

Administering, Managing and Overseeing Large Scale Portfolio Evaluations – Formulas for Success

*Nick Hall, President,
Evaluation Management and Oversight Services
Oregon, WI 53575*

ABSTRACT

As more evaluation contracts are structured to include evaluating multiple programs within a portfolio of energy efficiency programs there is a need to better understand the conditions that can be expected to impact those studies. The portfolio evaluation manager is subjected to a different level of pressures and conditions that influence that process compared to an evaluation of a single program within a portfolio. This paper highlights a number of issues of which the evaluation professional must be aware, and be ready to manage the associated evaluation efforts in a way that keeps the evaluations focused, objective, reliable and on-time.

Introduction

This paper discusses a number of issues that directly impact the design, implementation, management and oversight of the large portfolio evaluation function.

The move away from contracting single program evaluations to contracting for larger portfolio scale evaluations has changed the operations and management of the evaluation function. Evaluation professionals are exposed to new pressures and less flexible timelines. Evaluation studies must be able to prove that they were conducted using up-to-date methods that follow accepted protocols. Evaluation experts are experiencing new pressures from stakeholders who are not shy about expressing their opinions on what the evaluation function should be doing for them. Yet, it is still possible to design, plan and conduct objective, high-quality evaluations that present reliable objective findings.

The intent of this paper is to help evaluation managers understand some of the kinds of issues that may be faced as they design, direct, and manage these large projects that are not faced in the typical evaluations of smaller projects. Evaluating a single program does not carry the same level of impacts, risks, costs, regulatory attention, reporting, and stakeholder interests as a large portfolio of programs worth hundreds of millions or, even billions of dollars. Indeed, today the portfolio evaluation function alone can cost several millions of dollars. As the value of a portfolio increases, the degree of stakeholder focus on the evaluation's processes and results also increase. This increased focus can take the form of increased regulatory scrutiny, oversight and involvement, as well as increased attention on the evaluation objectives and approach, and on the evaluation findings and their associated impacts on the companies responsible for achieved cost effective savings. Much of this attention can be centered on the needs of the stakeholder which may or may not be directly linked to the need for reliable evaluation results. For these reasons the need for expert evaluation management and oversight has never been greater than it is today. This need continues to grow as the evaluation function, not only documents energy impacts, but also may become the primary way in which energy efficiency induced carbon reduction accomplishments are certified.

The information presented in this paper is derived from the experience gained through managing and overseeing four large energy efficiency program portfolios, as well as conducting, managing and overseeing several hundred program evaluations over the last 35 years. The topics included in this paper are:

1. Approaches should verify and estimate all impacts of the portfolio.
2. Establishing and maintaining the independence of the evaluation and oversight functions.
3. Effective collaboration and the collaboration process in evaluation planning.
4. Development, use and adherence to up-to-date evaluation protocols and evaluation guidelines.
5. The role of the technical resource manual.
6. Keeping an eye on the large and the small evaluation needs that impact the program portfolio.

7. Getting the right people on the evaluation team and in key positions.
8. Setting and managing timelines and expectations.
9. Policy shifts and refocusing of studies.
10. A plan for capturing errors and omissions
11. Ability to foresee regulatory needs including preparing testimony.

This paper by no means provides a thorough or exhaustive treatment of the portfolio evaluation and oversight function. However this paper does focus attention on many of the key issues that, if addressed well, can lead to a successful, useful and well managed portfolio evaluation effort.

In the past 35 years many of the evaluation studies our field has produced have been program-focused. That is, they focus on the evaluation of a single program or a small group of similar types of programs serving a single market sector (for example: a residential lighting program or a group of residential programs, or a couple of low-income programs). Large scale portfolio evaluation groups the evaluation of a portfolio of programs serving all types and size of customers into a single coordinated evaluation effort. Typically the portfolio evaluation effort is structured under a single budgeting process that distributes the evaluation resources across the portfolio of evaluation needs. Large scale portfolio evaluation is relatively new to the energy efficiency evaluation field and few papers are currently published that help inform or guide the portfolio evaluation management and oversight function.

This paper is structured by topic to allow a brief discussion on each of the listed topics. These discussions are not meant to be complete or comprehensive. Indeed, a single paper could be devoted to each of these topics. Instead this paper provides for a sharing of thoughts regarding each of the topics presented.

