stakeholders including the C&S program.

- **Weight Scores.** After the factors scores are developed, impact evaluators derive the weighted scores initially from secondary research then ask the IOUs to provide factor weights for each standard. This was done during the 2013-2015 Impact Evaluation report. The impact evaluators base the factor weights on their understanding of how resources were allocated across the factor areas for each code or standard. This assessment is based on the data collected through the review of rulemaking documents and stakeholder interviews. If the weights are relatively close, the impact evaluators use the weights developed internally. If large discrepancies exist between the impact evaluators weights and the IOUs' weights (generally 10% or more), the impact evaluators review the justification provided by the IOUs, conduct additional research, and make adjustments to the weights as necessary.

7.1 Key Issues

Themes appeared during the key stakeholder interviews regarding issues with accuracy, streamlining and transparency. The calculation of the net programs savings should take into consideration the critical influence that IOUs have on the codes and standards adoption process and interviewees voiced concern over how the existing attribution evaluation method actually attributes those activities effectively.

ACCURACY	STREAMLINING	TRANSPARENCY
Lack of accuracy: • The entire spectrum of code advocacy activities may not be captured • The three existing attribution factors cannot adequately reflect types of effort needed for standard adoption	Opportunities to streamline: • Attribution evaluation method and process can be improved by connecting the dots between program logic models and attribution factors • There is a need to reduce the heavy burden on program documentation	Lack of transparency: • Into how the impact evaluators consider program advocacy activities in determination of attribution factor scores • Regarding how the panel of experts are chosen and who is chosen

7.1.1 Accuracy

Accuracy of net program savings calculation may be reduced due to the fact that the three attribution factors 1) compliance, 2) cost-effectiveness and 3) feasibility may not be able to capture the entire spectrum of code advocacy activities. Stakeholder engagement, outreach, and advocacy is a large majority of the work completed by the IOU advocacy teams and not explicitly reflected in the three attribution factors. codes and standards planning and coordination efforts that happen prior to a specific standard development, such as measure scoping, and other over-arching program activity do not fit into any attribution factor. Therefore, the existing attribution factors do not fully reflect the types of activity and effort needed to push a specific standard through adoption process.

Furthermore, evaluator interviews with the IOU codes and standards advocacy team typically occur several years after code development activities due to the separate timelines of code development and evaluation processes; therefore, interview responses may not be comprehensive or accurate. IOUs have tried to address this time lag issue with the code change savings report (CCSR) on-going, real-time documentation of activities. The IOU team indicated during interviews that they did not receive confirmation from the evaluators that the real-time documentation improved the process and information used for the analysis.

7.1.2 Streamlining

C&S program has been designed and implemented based on program logic models, which show what and how program activities are planned to address barriers to standard development and adoption. However, past CPUC codes and standards evaluations did not consider the program logic models. Therefore, there is disconnect between codes and standards attribution evaluation and codes and standards program design and operation. The former is based on consideration of program activities, while the latter can provide comprehensive understanding of codes and standards program activities. Program logic models also provide holistic views of barriers to overcome for successful standard adoption. There are opportunities to connect the dots better between the logic models and attribution factors. The logic model currently is too focused on the chronological order of advocacy and that description of advocacy through the different parts of the rule-making is only a tiny part of the overall effort that goes into how advocacy draws from information collected by the utilities and the market transformation activities that utilities and other groups have had on the market.

The IOU codes and standards team needs to prepare very detailed documentation of program activities, stakeholder information, and related background information. This process is very time consuming. While the documentation process has been improved, some interviewees still suggested that CCSR development used too much of the program resources and it was unclear how effective the CCSRs are in supporting attribution evaluation. While the use of CCSR's is an important tool to capture the process, opportunities for improvement exist automate the process and improve feedback loops between the codes and standards evaluation team and IOUs codes and standards team. Figure 13 illustrates the strengths and challenges with the current code change savings reports as reported by interviewees.

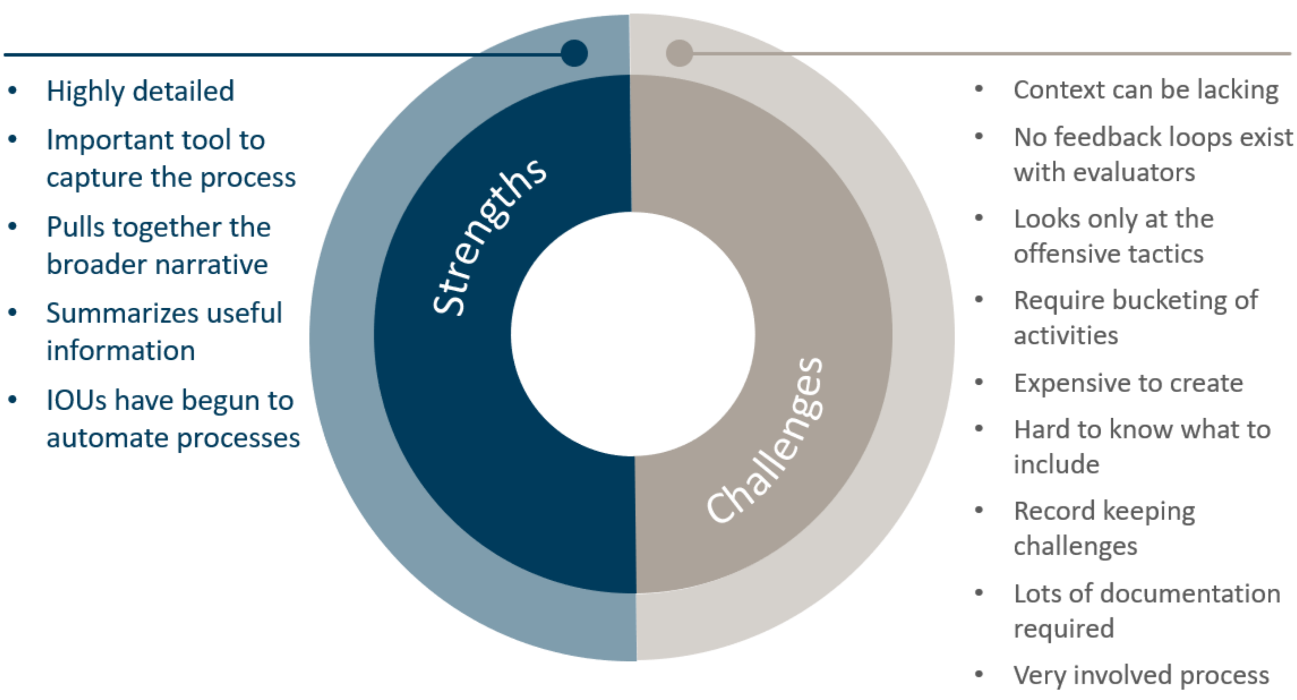


Figure 15: Code Change Savings Reports Strengths and Challenges

7.1.3 Transparency

Existing codes and standards attribution evaluation process is lack of transparency. Past evaluations didn't explain 1) what program activities were included in attribution analysis and 2) how program activities were considered to determine attribution scores. No guidelines were provided on how effectiveness of program activities was determined. It is unclear what types of program activities were considered to be more important than others, how different types of activities were rated in terms of relative importance for each attribution area, and how C&S Program advocacy efforts were evaluated against efforts by certain stakeholders who actively worked against increased regulations and standards adoption. Furthermore, there is limited information available on how the attribution panel selection process occurs.

7.2 Analysis

The TRC team investigated first how program activities are planned according to program logic model to achieve long-term success in standard development and then the connection between program activities and attribution factors.

7.2.1 Program Logic Models and Program Activities

It is critical that the attribution evaluation is based on an in-depth understanding of program design and activities. Program logic models provide comprehensive program design information and a program activity list. The C&S program has a logic model for the overall codes and standards program design and logic models for individual sub-programs, including for Title 20 appliance standards advocacy, federal appliance standards advocacy, building standards advocacy, and reach code advocacy, and code readiness sub-program. We analyzed the program activities based on program logic models and input from survey interviewees.

The overall C&S Program logic model shows that C&S Program is designed and implemented to facilitate long-term standard advancement through development, advocacy, and compliance intervention of individual standards. For Title 24 building standards development, the C&S program plans for the next two or three code cycles. For Title 20 appliance standards, C&S Program plans for potential standard topic for coming years. There are two distinct phases in which IOU C&S Program activities occur: rulemaking support and long-term planning, as shown in Figure 15. The rulemaking support phase includes activities for supporting a particular rulemaking; these activities could include CASE report development, stakeholder engagement and outreach, support to the CEC, etc. The CASE study activities are the most visible component of rulemaking support phase and often considered as the only activity of C&S Program. The existing attribution evaluation focuses primarily on this phase of codes and standards activities. The rulemaking phase is actually nested within the long-term development planning phase, as shown in 14. The IOUs' long-term development planning are continued efforts, even when there is no active rulemaking, to ensure long-term codes and standards development can successfully meet state energy policy goals. There is no indication that the C&S Program efforts in this phase have been captured, or can be captured by, the existing evaluation attribution factors.

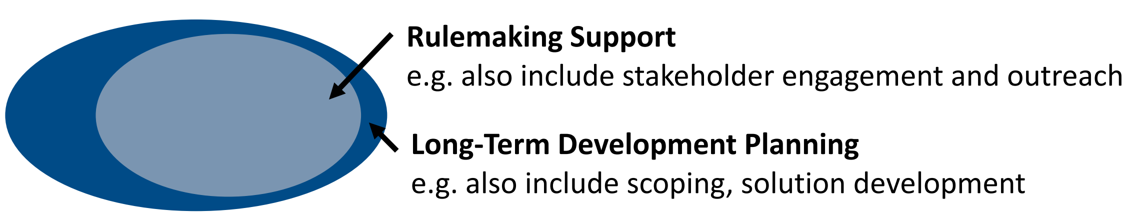


Figure 16: Program Efforts

To provide better understandings of the two phases of C&S Program efforts illustrated in Figure 16, we investigated the related program activities. We categorize detailed program activities presented in program logic model and provided by survey interviewees into several activity areas to provide a better charaterization of these activities. Figure 17 below provides a summary of these activity areas and how they are related to rulemaking support and long-term development planning. Figure 17 provides a high-level visualization of how different C&S Program activity areas are related to the two phase of C&S Program efforts. Appendix 9.5 contains more detail regarding program activity themes, description and example activities.