The remaining sections of this paper are devoted to sharing a few key items regarding the management and oversight of the portfolio evaluation effort.

Approaches Should Verify and Estimate All Impacts of the Portfolio.

The evaluation function is responsible for documenting the impacts of the entire portfolio. In most portfolios only a few programs, or types of programs, may be responsible for the majority of the impacts achieved. Within many of these programs only a few measures may be responsible for the majority of the savings within their respective programs. As a result, some evaluation professionals tend to focus their evaluation resources on those “high-impact” programs and measures and leave the rest for technical reference manual based estimates of savings.

This is a good policy as long as it does not reduce the accuracy of the savings being achieved from all programs, all actions, and all measures offered within the portfolio. In some portfolios, it is sometimes the case that almost no evaluation resources are focused on programs or measures that have minor impacts or on measures that have lower levels of adoption and use. In several cases, the estimates of savings have not been informed by participant-based technology use or baseline condition research. This is a mistake if this focus results in inaccurate program or measure level estimates of impacts to the extent that the organizations offering the programs do not have the information to effectively design and manage their programs or to understand if they need to change their programs to become more effective. When an evaluation effort uses a limited approach such as that described above, the evaluation function has to assume a portion of the liability for offering less effective programs within that portfolio.

In a very large portfolio evaluation effort in a western state, the evaluation teams were instructed by the commission’s assigned oversight agent to focus almost exclusively on high impact measures and only on those programs that captured the majority of the savings. Unfortunately this left the majority of the programs in the portfolio with impact estimate error bands so wide that the program designers and managers could not determine if their programs or measures within those programs were effective. In this case the commission’s oversight agent forgot that one of their key functions is to help program designers improve all programs in the portfolio by providing accurate impact estimates. What this means is that good portfolio evaluation starts with good program-level evaluation planning that also deals with the evaluation needs of the portfolio, as a whole.

The portfolio evaluation needs to provide the information needed by a wide range of stakeholders. Regulators and oversight officials need data to judge the performance of the portfolio as a whole, but also each of

the programs that make up that portfolio. This is important because changes or improvements to a portfolio are typically made at the program level. For example:

- Program implementers need information to design, manage and implement cost effective services and to understand which services to delete, keep or change.
- Utilities and regulators need to know if what they are doing is prudent and represents a sound public investment.
- Marketing managers need to know if they are reaching the right people with the right messages and incentive structures and if their efforts are developing a service pipeline that the programs can handle.
- Policy makers need to understand if the services are consistent with formed policy and performing in a way that supports policy objectives.
- Program and portfolio designers need to know if their assumptions about impacts are accurate and if their services are cost justified from a weather-sensitive comparative market acquisition perspective, and if not, what service design changes are needed.
- Commissioners need to know if energy efficiency supplies are reliable and if costs and benefits are accurate.
- Power planners and grid operators need to know if energy impacts are dispatchable, if they are available when needed.

The evaluation planning and implementation process needs to not only understand which objectives need to be designed into their studies, but also understand the individual stakeholder's metric reporting and use structures as well as the information timelines under which the evaluation results are needed.

Establishing and Maintaining the Independence of the Evaluation and Oversight Functions.

It is important that the portfolio evaluation function be established outside of communication and control operations or contractual agreements that allow stakeholders to influence the evaluation effort or the way findings are formulated and reported. The larger the risk to the organizations responsible for the performance of the programs, or to the various stakeholders associated with the evaluation findings, the more potential there is for attempts to influence the evaluation results. This is especially true in today's politically charged environment where legislative act or a commission sets goals to achieve more savings from energy efficiency programs, or where some see a portfolio as an affront to the operations of a theorized free market or as regulatory over-reach.

However, do not assume that the pressure to influence the evaluation effort will only be focused on increasing savings or on undermining the objectivity of the evaluation effort, all under the pretense of a commitment to evaluation excellence. The evaluation function today is just as likely to see stakeholder pressures focused on lowering the level of apparent savings as much as pressures to increase apparent savings, especially in states where utility regulation is seen as an unwarranted restriction on business operations or investment returns.