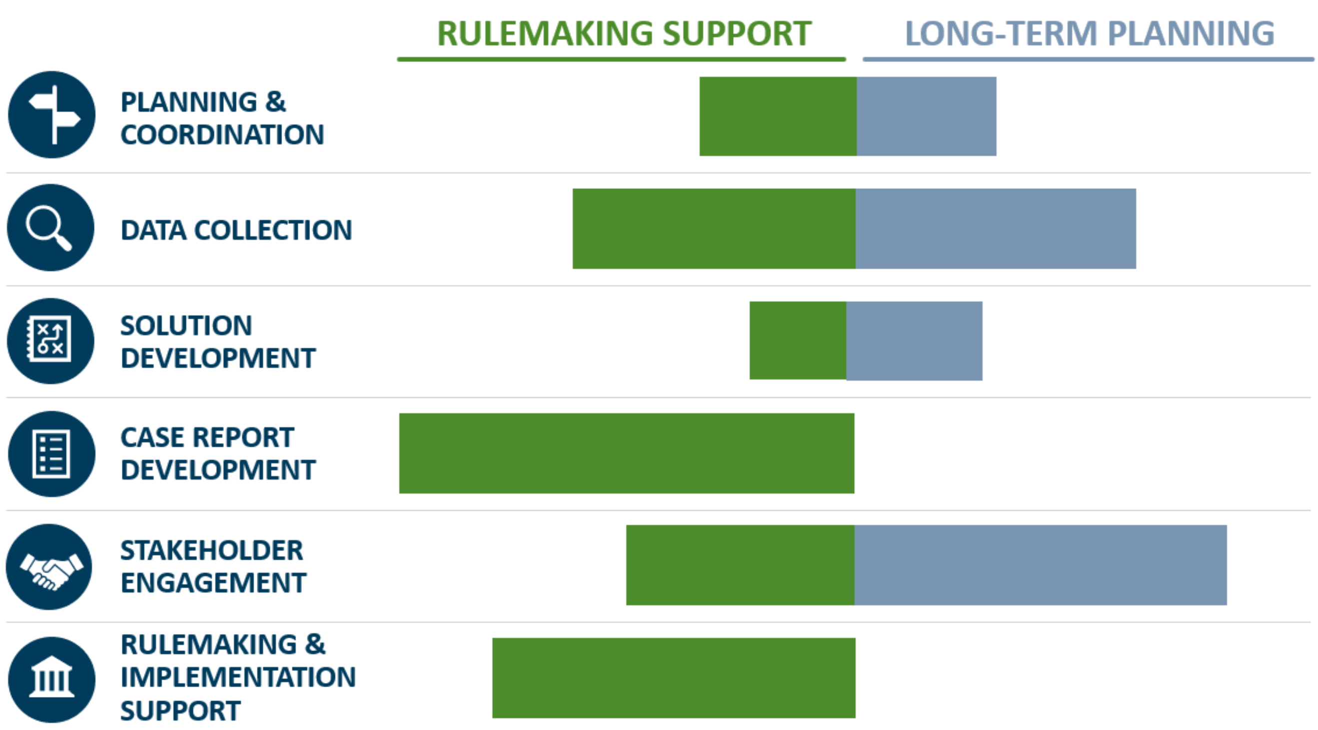


Figure 17: Rulemaking and Long-Term Planning Activity Themes

7.2.2 Connection Between Program Activities and Attribution Factors

As discussed earlier, the exiting attribution evaluation method utilizes a weight and scoring process to indicate relative level of effort spent in each attribution factor area and contribution from the codes and standards program activities. The model focuses on three areas of activity representing the fundamental requirements that must be met for the CEC to adopt a code. The factor weight indicates the relative effort required in each factor area. The factor score indicates the relative contribution of the C&S Program in the factor area. The model does not require collection of the actual total resources required. The CPUC evaluator suggested that one possible approach to score program efforts is to consider effort in terms of resources, such as labor hours or budgets.

We investigated if the activity themes described in the previous section have direct connection to the three attribution factors and, therefore, can receive attribution in these three attribution factor areas. The results are presented in Figure 17. From our perspective, there is a lack of factors available for those key activities that ensure IOU code advancement, including factors that could consider innovative ways the IOUs engaged in advocacy activities.

PROGRAM THEMES	ADOPTION REQUIREMENTS		
	Factor 1	Factor 2	Factor 3
PLANNING & COORDINATION	Mostly No	Mostly No	Mostly No
DATA COLLECTION	Some	Some	Some
SOLUTION DEVELOPMENT	Yes	Yes	Yes
CASE REPORT DEVELOPMENT	Yes	Yes	Yes
STAKEHOLDER ENGAGEMENT	Some	Some	Some
RULEMAKING & IMPLEMENTATION SUPPORT	Yes	Yes	Yes

Figure 18: Program Activity Themes and Adoption Requirements

7.3 Possible Improvements

Multiple stakeholders who we interviewed confirmed that most of the codes and standards improvement would not happen without the IOU participation and support. Currently, the attribution factors do not capture adequately the advocacy activities that the IOUs undergo to support the C&S program. In this section we provide information on possible improvements.

7.3.1 Develop Improved Methods to Determine Attribution Score in each Attribution Area

The method to determine the attribution scores could be improved to make the evaluation process more transparent. The analysis provided in the previous sections illustrates a potential list of program activity areas (see Appendix 9.5) that can be developed based on program logic models to represent the full spectrum of program activities reflecting how C&S program is designed and implemented. A more transparent attribution evaluation method can be constructed based on such a list of activity areas. Instead of only providing an attribution score for each attribution factor, codes and standards evaluators can provide attribution scores associated with each activities area and for each attribution factor, Figure 17 above, illustrates this.

In addition, evaluators can provide the following additions during the evaluation process to provide additional transparency to the evaluation process:

- Specific or highlighted activities considered within each activity area, and

This approach ensures all C&S Program activities, as planned in program logic model and presented in Code Change Savings Report, are considered by codes and standards evaluators or, at least, discloses what and how program activities are considered by codes and standards evaluators. The C&S program needs to ensure that program activity areas be developed based on updated program logic models. CCSRs need to document detailed program activities according to the category of activity areas.

7.3.2 Possible New Attribution Factors

The existing three attribution factors are based on the CEC's standard adoption criteria. Therefore, they reflect mostly technical requirements for standard adoption and specific rulemaking process. There are other requirements and pre-requisites for standard development and adoption, as listed below, even though they are not considered as official requirements and are not directly addressed the three attribution factors.

- Standard topics need to pass preliminary vetting through C&S Program interval planning efforts and coordination with CEC and relevant stakeholders in order to become a rulemaking candidate,
- General market data, e.g. plug loads in residential building, may need to be collected to support the development of a range of appliance and building standards,
- Creative regulation methods may need to be developed to enable the standard to be considered and analyzed, and

With the consideration of these additional requirements for standard adoptions, additional attribution factors may be added to fully capture C&S Program effects. The two possible additional attribution factors are:

- Strategic preparation: to capture planning, coordination, data collection, and solution development activities, and
- Stakeholder engagement.

7.3.3 Recommendations on Panel Selection Criteria

Interviews mentioned that lack of transparency into how the program attribution panel is selected that determines the attribution scores. As discussed earlier, the panel is convened in an in-person setting where the evaluators share documentation into how the IOUs underwent activities. The panel then discusses the evidence provide and allocates a score by each attribution factor. To improve the selection process, the evaluators need to make it clear how the panelists were selected, why they were selected by making this transparent to the IOUs. This will result in the IOUs being able to challenge the reasoning behind selecting specific panelists ensuring that a non-biased panel exists.

8. RECOMMENDATIONS

The project team identified many codes and standards evaluation issues based on interviews with a broad range of stakeholders. These issues reveal the complexity of the codes and standards impact and suggest that existing codes and standards evaluation methods are not adequate to accurately capture energy savings generated by IOUs' C&S Program. Resolving these issues may seem to lead to a more complex codes and standards attribution model and require more data to be collected for the impact analysis. The study revealed that these identified issues, excluding issues identified in the net program savings section of the model, are interconnected to each other and need to be addressed in a coherent manner. With the above considerations, the project team systematically analyzed issues to increase evaluation accuracy and streamline the evaluation process.

The CPUC codes and standards attribution model provides the framework of C&S Program impact evaluation; detailed evaluation methods for each model component determine how the codes and standards attribution model is implemented, including:

- selecting the baseline used to calculate potential savings,
- defining compliance and perform compliance adjustment to calculate gross savings,
- defining and assessing NOMAD to calculate net standard savings, and
- assessing codes and standards program activities to determine program attribution score.

These recommendations are directed towards the CPUC codes and standards evaluation team. The project team first investigated possible improvements to the codes and standards attribution model to ensure a robust evaluation framework is provided to guide the improvement of detailed evaluation methods. Then, the team analyzed evaluation issues related to each model component and developed improvement options according to the enhanced codes and standards attribution model. In addressing individual evaluation issues, the project team developed two overarching themes of improvement for model components:

- using energy-based impact assessment and
- collecting informative market data.

These themes allow different evaluation issues to be addressed coherently and, therefore, improve evaluation accuracy and streamline the evaluation process.

The following sections present recommendations for improving the codes and standards attribution model, overarching evaluation methods, and individual model components. The TRC Team developed these recommendations to provide comprehensive improvement to codes and standards evaluation and, therefore, we recommend they are adopted as a whole. However, because it might be challenging to implement these recommendations all at once, they may still be successfully selectively implemented, as discussed in the sections below.

This study did not directly address various savings double counting concerns related to building and appliance standards due to limitations of project scope. However, the recommended codes and standards attribution model enhancement and overarching evaluation methods provide a good framework to facilitate further development of solutions to resolve double counting issues.

8.1 Codes and Standards Attribution Model

The existing codes and standards attribution model includes factors, such as baseline, compliance rate, NOMAD, and program attribution, for determining impact assessment. The existing codes and standards attribution model is, in general, aligned with the savings calculation methods used by incentive programs and, therefore, allows easy integration between C&S Program and the rest of the energy efficiency portfolio. For this reason, the project team does not recommend any structural changes to the existing codes and standards attribution model. Without clearer definitions for each step, the existing codes and standards attribution model cannot provide adequate guidance on addressing complicated evaluation issues. The TRC project team recommends that definitions of the attribution model components be improved to clearly define the impact associated with each model component. The improved definition should allow the codes and standards evaluation team to accurately determine the energy savings attributable to IOUs, as well as codes and standards impact on California's energy demand, which is an important element for energy policy planning.