Stakeholders who do not want to offer energy efficiency services to customers for any number of reasons would like to see lower savings to support their positions. As one president of a utility told the author of this paper: *"We are in the business of building power plants and selling power, why would we want to reduce our sales or profits?"* Stakeholders who want increased energy efficiency would like to see higher levels of savings or more cost effective programs. From these stakeholders we hear such things as; *"Can we increase savings if we used a different evaluation approach?"* or; *"Can we evaluate this effort two different ways and pick the approach that provides the most savings?"*, or; *"We obtain more savings by applying the TRM¹ than by doing EM&V², why not just apply the TRM results and be done with it?"*.

In the author's experience, the rate at which attempts are made to influence evaluation findings has grown considerably as the size of the programs or portfolios has increased and as financial systems are established to reward performance. In some cases program implementers themselves push for the use of evaluation approaches

¹ TRM = Technical Reference Manual – Provides calculation approaches used to estimate typical savings from an assumed typical customer under assumed typical use conditions with typical weather.

² EM&V = Evaluation, Measurement and Verification. Evaluation efforts supported by in-field measurements and operational verification assessments.

that favorably influence the evaluation findings of their programs and services. In some cases they want to gain significant control over the evaluation effort, or have the evaluation function employ specific evaluation approaches which influence savings in a specific direction. In some cases program implementers are pressuring regulators to allow them, and only them, to be able to conduct evaluations of their programs, collect the evaluation data, or sort customers into test and control groups that are then handed to the evaluation team to be used to assess their program's performance. Other implementers are placing language in their implementation contracts to legally limit the collection or review of performance data necessary to conduct an independent evaluation.

The stakes are high, and they are only going to get higher if climate change issues influence, not only the price and supply of energy, but also the credits provided to reward energy or carbon reduction efforts.

In today's evaluation environment, it is important to establish and maintain an evaluation function that is independent of these pressures. It is important to establish the evaluation function to operate independently of the influence of stakeholders who have a direct interest in the evaluation results. The role of the independent evaluator needs to be similar to the role of the independent arms-length external auditor. Yet, even these relationships are vulnerable. Enron seemingly controlled, or influenced, how their external auditing service provider documented the financial health of the company, thereby establishing a false sense of accomplishment and financial health. This condition should not be allowed to occur in the field of energy efficiency evaluation, especially the evaluation of large portfolios of energy efficiency services. The independent evaluation function must be independent to establish the evaluation approach, identify the data needs for the study, independently collect that data, and be free to use the best and most reliable evaluation protocols or approaches within the allocated resources. That approach needs to be independent of all stakeholders, including the regulatory commissions and not just the companies or organizations offering services in the market.

In just the last 10 years the field has experienced the effects of a number of attempts to influence our portfolio and program evaluation efforts including, for example:

- For-profit program implementation firms trying to force evaluation experts to use only the study approaches approved by those firms;
- Firms want to "assist" in the collection of data or direct the allocation of participants into the test and control groups using approaches other than random assignment;
- Senior utility executives modifying findings and then developing press releases with false information making the portfolio appear not cost effective to voters and minimally-informed policy makers;
- Utility company executives clandestinely trying to force changes in evaluation management assignments when the original management team refused to accept biased guidance related to evaluation procedures;
- Commission staff pressuring evaluation experts to conduct studies using approaches that make programs appear less or more cost effective than they are;
- Commission staff hiring "expert" advisors that interfere with the evaluation designs and implementation strategies which leads to increased costs and lower findings accuracy;
- Utility managers asking to use outdated evaluation protocols or TRM applications because they provided higher levels of savings estimates.

This is the field in which evaluation professionals find themselves operating today. Fortunately these cases are not pervasive and in most all cases the evaluation community has withstood these attempts to influence our work. However, it is wise for the evaluation function associated with larger portfolios to be aware of the ways in which stakeholders will attempt to influence their work. There is an increased probability that these and other types of pressures will be placed on the portfolio evaluation effort because of its increased size and the implications associated with the evaluation findings.

Effective Collaboration and the Collaboration Process.

Portfolio level operations oversight and the evaluation function oversight typically involve a committee-structure, or some form of a multiple stakeholder collaborative management or oversight process that is often controlled or overseen by a regulatory commission or their agents. The agents can be in the form of a person, a team or an organization and can be tightly or loosely linked to the regulatory commission. In some cases the regulatory organization or their agents have served as the lead evaluation oversight structure, or have even acted as

organizational or operational directors, within that structure. In some states this approach has worked well allowing independently conducted evaluations while still engaging in a collaborative exchange of ideas relating to the evaluation effort. In other states these approaches were established in ways that allowed biased agents or organizations to control or influence not only the evaluation approach, but what data could be used to support the evaluation function and even specified how that data could be assessed. In other cases, the evaluation function has been able to maintain their independence from both the regulatory organization and from the oversight approach and have effectively observed or overseen the evaluation effort, at arms-length. There are other approaches as well.

Regardless of which approach is used to provide oversight the evaluation function must be well planned and structured with clear lines of communication, and established clarity about who is responsible for making what types of evaluation decisions. That process must be established in order to document decisions and decision rational, but also to avoid unclear decisions, resolve decision conflicts, avoid misrepresentation of the decisions, and reduce the level of over-dominance by the more vocal or controlling members of the collaborative or oversight structure.

Such processes should include ‘checks and balances’ that 1) ensure that individuals, representing special interests, do not have the opportunity to exercise greater influence over the evaluation effort outside of the established decision making process. e.g., without the knowledge or consent of other members of the associated decision body; 2) prevent key stakeholders from influencing program decision makers to avoid a type of evaluation that would tend to over or underestimate savings impacts; and 3) ensure the effective independence of the evaluation function.

For these reasons, and others, it is important to establish a protocol or standards of conduct for committee membership or stakeholder contact with the evaluation function that can occur outside of formal documentation or meetings. That protocol should contain a formal process for documenting contacts with the evaluation leadership, identifying the reason for each contact, and indicate the results (or expected results) of each contact. The protocol needs to be established so that the entire collaborative or committee structure is informed of those communications to insure transparency. If possible, the approach should be established so that the communication occurs within an open meeting with formal documentation of the discussion and the decision.

While there needs to be a clear documentation trail of communications and decisions that influence the evaluation effort. This should not interfere with the evaluation team’s access to key stakeholders who must provide the information needed to support the evaluation. Evaluation experts need immediate and direct access to a wide range of stakeholders. The evaluation professional will typically need to have rapid and direct access to and contact with program and policy professionals during the evaluation planning and implementation stages. The ability to have rapid access is important if timelines are going to hold. Delays in access can often result in delays in task launch or completion. This is especially true for programs within portfolios that are constantly being changed in order to maintain cost effectiveness or meet a stakeholder need. Evaluation professionals need to be embedded in the information stream associated with any program related activity so that evaluation efforts can be adjusted in real-time.

Development, Use and Adherence to Up-to-Date Evaluation Protocols and Evaluation Guidelines.

The use of accepted evaluation protocols is important to the evaluation planning function, as they help inform stakeholder understanding of and confidence in the reliability of the evaluation study results. The studies need to be conducted in a way that can be supported by the evaluation community, especially under formal cross-examination of the approach used and of the study results. When studies deviate from accepted methods, the results of the findings can be challenged, placing the evaluation professional in a position of needing to defend the study approach under public testimony and cross-examination.

When study methods represent an improvement to typical approaches, it is easier to defend the new approach. However, when an approach deviates from an accepted protocol in an effort to control costs, for example, the findings may not be as defensible compared with more traditional approaches. For example, some refrigerator recycling programs are in this situation because the evaluation limits the savings estimate to only the energy impacts occurring within the participant’s home and not across all market channels. Of course, the higher the percent of portfolio savings, the more important it is to follow established protocols or implement evaluations that are more rigorous than what might be indicated in a protocol.

Energy efficiency evaluation protocols appear to be heading for an up-date process associated with the need to rigorously document the level of energy and demand reduction achieved in the context of meeting state savings goals and system planning, but also to assess the carbon impacts associated with reaching state and/or federal carbon reduction objectives. These changes are focusing on both the approach for establishing adequate and representative sample sizes, on the use of on-site metering and monitoring (M&V), and on the use of participant-level weather data and equipment use applications. While these up-dates tend to focus on impact evaluation, there is a continued need for protocols to focus on process evaluations as well. While the impact evaluation documents impacts, it is the process evaluation that helps improve programs and technology services, help programs reach more customers, improve participation rates, and increase the cost effectiveness of those services.