The TRC team recommends that the CPUC codes and standards impact evaluation team recognize the importance of assessing overall market changes. The codes and standards advocacy program is a resource program aiming to reduce energy demand through mandatory changes of the whole market, instead of selected population of the market, e.g. participants of incentive programs. Energy savings attributable to the C&S Program should be based on actual energy use reduction of the whole market. Past CPUC codes and standards impact evaluations focused on verifying energy savings of standards, not market changes. In theory, energy savings are the same as the corresponding energy use reduction of the market. In practice, considering energy impact from the perspective of market changes provides a better understanding of actual energy savings achieved by standards in a complicated market. Following this approach of considering overall market effects of standards, the TRC team recommends the following definitions of the codes and standards attribution model components:

- Potential savings are the standard's technical potential assuming idealized whole market change from baseline to the new standard;
- Gross savings represent actual market change, with consideration of how the average market condition deviated from the baseline before the new standard took effect and the how the average market condition deviates from the new standard after it took effect;
- Net standard savings correspond to the market change purely due to the standard, with exclusion of market changes introduced by all non-regulatory forces including those associated with voluntary standards and other utility efforts, including incentive programs; and
- Net program savings reflect the market changes attributable to utilities.

The net standard savings defined above is slightly different from the existing calculation method, which excludes market effects associated with utilities' incentive programs from NOMAD or, in other words, includes market effects of incentive programs as net standard savings. We recommend that impact associated with all utility efforts, including market effects of incentive programs, are considered in net programs savings calculation. With this revised definition, the net standard savings purely represents the net energy impact of standards and, therefore, can be used to inform procurement planning and net reduction of greenhouse gas emissions. Net program savings should include impact associated with all utility efforts, including both advocacy and non-advocacy efforts.

This enhanced codes and standards savings definitions was used to resolve specific codes and standards attribution issues. For example, in consideration of compliance adjustment method, we explained in Chapter 5 (illustrated in Figure 10) that the gross savings should reflect actual market change associated with transformation from the market average under the prior standard to the market average under the new standard. The former reflects average compliance of the prior standard and the latter reflects average compliance of the new standard. We provided an improved gross savings calculation method by incorporating compliance adjustment of the prior standard. Recommendations for other evaluation issues were also developed based on the improved definitions of codes and standards attribution model components.

8.2 Overarching Evaluation Method

Two themes of overarching improvement are developed based on enhancement of the codes and standards attribution model and analysis of evaluation issues related to each model components.

8.2.1 Perform Energy-based Assessment

The TRC team recommends that assessment of market change, as emphasized in codes and standards attribution model improvement, should be based on energy performance of the overall market, instead of market share of qualified installation, i.e. those meeting or exceeding the standard. The latter cannot capture energy impact variations associated with different efficiency levels.

Past codes and standards evaluation studies used energy-based evaluation for building compliance evaluation. Analysis in Chapter 5 showed that compliance evaluation based on compliance rate, which is market share of qualified installation, under-estimates compliance in terms of energy impact. Analysis in Chapter 6 indicated the NOMAD assessment has been based on market share of qualified installation and is incompatible with energy-based compliance assessment used for building standards. NOMAD based on market share of qualified installation cannot adequately reflect the market trend of energy use in absence of the new standard. Therefore, it is necessary that energy-based evaluation be consistently used for all standards and for all model components.

The TRC team further recommends that energy-based evaluation be performed by assessing market average UEC of related end-uses for installations covered by the standard. For example, evaluation of a Title 24 standard addressing LPD and controls in warehouse buildings would assess the reduction of average indoor lighting UEC in warehouse buildings; evaluation of Title 24 whole building improvement for nonresidential new construction building would assess the reduction of average whole building UEC among all nonresidential new construction buildings.

UEC has been used to determine UES, which is the UEC difference between the baseline and the new standard. Following the above recommendation, compliance evaluation would assess actual market average UEC and compare it to the UEC corresponding to the new standard. This compliance assessment method is the same as compliance margin approach described in Chapter 5, which indicated that compliance margin is a better indicator of compliance status than ESAF or CAF. Following the above recommendation, NOMAD evaluation would estimate the market average UEC trend in absence of the new standard. The difference between the baseline UEC and NOMAD UEC is the energy savings that would happen without the new standard and should be excluded from gross savings as NOMAD adjustment. Analysis in Chapter 6 indicates that the market average UEC trend in the absence of any related codes and standards development is a better representation of NOMAD for all related standards.

8.2.2 Collect Informative Market Data

The TRC team recommends that the CPUC codes and standards evaluation collect informative market data to support energy-based evaluation discussed in the prior section.

Past CPUC codes and standards evaluations spent significant resources on data collection and analysis, not all of which provide informative data on energy use characteristics of the market to support energy-based evaluation. Data collected to perform whole building compliance evaluation provides detailed building energy performance information and can be used to assess market average UEC under the new building standards. However, past codes and standards evaluations only used the collected data to estimate energy savings but did not provide UEC information. For appliance compliance evaluation, past codes and standards evaluations did not collect enough data to assess energy performance distribution among covered annual shipment to determine market average energy performance. NOMAD evaluation conducted by past codes and standards evaluations were based on input from industry experts, not market data.

To implement the above recommendation, the evaluators would continue to collect building performance data to develop average UEC, which would then be used to determine average compliance margin to inform market compliance status. Compliance evaluation for appliance standards needs to collect appliance shipment data for different performance levels. For NOMAD assessment, the codes and standards evaluation team needs to collect market trend data to estimate efficiency improvement trends to inform how market average UEC for related end-use areas has been reduced overtime. Market trend data needs to be collected from past studies, e.g. California Commercial End-Use Survey (CEUS), Residential Appliance Saturation Study (RASS), and California Commercial Saturation Survey. Data collected at the national level and for other states could be used as supplemental information. The CPUC evaluation team must anticipate that data gaps will exist when developing the completed market trend. However, the TRC team believes it is better to estimate NOMAD based on limited market data than not based on any market data. If future codes and standards evaluations collect market data to assess market average UEC on a periodic basis, there would be adequate data over time, along with data provided by other market studies, to provide more reliable market trend information.

8.3 Components of Codes and Standards Attribution Model

This section provides recommendations for each of the attribution model components. These recommendations address detailed evaluation issues following the framework of the enhanced codes and standards attribution model and overarching evaluation methods.

8.3.1 Potential Savings

Estimating baseline is important for determining potential savings. Current CPUC policy requires different approaches to determining baseline based on whether a prior standard exists. A prior baseline standard is straightforward. When a prior standard does not exist, an ISP baseline is required. There are complexities in selecting appropriate baseline options that carry through the rest of the attribution model. We propose two recommendations based on Chapter 4 analysis:

- When there is not a prior standard and ISP is used as the baseline per CPUC policy, ISP is defined as market average UEC. In addition, in net standard savings calculation, only incremental NOMAD, which is the additional NOMAD beyond that included in the market average, should be used for NOMAD adjustment.
- For Title 24 building standards with multiple sub-measures applied to multiple building spaces, UES should be defined as the average UEC reduction for related building types and annual installation should be defined as the construction floor area of related building types.

8.3.2 Gross Savings

Past codes and standards evaluation studies spent significant resources on compliance assessment, yet still ultimately had small sample sizes used to determine statewide compliance. For example, the evaluation approach used permit data for newly constructed projects permitted under the most recent code, which most likely biased the sample to smaller nonresidential buildings. The approach does not provide robust compliance results to meet the C&S Program attribution evaluation objectives, and therefore other alternative should be considered.

Additionally, over-compliance must be understood and addressed correctly. Code compliance of individual buildings can exceed minimum code requirements due to normal market conditions or other parameters that are not a direct result of the code requirements. Efficiency improvement occurs for various and multiple reasons depending on the product; not specifically to improve efficiency. Over-compliance at the building level should not be capped. Some over-compliance may represent normal market conditions.

The TRC team has the following recommendations based on previous analysis:

- Use compliance margin as the metric to determine compliance status of Title 24 standards and compliance adjustment can be performed using the method provided in Section 5.2.
- Incorporate compliance adjustment for compliance of prior standard, when it exists and used as baseline. This recommendation ensures that there is no savings overlap and gap between two consecutive standards.

- Include over-compliance in calculating average compliance, i.e. not cap measured compliance to exclude over-compliance. Some of the over-compliance might be associated with NOMAD and the corresponding savings would be removed through NOMAD adjustment. Also, by incorporating compliance adjustment for prior standard, savings associated with over-compliance of prior standard, would be excluded from savings of the new standard.

8.3.3 Net Standards Savings/NOMAD

Market conditions are intertwined with baseline and compliance and affect potential and gross savings as described above. Quantification of naturally occurring market influences is very challenging. Further, there are limited data sources for providing actual NOMAD information and the current evaluation approach lacks transparency on how the final NOMAD curves are derived. The approach looks at each measure in isolation, the "market" is too complex and ambiguous for this to be a useful approach. The resulting NOMAD curves are difficult to test for accuracy since the process and results are not replicable. Because the approach cannot be tested, and the resulting estimates cannot be systematically replicated it is difficult to determine if the process is sound and the final estimates are valid.

The TRC team has identified opportunities for improving the net standards savings evaluation approach by developing long-term energy-based NOMAD trends by technology areas to better reflect codes and standards impact over time. NOMAD should be reduction of energy, not adoption of a particular product or technology. Energy-based NOMAD indicates how UEC reduction changes over time for a technology or end-use area. Such a NOMAD curve can be used for different standards belonging to the same technology areas or serving the same end-use area.

The TRC team has the following recommendations based on previous analysis:

- Perform NOMAD based on market trend of energy use, i.e. UEC reduction in absence of the new standard, not based on adoption rate of installations meeting or exceeding the efficiency level determined by the standard.
- Estimate long-term market trend of UEC for specific end-use areas and whole building without any standard improvement as NOMAD for related standards.
- Do not exclude market effects caused by utility incentive programs in NOMAD assessment. These market efforts should be attributed to IOUs during net program savings calculation. Without excluding these market effects, net standard savings purely reflect energy use reduction due to the adoption of the new standard beyond the business-as-usual condition.

Evaluators should collect market data to develop efficiency improvement trends to inform NOMAD limits. Evaluators should utilize existing resources to collect more market data to develop efficiency improvement trends to inform NOMAD curves. Evaluators could base market adoption trends on a review of previous evaluation and market studies, planned market and potential studies as well as detailed energy efficiency program data available through the CPUC.

The TRC team has identified opportunities for improving the transparency of expert recruitment process. The robustness of the final NOMAD results is dependent on the strength and credibility of the Delphi panel. Recruitment of appropriate panelists is critical to a successful outcome. Depending on the nature of the product or markets addressed, it may be difficult to identify and recruit an appropriate expert panel. All panel members bring inherent biases and opinions that need to be made as explicit as possible to assist with the interpretation of results. The CPUC Codes and Standards evaluation team needs to provide transparency throughout the process in the development of the Delphi panel. The TRC Team recommends that the evaluation plan includes updates to the IOU Statewide codes and standards team on the composition of the final Delphi panel(s) for review and input.

8.3.4 Net Program Savings/Attribution

Within this step in the model, the critical component is capturing the spectrum of IOU advocacy activities into the current attribution factors. The evaluators approach to assessing program advocacy activities in determining attribution factor scores lacks transparency and accuracy. It is unclear how the attribution assessment might overlap with the net standards savings estimates. The attribution factors may not reflect the full range and types of activities needed to achieve adoption and therefore may be limiting. In prior evaluations, no guidelines were provided on how effectiveness of program activities were determined and how ultimately the factors were scored and weighted. Trade-offs do exist with employing a more accurate approach instead of a more streamlined approach. These recommendations posed below take a more concentrated and dedicated approach in ensuring accurate net program savings.

Therefore, the TRC team proposes a set of possible resolutions based on the analysis in Chapter 7. The party responsible for adopting and implementing these changes lies with the CPUC evaluators.

- Use C&S Program logic model to establish a list of program activity areas and determine attribution score associated with each activity area. Program attribution evaluation needs to perform detailed consideration of codes and standards program activities and this recommendation simply provides a better organized evaluation process without requiring additional data collection or analysis. This recommendation enables the attribution evaluation to be transparent and providing effective feedbacks to codes and standards program staff.
- Use additional attribution factors or develop clear evaluation method using the existing attribution factors to determine attribution for the following two important program activity areas:
  1. Strategic preparation, including short-term and long-term standard development planning, coordination, data collection, and solution development, and
  2. Stakeholder engagement and advocacy.
- Recruitment of panelists for program attribution evaluation should be conducted through a transparent process including using selection criteria vetted by the CEC, IOUs, and other key stakeholders involved in standard development and adoption.
- Include market effects introduced by incentive and codes and standards programs in calculating net program savings.

9. Appendix

9.1 CPUC Evaluation Approach to Establishing Statewide Compliance Rate

9.1.1 Sampling Strategy

The CPUC evaluation team used the following statewide sampling strategy:

- The sample population will include only the customers in electric and gas IOUs' service territories.
- The sample targets will be determined based on:
  - Recent volume of new construction or alteration activity/valuation;
  - Level of significance of contribution to overall savings at measure-level
  - Level of significance of contribution to overall savings by building type,
  - Sufficient representation of buildings in various California climate zones (or climate regions).

The CPUC evaluation sampling strategy used a multiple stage approach to establish a statewide compliance rate:

- **Stage 1.** Sample points will be proportionally divided among a number of climate regions that cover the entire state. The target sample points will be allocated based on the reported total recent square feet of newly-constructed space and valuation of alterations of existing buildings within each region that occurred within the IOU territories.
- **Stage 2.** Within each climate region, municipalities will be randomly selected and ordered using a proportional-to-size method. The jurisdictions will then be contacted for recruitment based on their position in the list.
- **Stage 3.** Once recruited, a staff member will contact each building department to develop a population list of recently issued permits under the 2013 Title 24 energy code. Once the jurisdiction's permit list has been developed, the list will be cleansed based on the relevance of specific measures being upgraded for alteration projects. All sites will be given an equal chance to be selected (unless too small to produce tangible energy savings). Projects will be randomly selected for recruitment until the sample targets are achieved with sufficient back-ups.

This phased approach using proportional-to-size sampling was intended to provide an efficient sample set that covers the entire geographic (and climatic) areas of the state. It is assumed that a statistically significant number of data points will occur for the most frequently altered/constructed buildings.

Sampling Challenges

Cadmus initially identified a sample of 42 jurisdictions across the three climate regions. However, the sampled jurisdictions did not contain enough sites that met the criteria of the evaluation to reach the target number of site visits. To supplement the original sample, Cadmus added 24 additional jurisdictions, but once again fell short of the target number of buildings. As described earlier in this section, we additionally contacted building contractors for sites located in the sampled jurisdictions to obtain a list of sites outside of the sampled jurisdictions they had built that met the criteria required to be included in this evaluation.

9.1.2 Recruitment Process

Building permits are issued by local plan review jurisdictions for authorization to construct a new building or make alterations to an existing building. We will leverage our experience from the past Title 24 impact evaluation study to request commercial building permit data from July 1, 2014, for the California jurisdictions within the sampling frame. We will request data for more than the sampled jurisdictions to have replacements available when a jurisdiction is unable to provide permit data. Building permit records typically consist of the following information:

- Permit type and number;
- Building address;
- Owner's name and contact information;
- Project valuation;
- Work/job description or category; and
- Plan submittal and/or issue date.

After compiling the data provided by jurisdictions, we will screen them for eligibility to participate in the study, mainly to ensure that they have been permitted under the 2013 Title 24 energy code and include the significant standards in the case of alteration projects. We will then use the records to recruit potentially eligible participants for conducting the site visits.

The Cadmus team will leverage the unique recruiting process developed and tailored during the past study, which was specific to the study criteria and constraints on the availability of the building department permit data. We will conduct on-site data collection for recruited eligible sites that agree to participate in this study. The field auditors will conduct additional research and provide field data, photos, and any available as-built architectural drawings and Title 24 documentation to the analysis team.

This process includes the following recruitment steps:

- **Request nonresidential building permit data.** Obtain records from plan review jurisdictions within the sampling frame.
- **Conduct building permit data research and eligibility screen.** Review and clean data, conduct eligibility screen of data, and finalize potentially eligible permit records for recruitment.
- **Recruit and schedule eligible sites for visits**. Screen the sites for eligibility regarding measure type and project type over the phone and schedule audit appointments. Track recruiting progress against goals for each climate region, jurisdiction, and measure.

Nonresidential Recruitment Approach and Outcome

The evaluation team completed total of 66 surveys (49 new construction and 17 lighting alteration sites).

- Request Commercial Building Permit Data: 153,000 permit records received
- Conduct Building Permit Data Research & Eligibility Screen: 4,895 potentially eligible permit records identified
- Recruit Eligible Sites: 93 sites were recruited
- Schedule and Conduct Field Surveys: 83 field surveys were scheduled and completed
- Conduct Analysis of Completed Field Data from Audits: 66 sites confirmed by analysis

9.1.3 On-Site Data Collection

The objectives of our field data collection include the following: (1) perform rigorous data collection based on the specifications of the critical measures covered by 2013 Title 24, (2) inform the analysis by incorporating all building parameters and characteristics that impact the savings associated with those measures in a measurable way, and (3) provide insights to improve savings estimates.

The Cadmus team will deploy a range of methods and tools to achieve these objectives through a consistent, integrated, and transparent approach. We will determine the appropriate measurement and verification (M&V) methods for each project and measure type by performing an in-depth review of the code compliance requirements and the scale of savings estimated by the IOUs. From the code review, we identified the building and measure parameters that affect the compliance of a particular measure with the code; this will inform the development of a custom data-collection form for this study.

Research of Building Department Records

The Cadmus team will research all available documents kept by the building department related to the plan review and permitting process for each surveyed site. The documentation will include, but is not limited to:

- Architectural, electrical, and mechanical drawings
- Construction details and specification books
- Title 24 documentation (envelope, lighting and mechanical)

Site Measurement and Verification

The Cadmus team will conduct site visits to physically verify the building's parameters and characteristics for new construction and alteration commercial project types. The data collected in the field will inform the input variables that are specified in the whole building energy modeling or engineering analysis on a per-site basis.

9.2 NOMAD Research Methodology

Our approach for estimating both initial market penetration and naturally-occurring market adoption was to solicit expert opinions to estimate a market adoption curve, using a typical S-shaped curve. We chose the Bass model to approximate the process.

We used a unique approach to obtain expert judgment about the market adoption curve. An interactive web-based tool was created that allowed experts to use sliders moved by their mouse to input their selection of leading and following behavior and maximum market penetration parameters. They were able to view the adoption curve in real time and make adjustments until they were satisfied with its shape. The web tool was interactive in real time as it enabled the following:

1) Allowed the experts to see the shape of the diffusion into the market over time: As the expert moves the sliders that determine the values of the three needed parameters, (leading behavior, following behavior, and maximum market penetration), the curve starts to take shape on the screen.

2) Provided the expert with a verbal description of the selected shape: Once the expert decides the shape looks right, he or she clicks a selection button and a verbal description of the curve is displayed. For example, the description might be that the selected curve implies that the market penetration will never exceed 50%, current market penetration is at 20%, and the market is expected to take off in three years.

When the expert is satisfied with the shape and the verbal translation, the selection is submitted, and the data are saved. All opinions are aggregated to produce the average S curve using a mathematical procedure that best fits the average values at each point. A second round allowed the participating experts to review the average curve and provide feedback on the differences between their curve and the average. Based on this feedback and follow-ups with the participants, we finalized the naturally occurring market adoption curve.

9.3 Literature Review: California Codes & Standards Program Advocacy and Attribution Study

9.3.1 Introduction and Literature Review Objectives

The first three study objectives listed in the scope of work and research plan involve analysis of the existing evaluation process. Our initial data collection approach is to review literature associated with the IOU Codes and Standards Program advocacy and Attribution Model. The literature review provides information on the history and evolution of the Attribution Model and explores alternative approaches to model components and processes. The findings will inform and direct further investigation through interviews with key stakeholders. As part of the literature review, we document changes that have been made to the Attribution Model to improve transparency, accuracy, and streamlining. Additionally, we researched the Bass diffusion model in the context of building and appliance energy measures and innovations, as well as investigated alternative approaches.