In addition, there is currently an effort in place to determine the need for and level of “certification” needed to accept study findings from evaluators or evaluation firms who heretofore have not been certified. There are strong opinions within the evaluation community regarding the need for various levels of certification, and for the certification approach, if needed. The trend for this process seems to be following two lines of development.

- The first stems from evaluation experts themselves. It appears that the more experienced the evaluation professional is, the less likely they are to support or see a need for certification efforts.
- The second stems from policy professionals. It appears that the more years involved as a policy professional, the more likely they are to support a need for certification efforts.

Where this effort will end within a possible certification process is unclear, however what is clear is that evaluations conducted today will need to be designed and conducted in adherence to protocols and potential certification approaches that may be adopted. Evaluation experts now find themselves in the position of needing to design and apply methodologies that may have to hold up under carbon certification protocols not yet established in order for carbon credits to be applied retroactively to programs or portfolios implemented today. This condition is anticipated because it appears that previous program efforts may be allowed to be credited toward currently forming carbon reduction objectives.

The subject of protocol deviation is becoming a hot-topic in the evaluation industry. Not every program within a portfolio can be evaluated using rigorous protocols as long as the evaluation budget is established at less than 5% of the portfolio costs. Likewise, not all states require or need the same evaluation rigor for their portfolio of programs. This means that there needs to be acceptable trade-offs for applying the rigor levels of protocols, such as the California Evaluation Protocols (which has different rigor paths for each type of evaluation need), or there needs to be a structured decision process for how and when to deviate from protocols such as the Uniform Methods Protocols. These decisions need to be made transparent so that information consumers understand why deviation occurs and the impacts of that deviation on the reliability of evaluation findings.

It should also be noted that efforts are currently underway in the Northeast states to make EM&V practices transparent using a standardized reporting format, developed by the Regional EM&V Forum (see complementary paper (Michaels and Reynold 2015) to this panel by Julie Michals and Arlis Reynolds *What's in Your EM&V Genome?*)

In the context of portfolio evaluation, transparency of EM&V practices should focus on: 1) programs with large budgets, 2) programs with high energy savings, 3) higher risk programs, 4) programs for which there are limited studies in the literature, 5) programs that need to be redesigned, 6) pilot or trial programs that have the potential to be expanded, 7) programs that serve a specialized market or participant base, and 8) programs that have high-impact measures with significant potential for variance in use conditions need higher levels of rigor than other programs. In addition, there is currently an effort in place within the IEPEC to determine options for and level of “certification” necessary to accept study findings and credit those accomplishments. These efforts are may result in certification that addresses any or all of the following: the evaluation methods, data collection approach, calibration and use of the equipment used in data collection, and the skills and the level of expertise for the evaluation managers, or of firms conducting the studies.

The Role of TRMs.

Evaluation experts are often surprised when non-evaluation professionals expect to see evaluated savings results that are consistent with the savings results generated using standard inputs or average savings estimates from

a Technical Reference Manual's (TRM). TRM savings estimates are based on expected levels of typical application conditions for the average participant over a wide range of programs, markets, and use conditions. They are not specific to the conditions associated with any specific program or participant population. In addition, a TRM often assumes a level of baseline or pre-existing equipment that is operational prior to program participation that may not actually be the level associated with a specific mix of participants. Of course, in reality, there is no such thing as an accurate TRM-assumed typical use condition, when it applies to a specific market or the set of use conditions experienced by a specific program.

The bottom line is that TRMs can provide forward-looking estimates of typical (expected) impacts; however, the evaluation effort provides more precise estimates because average conditions are likely to change over time and with each program or portfolio experience. For this reason, the TRM up-date process should include assessments from the results of rigorous evaluations. As program and portfolio evaluations are published, these serve as the primary input data for up-dating a TRM for use in future program planning. That is, the TRM is not a retrospective looking evaluation tool, but rather is a prospective program planning tool. The use of portfolio evaluations to up-date a TRM is much preferable over the use of program evaluation based up-dates. The portfolio will experience a wider range of technology use conditions and customer types than a program evaluation, and as a result, can better inform the TRM up-date process.