The literature review accomplishes the following goals:

- Document the history and evolution of the codes and standards advocacy process and attribution model
- Informing prioritization of Attribution Model research scope

To identify where improvements have already occurred, or potential modifications could address study objectives to improve accuracy, transparency, and streamlining, we have used the following symbols:

A **Improve accuracy[1]:** Improved confidence in correctness of estimated savings

S **Streamlined model or process:** Too many steps in the process causing redundancy in the model resulting in confusion and longer than necessary timelines

T **Improve transparency:** Ability to see and comprehend how the estimated savings were determined

9.3.2 History of Codes and Standards Program Advocacy and Attribution Model

We investigated the evolution of the C&S Program from its start in 2004 until the most recent evaluation report in 2017. While improvements occurred throughout the fourteen years, the most significant contributions to the current attribution model used to date occurred in 2009. Evaluators continue to align with the core CPUC evaluation protocols established in 2006. Figure 1, below, illustrates milestone activities along the C&S Program Advocacy continuum.

[1] Accuracy is "the condition or quality of being true, correct, or exact; freedom from error or defect; precision or exactness; correctness."

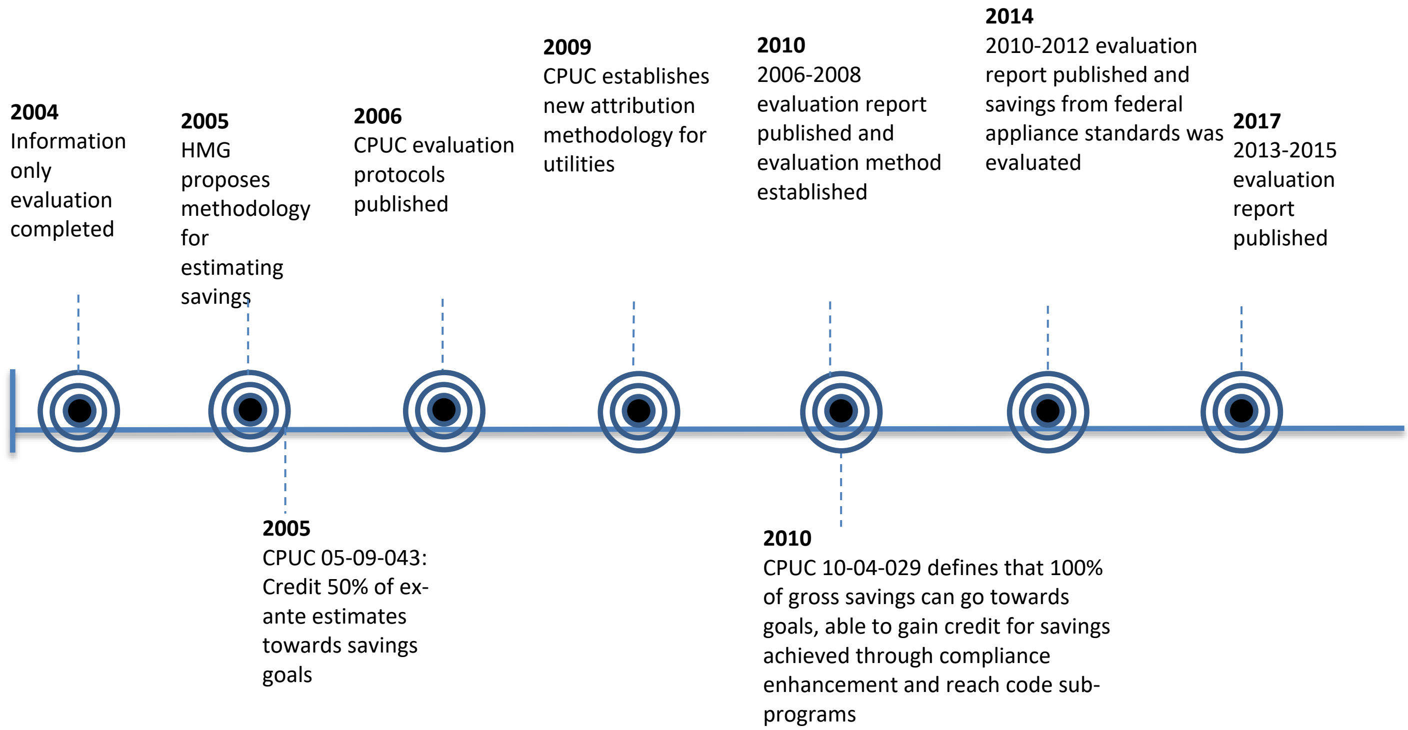


Figure 19: C&S Program Advocacy and Attribution Model Milestone Timeline

Prior to 2004, California Public Utility Commission (CPUC) considered the Statewide IOU C&S Program a non-resource program, meaning it did not contribute to IOU savings claims, nor was it required to go through a rigorous evaluation to determine savings impacts. Rather, the evaluation process was information based only and in 2004, an information only evaluation was completed on the 2002 Statewide IOU C&S Program. This evaluation report also contained the first suggestions for how to attribute savings to IOUs based on specific adoption factors.

In 2005, with the Statewide IOU C&S Program joining the class of resource-based programs, Heschong Mahone Group (HMG) proposed an attribution methodology to estimate statewide codes and standards savings resulting from the Statewide IOU C&S Program, as well as how to allocate those savings amongst the IOUs (Mahone, Hall, Megdal, Keating, & Ridge, 2005). In the same year, CPUC filed a decision (D-05-09-043) that credits 50% of ex-ante codes and standards savings estimates towards IOU savings goals (CPUC, 2005).

In 2006, the CPUC published the California Energy Efficiency Evaluation Protocols that contain guidelines for IOU C&S Program impact evaluations and attribution; however, these protocols are not mandatory and explicitly state that evaluators should modify as appropriate (The TecMarket Works Team, 2006, p. 82). The protocol steps include:

- Identify the codes and standards that the evaluation will cover
- Conduct a codes and standards gross-market level energy impact assessment, which includes techniques such as simple engineering model estimation approaches and normalized annual consumptions approaches
- Estimate the IOU C&S Program's influence on the adoption of codes and standards
    - The suggested process was through stakeholder interviews
- Estimate net IOU C&S Program induced energy impact

  - This step introduced normal market adoption (NOMAD), normal standard adoption (NOSAD), Non-Compliance Adjustment, Actual construction retrofit true-up, multiple counting of energy savings adjustments, and measure life adjustments

Once the protocols were established and the IOU C&S Program and evaluators were able to review and understand the potential implications of the protocols, consideration of compliance and compliance improvement efforts were brought to the CPUC. In 2007, CPUC included guidelines for adjusting savings based on compliance rate and allowed for IOU compliance improvement efforts to be included in the evaluation process and Attribution Model (Decision: 07-10-032 , 2006).

Guided by the CPUC Energy Efficiency Evaluation Protocols and having reviewed the HMG proposed methodology, Cadmus, who was the CPUC third-party evaluator at the time, published a proposed evaluation methodology in 2009 prior to evaluating the 2006-2008 code cycle (The Cadmus Group, Inc. & Quantec, 2009). This methodology set the base for future evaluations, with minor modifications occurring each iteration. The proposed methodology included an adjustment to reduce double counting of savings ●, introduced a Cadmus developed tool the Integrated Standards Savings Model (ISSM) tool ■, and incorporated attribution factors and weights into attribution score calculations ● ■. The Cadmus evaluation method established the following:

- Defined potential energy savings as the magnitude of energy savings that would result from every appliance or building standard measure covered by new standards
- Established the use of whole-building energy savings and compliance approach ●
- Eliminated NOSAD ■ ▼

Cadmus completed the first code cycle evaluation in 2010 for the 2006-2008 codes and standards cycle using the Attribution Model and their proposed evaluation methodology (KEMA, Inc.; The Cadmus Group, Inc.; Itron, Inc.; Nexus Market Research, Inc.; ENRG, Inc., 2010)[Also in 2010, the CPUC redefined in D-10-04-029 that 100% of gross savings can go towards IOU savings goals and that IOUs can gain credit for savings achieved through compliance enhancement and reach codes ● (Decision 10-04-029, 2010)].

Finally, in 2014 and 2017, two code cycles of evaluations were published using the Attribution Model and the Cadmus evaluation methodology, with modifications, such as adjusting for IOU Incentive Programs in NOMAD.

9.3.3 Attribution Model Investigation

Through an in-depth analysis of the literature, TRC has identified specific evolution and modifications to Attribution Model components and the methodologies used to implement the model. The following subsections address each main Attribution Model component and the subcomponents that are necessary for establishing estimates of this main component. The investigation considers alternative approaches that have been previously explored during the evaluation process and those that are cited in literature and could have possible application within the Attribution Model and evaluation framework.

Potential Savings: Energy Baseline, Unit Energy Savings, Market Baseline

Potential savings incorporates estimates of the energy baseline, UES (i.e. per measure savings), and the market baseline. To determine UES, the evaluation team collects field data and generates energy simulations for Title 24 measures and uses field data collection and rated energy performance for Title 20 appliance standards. Based on a CPUC decision, the energy and market baseline are set at the previous code or standard, if one exists, or the ISP if one does not exist. This becomes complex when dealing with alterations and retrofits to buildings that may have been built prior to energy codes, considering older appliances in existing buildings, and acknowledging that compliance with the previous code likely had a distribution above and below 100%. For a previous evaluation plan, Cadmus introduced the idea of establishing the energy and market baselines based on previous evaluation cycle findings for compliance rates ● (The CADMUS GROUP, Inc., 2012). This has not been implemented in an evaluation and is not in line with the CPUC decision.

Methods and opportunities to streamline the energy savings data collection have been considered in previous evaluation cycles, such as referencing Codes and Standards Enhancement (CASE) Reports UES, sales forecasts, and building forecasts. However, the CASE Report data has not been successfully integrated into the evaluation because savings are on a per measure level and do not include interactive effects between measures, such as done in the whole building evaluation approach. Additionally, CASE report sales and building forecasts, which are estimated several years before the code or standard takes effect, as well as other data sources, are often outdated by the time the evaluation takes place (DNV GL & CADMUS, 2014). The evaluation does use the CEC building forecast data for T24 savings because there would be too large of a time lag if actual construction were used. A previous evaluation team has considered adjusting prior evaluation results based on actual construction during the next evaluation cycle ●, but this modification has not been implemented (CADMUS, 2017).

To improve the accuracy ● of the market baseline and potential savings estimates, Cadmus considered an alternative approach to estimate savings based on pre- and post-adoption average savings; however, it was noted that sufficient data to support this analysis is not readily available and may require more data collection than is currently conducted (CADMUS, 2017; Antonopoulos, 2013).

Gross Savings: Compliance Rate

To estimate Gross Savings, the Attribution Model incorporates Compliance Rate to reduce Potential Savings for buildings or appliances that do not meet the new code or standard. Compliance rate is one of the most challenging components to estimate and there have been different perspectives on how to approach the definition of compliance. The current method to estimate compliance rate for Title 24 is at the whole building level, which allows for assessment of projects using the performance method; a measure-level approach would only be appropriate for projects using the prescriptive approach. The evaluation team has assessed whole building compliance through field data collection and energy models, which are then compared to a code baseline building. For Title 20, appliances observed in the field (e.g. retail stores) are compared to the CEC appliance database.

Compliance can be considered one of two ways: 1) binary, where a building or appliance either complies or does not (this perspective does not award savings for incompliance and does not consider the magnitude of under- or over-compliance), and 2) sliding scale, where a range of efficiencies, and therefore compliance levels, can exist (this perspective considers under- and over-compliance and adjusts savings accordingly). More recently, CPUC has stated preference for a hybrid approach which caps compliance at 100%. With this approach, building performance and appliances cannot achieve more than 100% compliance for exceeding the code baseline, but less than 100% compliance can be reported.

Although whole building compliance is used to estimate the compliance rate for Title 24, Cadmus investigated measure level compliance and energy savings during the 2013-2015 Impact Evaluation in response to stakeholder interest for measure level impacts. To determine measure level energy savings Cadmus used field data-informed energy simulations and adjusted individual measures to code baseline, while keeping all other building measures static ▼ (CADMUS & DNV GL, 2017).

Net Standard Savings: NOMAD

The current method to estimate NOMAD is implementing the Bass Diffusion model through Delphi panels using the Cadmus Market Adoption Tool. In this approach, a panel of experts individually develop Bass curves that graphically represent how the market would have adopted the code or standard level of efficiency without the codes and standards process. This is done through multiple rounds, with each round feedback to each respondent from the evaluation team. The Cadmus tool has several inputs for describing the shape of expected market adoption (Antonopoulos, 2013). The tool provides visualization to how these factors impacted the shape of the adoption curve. This method is straightforward and simple, which streamlines the process ■, but it may not be completely appropriate or accurate for measures that allow for a range of efficiency, nor does it document what factors experts are considering in their responses. For example, experts are not told to either include or exclude incentive program influences. For the 2010-2012 Impact Evaluation, CPUC acknowledged that

IOU incentive programs can have significant influence on the market and this should be excluded from NOMAD (DNV GL & CADMUS, 2014). Cadmus adjusted NOMAD after the Delphi panel process for IOU incentive program influence; therefore, crediting the IOUs for pre-code market influence, but it is unknown if this market influence was already excluded by experts. Additionally, there were inconsistencies in applying the Delphi panel method within and across the evaluation years; for instance, expert panels consist of different number of individuals from measure to measure, meaning that individual responses have greater weight in smaller panels.

In the 2005 HMG white paper, a few alternative methods were proposed:

1. Consider three scenarios of market adoption: low, medium, and high. This method may require more resources to estimate and would result in a range of possible net standard energy savings rather than a single value, as is accomplished with the Delphi panel Bass curve.
2. Look at new construction program data to estimate how technologies and energy efficient measures are being voluntarily adopted.
3. Estimate what the market saturation of a technology or measure was before code adoption to inform the likelihood of market adoption without a code (e.g., if a technology has captured a large portion of the market, there is a greater probability that the market would have standardized that level of efficiency without a code).
4. Using the Delphi panel method and weighing responses from experts based on their level of knowledge. However, this could introduce an additional layer of subjectivity or bias into the NOMAD estimate.

In the 2010-2012 Draft Evaluation Plan, Cadmus acknowledges the possibility that codes and standards may actually impact the market by altering the shape of efficiency distribution rather than just eliminating the lower tail of products or building practices. Although an understood market dynamic, it is challenging to assess a shifting market, whereas the Delphi panel approach attempts to take a snapshot of market adoption at a time prior to code adoption (The CADMUS GROUP, Inc., 2012).

TRC reviewed Diffusion of Innovation and Bass model literature and conducted an initial investigation of the Cadmus Market Adoption Tool to understand the impact of different inputs on the shape of the Bass curve. The most important inputs are the starting and end (maximum) market adoption and the leading adoption behavior. The literature is sparse on examples or assessment of applying a Bass model or other diffusion of innovation theories for energy efficiency measures, especially for buildings. In general, most of the literature either explicitly or implicitly indicated that the Bass model is not ideal for non-binary choices (such as efficiency measure options) but is still the best option available. From Diffusion of Energy Efficient Technology in Commercial Buildings: An Analysis of the Commercial Building Partnerships Program study, the assessment concluded that the Bass model may be more appropriate for whole building energy efficiency rather than looking at individual measures or combinations of measures which becomes complex (Antonopoulos, 2013).

Net Program Savings: Attribution

In the first C&S Program Evaluation from 2002 by ADM Associates, 100% of savings for each code revision was attributed to the primary funding organization, which was either CEC or the IOUs for this code cycle. At this time, the IOU codes and standards Program was a non-resource based program, so attribution was not considered as rigorously. This simplistic method worked at the time when code development was less collaborative, but in the context of the current code development and advocacy process, CEC stated that this method would ignore significant work and influence from a variety of stakeholders, including CEC, IOUs and industry representatives.

ADM Associates considered Decision Theory Approach, which is a statistical approach that identifies factors that are associated with a decision, how a change in various factors will affect the decision, and the overall predictive power of the statistical equation that links the decision indicator to the explanatory factors. Using this method, ADM Associates proposed analyzing two scenarios: 1) the probability of adoption if a code has support from the

IOU C&S Program and 2) the probability of adoption if a code does not. In the end, this approach was not pursued because it is typically used as an *ex ante* approach, whereas the evaluation is *ex post*.

For the 2006-2008 evaluation and onward, Cadmus used a method that attempted to systematically assess attribution, which is inherently subjective. The method used an attribution score calculated from three attribution factors: 1) the development of compliance determination methods and other special analytic techniques; 2) the development of code language and technical, scientific, and economic information in support of the standard; and 3) demonstrating the feasibility or market acceptance of code adoption. To assess the attribution factors and final score, the evaluation team reviewed documentation from the IOU CASE team that details their advocacy efforts and conducted interviews with key stakeholders to corroborate or expand on the CASE team documentation. Interviews typically occurred several years after code development activities due to the separate timelines of code development and evaluation processes; therefore, interview responses may not be comprehensive or accurate. This process can be time consuming for both the CASE team, which produces a separate report documenting their activities and accomplishments, and for the evaluation team to conduct several interviews. In past evaluations, the evaluation team focused on assessing attribution through interviews only for measures which produced the majority of energy savings, and then relying on documentation for the remaining measures.

The HMG white paper pointed out that deemed attribution, which attributes a predetermined portion of all codes and standards savings, is likely too simplistic and it is impossible to tell if the overestimations and underestimations weigh each other out. The white paper recommended a similar approach as the one which Cadmus eventually implemented, but the HMG method included five attribution factors, two more than Cadmus (Mahone, Hall, Megdal, Keating, & Ridge, 2005). The proposed factors are slightly different in nature between the HMG white paper and the Cadmus methodology; the intent of four out of five of HMG's factors are captured in Cadmus' factors; the one factor not included is to provide credit for innovativeness. This factor considers the likelihood that CEC would have identified or pursued a measure without promotion from the IOUs. This factor was not included in the final methodology because innovativeness of a standard is not criteria considered for code adoption in the CEC process.

Savings by Utility: Allocation

The proposed method to allocate energy savings by utility territory has remained unchanged and is straightforward. For federal appliance standards under the evaluation, the allocated California savings are proportional to California's population to that of the United States' (i.e., 12%).

9.3.4 Logic Model Alignment

TRC reviewed the C&S Program theory and logic model to summarize how program activities were planned to support California and federal standards development, and whether the logic model accurately represents the activities and process. The purpose of the program theory review is to provide a succinct refresher of the program process conceptualization. The program theory describes:

- Explicit and implicit assumptions made by program stakeholders about the actions required to achieve the program goals, and
- How these actions lead to specific outcomes that result in the program accomplishing the desirable goals

Based on initial review and comparison of known IOU C&S Program activities and the logic model, we have identified the following potential modifications where the logic model may not appropriately capture activities:

- Missing activities from Pre-Docketing and T-24 Draft CASE Study Development:
    - Support to voluntary standards (e.g. ASHRAE and Energy Star®) that lead to and facilitate market readiness for mandatory standards
    - Funded research projects that lead or support standards development, including:

    - ZNE characterization studies
    - ZNE road-mapping studies
    - Opportunities and barriers studies on specific market sectors, such as multifamily
    - Technology studies
- Missing activities from Post-Adoption and T-24 Code Implementation Support:
    - Support to program staff for code-readiness implementation, including workshops, presentations, and developing program-specific code change summaries
- Missing Outputs:
    - Analysis results such as cost effectiveness and energy model simulations that are used by organizations other than CEC for code adoption, such as CBIA and product manufacturers
- Missing Intermediate Outcomes:
    - Drive further efficiency in programs from improved baseline

9.3.5 Comparison to Other Codes and Standards Program Evaluations

Most states adopt and enforce versions of the IECC and ASHRAE 90.1 national building standards as well as incorporating the Department of Energy's national appliances standards. Few states have developed and adopted their own building codes, and even fewer have adopted their own appliance standards. Arizona, California, Oregon and Washington are among exceptions to states that have adopted their own appliance and building codes.

After developing and adopting standards, there is no agreed upon methodology for evaluating energy savings impacts from this advocacy. As stated in *Energy Efficiency Policy in the United States: Overview and Trends at Different Levels of Governments*, "at state and local levels, policies are evaluated to the extent that jurisdictions have the interest and financial capability to do so. These evaluations are generally focused on the ancillary benefits of energy savings, e.g., job creation, electricity prices, and environmental impacts. Typically, modeled energy savings (as opposed to actual savings) are the basis for these evaluations" (Doris, Cochran, & Vorum, 2009).

The organizations that develop and support code adoption can vary by state or jurisdiction, which can influence the level of rigor necessary for assessing energy savings. For example, in California, CPUC mandates that IOUs participate in developing and supporting codes and standards as part of their portfolios and are held to strict cost-effectiveness criteria. However, in Oregon and Washington, the Northwest Energy Efficiency Alliance (NEEA) participates in codes and standards but is not mandated to do so by any regulating body. NEEA must still report their codes and standards impact to their regional funders on impacts from activities.

Figure 20 and the following descriptions provide an overview of the method used in other states that conduct codes and standards impact evaluations or assessments.

	Attribution Model Component				
	Potential Savings	**Gross Savings**	**Net Standards Savings**	**Net Program Savings**	**Savings by Utility**
Arizona: AZ Corporation Commission	Yes	Yes, verified savings	Unknown	Deemed savings: up to 1/3 of as-built verified savings from codes and standards	Unknown
Arizona: Salt River Project	Yes	Yes, verified savings	Unknown	Deemed savings: up to 50% of as-built verified savings from codes and standards using point system	NA
BC Hydro	Yes	Unknown	Unknown	Deemed: up to 1/3 of total portfolio savings	NA
NEEA: OR, WA, ID, MT (Appliance Standards)	Yes (from DOE)	Yes (from DOE)	Yes (from DOE)	Based on weighting factors for market barriers and activity effectiveness	Based on participation in activities
NEEA Commercial Code Evaluation Pilot Study	Yes	Yes, verified compliance by end use/building system through field studies and assessed energy bills	Unknown	No (not part of the scope)	No
National Model for Building Energy Codes, PNNL	Yes	Yes (realization rate): 80% to 100% over 10 years for residential; 50% to 80% over 10 years for commercial	Rolling baseline (previous code)	All code-to-code savings	No
DOE Appliance Standard Approach	Yes	Yes	Yes	N/A	N/A

Figure 20: Comparing Codes and Standards Impact Evaluation Methodologies Across Other Jurisdictions

Arizona

Arizona is home rule, which means that each jurisdiction adopts and enforces energy codes. There are at least two evaluation methods used in Arizona: the AZ Corporation Commission method and the Salt River Project method, both use a deemed savings approach. The AZ Corporation Commission allows for the jurisdiction to claim up to 1/3 of verified (i.e., in operation) savings from evaluated codes and standards. It is unclear if compliance rate and NOMAD are included in the net savings. The Salt River Project method allows up to 50% of verified (i.e., in operation) savings from evaluated codes and standards based on a point system. The savings are adjusted for compliance rate prior to attribution assessment, but it is unclear if NOMAD is included in the net savings.

BC Hydro

Similar to Arizona, BC Hydro method uses a deemed savings approach and allows up to 1/3 of total codes and standards portfolio savings to be attributed to the program. It is unclear if compliance rate or NOMAD adjustments are included in net savings.

Northwest Energy Efficiency Alliance

NEEA, which represents utilities across OR, WA, ID, and MT has used different approaches based on need. For appliance standards, NEEA has used a weighting approach to assess percentage of savings for supporting DOE Appliance Standards. The net savings estimated by DOE are used as the basis for attribution, and account for compliance rate and NOMAD. The weighting method applies weighting factors to barriers to code adoption and separate weight factors to assess the effectiveness of energy efficiency advocacy activity (including NEEA) to address the barriers. Finally, NEEA's participation in these activities is used to obtain a final percentage of total codes and standards savings per appliance standard. For commercial building energy codes, NEEA piloted a method that assesses savings through on-site compliance verification and energy bill analysis. The as-built projects were compared to prototypes to assess savings. "Building energy use was normalized, summarized, and disaggregated by end use. Total and end-use energy use intensities (EUI) were used to correlate energy use and compliance in the buildings (Storm, Baylon, Hannas, & Hogan, 2016)." Compliance is assessed at the end use/building system level rather than whole building or individual measure level. It is unclear if NOMAD is included in net savings.

Pacific Northwest National Laboratory (PNNL)

PNNL assessed national building energy codes for all states which have adopted the model code (i.e., states that do not mandate the national code or have their own code were not included). PNNL's method included calculating potential savings using EUIs from baseline (current) code and comparing to EUIs from adopted (new) codes, usually at an end-use category rather than whole building or individual measures. The EUIs from adopted codes are de-rated using realization (aka compliance rates) based on literature review and field studies. For residential measures, the first-year realization rate is 80% and increases to 100% over 10 years; for commercial measures, the first-year realization rate is 50% and increases to 80% over 10 years. The net savings do not account for NOMAD, but instead PNNL uses a rolling baseline so that each new code is assessed against the previous code, which PNNL suggests captures improvements in the market over time. As stated in the report, "energy codes remain the primary mechanism through which improvements in energy efficiency are enforced on the majority of the building stock. In this analysis, potential improvement between successive codes is entirely attributed to the new code (Athalye, Sivaraman, Liu, & Elliott, 2016)."

Department of Energy Appliance Standard Approach

The DOE assesses energy impact as part of technical analysis for federal appliance standards development. The assessment provides annual national energy savings, which are equivalent to California's net appliance standard savings. The overall approach of DOE's savings assessment is to compare annual national energy consumption by the appliance of interest with and without the proposed standard options in future years.

Energy consumption of an individual appliance is determined by its efficiency rating and operational parameters. For estimating annual national energy consumption, DOE considers possible combinations of efficiency rating and operational parameters for existing and new appliances. DOE first identifies possible levels of energy efficiency rating and values of operational parameters and then assess their market distribution among installed appliance based on market study results or assumptions. DOE then uses a Monte Carlo method to develop a sample of appliances with their efficiency rating and values of operational parameters reflecting market distribution. These appliance samples, along with their efficiency rating and values of operational parameters, are used to estimate the distribution of energy consumption for the entire stock of installed appliances.

DOE assesses the annual national energy consumption for the baseline market condition, under which no new standard is adopted, and different standard adoption scenarios. The difference in annual national energy consumption between the baseline and a standard adoption scenario for a future year represents the annual national energy savings of the corresponding standard option for that year. It should be noted that the baseline market condition includes a forecast of future adoption trend of high-efficiency options and, therefore, NOMAD is included in the baseline. In this way, the baseline represents a business-as-usual case, reflecting market trend without standard improvement. DOE does not explicitly consider compliance of new standards. However, for each adoption scenario, DOE forecasts future market adoption of different efficiency options and, therefore, effectively includes compliance forecast for future years.

9.3.6 Conclusion

This review uncovered the evolution of the codes and standards Program illustrating where accuracy, transparency and streamlining has occurred over the last decade as well as identified where improvement is still needed within the model. Furthermore, the attribution models used outside of California (by state and federally) illustrate the array of potential options that could be adopted or integration into the California codes and standards approach.

9.4 Interview Insights: Attribution Model

9.4.1 Transparency

Potential Savings

- The calculation methodologies are not transparent enough, a consistent methodology should be used throughout the impact evaluation for determining potential savings.
- Due to the complexity of the California climate zones and multiple types of buildings, many CASE authors will model the worst-case scenario versus the actual energy savings and costs. Other programs use the CASE reports for their potential savings estimates. In addition, the CASE reports some of the standards can be adopted a little bit differently and a mismatch occurs because of that deviation.
- Looking at actually where the market stands instead of assuming that the baseline year isn't necessarily where the market is.

Gross Savings

- The broader stakeholder group does not understand how the impact evaluators calculate the compliance rate. The compliance rate should line up with other programs that undergo compliance calculations. Different data sets gathered from field tests are being used and resulting in different compliance rates.

Net Savings

- Significant lack of transparency in how the NOMAD curves are determined as well as how the Delphi panel was selected. IOUs don't know who draws the curves, who was selected and the criteria that was used to select the panelists.
- The purpose and understanding of NOMAD need to be shared more broadly and its potential utilized more to help with layering issues.
- NOMAD should not include the activities of EE programs that influence market transformation programs. NOMAD should consider a shift from looking at the adoption curves of technology into the actual energy consumption of those technologies.

Net Program Savings

- Selection of the attribution panelists is not transparent and there is little information available on how the panelists were selected. There is a sense that those who are selected are new to their jobs.

9.4.2 Streamlining

Potential Savings

- CASE reports should align their estimates for potential savings with other studies to reduce redundancy.
- The process is just repeated by the impact evaluators after the IOUs have completed their potential savings estimates.
- There could be an opportunity to use the CEC numbers for potential savings estimates for both T24 and T20.

Gross Savings

- Utilize a default value for compliance that stakeholders can agree on if determining the compliance rate is too costly and too lengthy to undergo. There is a big cost concern related to the in-depth knowledge that is needed to determine the compliance rate.

Net Savings

- NOMAD can be used as layering tool.

Net Program Savings

- The feasibility attribution factor is focused on demonstrating feasibility but not getting IOUs to complete activities to make it more feasible.
- Need to connect the dots better between the logic model and the attribution factors.

9.4.3 Accuracy

Potential Savings

- Estimating the market baseline can be a challenge for both T20 and T24. For example, for T20, the market baseline shifts depending on how many hours an individual watches TV. For T24, there is a need to focus a bit more on developing an accurate market baseline understanding the complexity and utilize more surveys.
- The accuracy of potential savings is critical as if an error emerges during this process the whole attribution model could be off.
- The actual potential savings estimates may not include the potential for prescriptive measures that have helped raise the bar of overall energy efficiency in buildings.
- There is inconsistency with how the statewide teams goes from per unit to statewide estimates.

Gross Savings

- Need to ensure that this process includes the integration of interactive effects. The compliance rate is linked to the baseline estimate and the IOU codes and standards staff are not aligned on the methodology. Assumptions used for the compliance rate are derived with a lack of data sources.
- It takes a while for old appliance stocks to get washed out of the system and the compliance rate doesn't reflect this current market reality and how long it takes to penetrate the market.
- The timing of the compliance rate process adds accuracy limitations. As the building construction timeline does not align with the impact evaluators timeline for determining overall building stock compliance.

Net Savings

- Market adoption is already occurring pre-rulemaking when the IOUs get involved. Currently, the NOMAD curves do not consider this effect.
- NOMAD curves are a really tricky thing to quantify. Need to look at the NOMAD curve more carefully and potentially take a look at what other states are doing as a baseline. Furthermore, an ability to test the accuracy of the NOMAD curves needs to be integrated.

Net Program Savings

- Identification of more attribution factors that accurately reflect the activities completed by the IOUs. Including the activities that the IOUs participate in at the federal level.
- The current model could incorporate attribution to how the IOUs have reduced the cost of high-efficiency energy measures over time.
- Incorporating how important each factor was in supporting the new codes and standards as well as the amount of effort that each measure takes to get into code or standard.

9.5 Net Program Savings: Program Attribution Theme and Activity List

Theme	Description	Example Activities	
		Rulemaking Support	**Long-Term Development Planning**
Planning and coordination	**For Specific Rulemaking Support**: Before and during the rulemaking process, coordinate with CEC on standard development scope and detailed measures, strategy, resources, specific adoption issues, and schedule. **For Long-term Development Planning**: Develop long-term standard development strategies and plans to support the achievement of state policy goals. Continuously work with other entities, including voluntary standard organizations and efficiency advocacy groups, to promote future voluntary and mandatory standards development. Coordination with industry on standard development plans.	Develop CASE topics, potential standard requirements, and supporting documentation for CEC to consider. Work with the CEC to determine CASE topics for a specific rulemaking and potential efficiency improvement requirements for selected CASE topics. Coordinate with the CEC to adjust CASE study topics and standard development directions based on CASE study findings, stakeholders' feedback. Coordinate regular meetings with the CEC to discuss detailed standards development issues scheduling, arranging, setting agendas for meetings, hosting meetings.	Conduct interval development of future standard improvement strategies to achieve ZNE goals and climate policy goals. Consistent messaging with stakeholders on future codes and standards development. For example, on-going communication with manufacturers on product development and market trends. Continued engagement with working groups and committee members via symposiums, conferences, forum beyond (outside) the code development process, etc. Support and advocate for development of voluntary standards, e.g. ASHRAE standards, IECC standards, and EnergyStar specifications.

Theme	Description	Example Activities	
		Rulemaking Support	**Long-Term Development Planning**
Data collection	**For Specific Rulemaking Support**: Data collection for technical development and advocacy support of specific code or standard during pre-rulemaking and rulemaking process based on. **For Long-term Development Planning** Data collection on topics/technologies and market intelligence gathering activities Identifying additional data needs for future codes and standards updates	Collect industry (manufacturers, builders, designers, etc.) information on specific CASE topics, including product specifications for standard development decisions. Collect cost data from web scrubbing, and from manufacturers, distributors, and suppliers. Collect stakeholder concerns on specific CASE topics Conduct lab tests Conduct field tests	Conduct residential and commercial building surveys to collect data needed to address increasing requirements of providing detailed energy use data to demonstrate savings and cost-effectiveness of proposed standards Market research and technology trends for market awareness and readiness Gather stakeholder concerns on policy issues Develop guidelines for test procedures Conduct early test procedure review Identify drawbacks of existing appliance test procedures and develop improvement methods, e.g. how to test efficiency degradation caused by short cycle of HVAC equipment. Conduct field tests through code readiness programs to collect performance and cost data for measures, e.g. VRF, mini-split, HPWH, and DOAS, for future standard and compliance model development.

Theme	Description	Example Activities	
		Rulemaking Support	Long-Term Development Planning
Method development	**For Specific Rulemaking Support**: Identification of CASE study topics and standard development approaches prior to CEC selection of specific CASE topics. Develop innovative standard development method sand requirements. Develop solutions to standard adoption barriers. **For Long-term Development Planning** Develop standard development methods and plan to improve efficiency of certain end use and technology areas.	• Conduct internal discussion and seek input from consultants to vet potential CASE topics based on technology trends and market readiness information. • Develop Title 24 prescriptive package with special requirements to overcome barriers related to federal pre-emption. • Develop new and advanced building design solutions to demonstrate feasibility and provide compliance method to enable new Title 24 requirements. Examples include residential high-performance attic requirements and non-residential building plug load controls.	• Investigate possible measure packages for achieving residential and nonresidential building ZNE. • Develop methods to regulate energy use of plug load. • Develop standard solutions to significantly reduce HVAC energy use in commercial buildings. • Develop standard solutions to reduce greenhouse gas emission from refrigeration systems. • Conduct gap study on Title 24 compliance software to identify areas for improvement. • Develop compliance options and modeling methods for possible future standards.

Theme	Description	Example Activities: Rulemaking Support	Example Activities: Long-Term Development Planning
CASE report development	**For Specific Rulemaking Support**: Develop draft and final CASE report to advocate and support standard development. Solution Development: Identify probable outcomes and conducting data analysis to support the proposed outcome Technical development data analysis: Analyze all relevant data collected to develop technical documentation for code and standard adoption such as CASE studies and technical analysis.	Market research data analysis to develop market trends for specific CASE topics, measure specific savings analysis Cost data analysis for measure specific cost-effectiveness. Cost data analysis for understanding market trends. Conduct research and data analysis based on stakeholder requests. Act as SMEs on specific CASE topics. Act as key SMEs on specific technologies (appliances, lighting, HVAC, ZNE, etc).	
Stakeholder outreach, engagement, and advocacy	**For Specific Rulemaking Support**: Conduct stakeholder engagement, outreach, and advocacy to ensure successful standard development. **For Long-term Development Planning** Conduct on-going engagement to ensure continued working relationship and communication channels to support future standard development.	Maintain healthy and trustful relationships with key stakeholders. Collect feedback on specific CASE topics. Collect feedback on industry players understanding of market trends. Seek support for specific CASE topics. Seek reconciliation with adversaries. Gain agreement and alignment on specific codes and standards. Work with working groups and standards committees during codes and standards development cycle	Continue to improve the stakeholder coordination and engagement. Utilize and develop key management and coordination tools to integrate into the overall process (e.g., T24stakeholders.com). Integrate the key stakeholder communication and findings from previous code cycles into current code cycle Act as trusted advisor with the stakeholder community

Theme	Description	Example Activities	
		Rulemaking Support	**Long-Term Development Planning**
Rulemaking and implementation support	**For Specific Rulemaking Support**: Support CEC during pre-rulemaking, rulemaking, and post-rulemaking processes Support federal appliance standard development conducted by DOE and legislation development process	Participate CEC rulemaking meetings to present CASE proposal, answer questions, address concerns and objections, and advocate stringent standard options. Develop response to stakeholders' and CEC's questions on CASE proposal. Develop comment letters to address CEC's findings, proposals, and decisions. Conduct additional data collection and technical analysis to address questions, concerns, and objections brought up during the rulemaking process. Support the compliance software development need to implement the new Title 24 standard. Review federal rulemaking documents and develop comment letters to provide in-depth technical input, analysis results, and/or additional data to advocate for adoption of stringent standards. Work and share information with other advocacy groups to develop response and proposals for federal appliance standards development.	Understand the positions of the various actors, most of these below are variations on this idea. Understand industry players knowledge and perception of market trends, and their position on policies. Understand their position in the industry/market (get to know the industry players). Seek support for general advocacy (know allies and adversaries) Seek reconciliation with adversaries.

10. References

ADM Associates, Inc. (2006). *Process Evaluation of 2003 Statewide Codes and Standards Program.* Sacramento.

ADM Associates, Inc. (2006). *Process Evaluation Of 2003 Statewide Codes and Standards Program.* Sacramento.

Antonopoulos, C. A. (2013). Difussion of Energy Efficient Technology in Commercial Buildings: An Analysis of the Commercial Building Partnerships Program.

Athalye, R., Sivaraman, D., Liu, B. B., & Elliott, D. (2016). *Impacts of Model Building Energy Codes.* Retrieved from Energy Codes: https://www.energycodes.gov/sites/default/files/documents/Impacts_Of_Model_Energy_Codes.pdf

CADMUS & DNV GL. (2017). *California Statewide Codes and Standards Program Impact Evaluation Phase Two, Volume Two: 2013 Title 24.* Portland.

CADMUS. (2017). *California Statewide Codes and Standards Program Impact Evaluation Plan Phase Two, Volume Two: Appendices E-P Impact Evaluation Report.* San Francisco.

CADMUS Group Inc. (2012). *2010-2012 California Statewide Codes and Standards Program Process Evaluation Final Report.* Portland.

CPUC. (2005, April 21). Decision 05-04-051-Interim Opinion: Updated Policy Rules for Post-2005 Energy Efficiency and Threshold Issues related to Evaluation, Measurement and Veritification of Energy Efficiency Programs. California.

CPUC. (2018). *Decision 18-05-041.* CPUC.

Decision 10-04-029, Decision 10-04-029: Determining Evaluation, Measurement and Verification Processes for 2010 Through 2012 Energy Efficiency Portfolios (Public Utilities Commission of the State of California April 8, 2010).

Decision: 07-10-032 , 07-10-032 Interim Opinion on Issues Relating to Future Savings Goals and Program Planning For 2009-2011 Energy Efficiency and Beyond (Public Utilities Commission of State of California April 13, 2006).

DNV GL & CADMUS. (2014). *Statewide Codes and Standards Program Impact Evaluation Report For Program Years 2010-2012.* Portland.

DNV GL & CADMUS. (2016). *California Statewide Codes and Standards Program Impact Evaluation Report: Phase One Appliances.* Portland.

Doris, E., Cochran, J., & Vorum, M. (2009). *Energy Efficiency Policy in the United States: Overview of Trends at Different Levels of Government.* Retrieved from National Renewable Energy Labratory: https://www.nrel.gov/docs/fy10osti/46532.pdf

Energy Efficiency Policy Manual R.09-11.014 Version 5 (July 2013).

KEMA, I. (2009). *Market Effects and Maket Transformation: Their role in Energy Efficiency Program Design and Evaluation (Calmac Study ID CPU0045.01).* CPUC.

KEMA, Inc.; The Cadmus Group, Inc.; Itron, Inc.; Nexus Market Research, Inc.; ENRG, Inc. (2010). *California Investor Owned Utilities' Codes and Standards Program Evaluation for Program Years 2006-2008 Final Evaluation Report.*

Mahone, D., Hall, N., Megdal, L., Keating, K., & Ridge, R. (2005, April 7). Codes and Standards White Paper on Methods for Estimating Savings. 1-61.

Services, T. E. (2016). *Whitepaper on Voluntary Standards and Specific Advocacy.*

Storm, P., Baylon, D., Hannas, B., & Hogan, J. (2016). *Commercial Code Evaluation Pilot Study Final Report.* Retrieved from Northwest Energy Efficiency Alliance: https://neea.org/img/uploads/commercial-code-evaluation-pilot-study-final-report.pdf

The Cadmus Group, Inc. & Quantec. (2009). *The Proposed Cadmus Attribution Methodology.* Watertown.

The Cadmus Group, Inc. (2012). *California Statewide Codes and Standards (C&S) Draft Evaluation Plan.* Watertown.

The TecMarket Works Team. (2006, April). California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals. San Francisco, California.