Getting the Right People on the Evaluation Team and Placed In Key Positions.

Portfolio evaluations need expert-level evaluation professionals that are experienced at working across a wide range of evaluation associated tasks. Getting the right people on the team can be challenging when those individuals are in high demand, which means that they typically also cost more to acquire, potentially decreasing profits from a given fixed-price study. The vast majority of the costs of an evaluation are labor costs. Because evaluations are typically competitively bid, evaluation contractors that bid a large component of their best staff, which are typically the highest cost staff, do not win many low-bid contracts. This means that when evaluation contracts based on lowest price are awarded, the management of that evaluation effort may not have the level of experience and knowledge required. It may also mean that the evaluation manager may need to focus a good bit of their time on making sure that the people doing the work and the people supervising that work are doing their jobs well.

This is one of the largest problems within our industry; making sure the right people with the right skills, knowledge and experience are engaged in the key tasks that lead to accurate impact estimates (Saxonis 2015). In several cases in recent years, it has been observed that evaluation teams responsible for key tasks have conducted studies without experienced or senior evaluation professionals overseeing their work or performing adequate quality control checks. In one study, for example, the impact estimates for the portfolio came into the independent oversight process with almost 30% lower impacts than what they should have reported. This single error represented hundreds of millions of dollars in lost savings that were not reported in the initial reports. In other reports the evaluation team used experimental design approaches that provided net savings, and then subtracted out freerider savings based on survey responses, substantially underestimating savings by double counting freeriders. Other examples of both over and under reported estimates of savings are not uncommon, resulting from inadequate or unqualified staff responsible for those studies, or the quality control process that was employed.

This is not the environment in which the evaluation function should operate. Fortunately these problems are easily solved by making sure experienced experts are in charge of key tasks and by making sure that there is a quality control function in place that is designed to assure high quality accurate work is performed. However, there are two other aspects that the senior evaluation experts must consider. The first is making sure that the senior management of a portfolio evaluation is continually and proactively reviewing methods, approaches, tasks and task procedures with all individuals responsible for the completion of those tasks. The second is to make sure that the budget is set up to allow for senior experienced experts to oversee key work and to support key management and oversight communications. Dollars are the input resources required for the evaluation function. Low-bid studies get low-bid results and end up with the level of staff expertise that the accepted price can afford. The old saying that you get what you pay for is not fiction.

Setting and Managing Timelines and Expectations

Evaluations always seem to take longer than stakeholders want or need. However, evaluations also seem to take longer than planned, even when planned by experienced evaluation professionals. Every portfolio study has hundreds of tasks and thousands of sub-tasks spread across a wide range of staff expertise and availability. All of these tasks and sub-tasks need to come together in a staged sequence of events that have to occur in the right order at the right time. This is a challenge to even the most experienced evaluation professionals. These studies typically have to be delivered on time in order to meet internal review and commission submission deadlines. This may mean that a portfolio evaluation requires more than a full time evaluation expert to manage the timeline of tasks to meet the commission-ordered delivery timeline.

While it is important to manage timelines and expectations for all studies, it is especially important for portfolio evaluations in which the results of a study are needed to inform a regulatory process with an ordered timeline. This applies to under-budgeted studies just as it does to studies conducted with sufficient resources to hire the right professionals. The tools that any given evaluation expert uses to manage an evaluation timeline are not important, as long as those tools allow that expert to intervene in the evaluation process when an intervention is needed in order to meet a delivery timeline or deliver on the rigor needed.

The experienced evaluation professional will expect delays in about one-third to one-half of the tasks associated with any given study. This expectation needs to be planned into the portfolio planning process so that the studies and their associated tasks have room to move within the established timeline without delaying the deliverables. This may mean that the evaluation expert in charge of a portfolio evaluation may need to reassign tasks to different people within a firm, or even different firms altogether. These actions will not be well received by those firms; however, they may be necessary to assure an on-time delivery. It also means that evaluation contracts or sub-contracts between team members need to be structured in a way that easily allows change. It also means that the clients and other information consumers need to understand that directing an evaluation that has to be on time may require hard choices that are not popular with the team members and the owners of those companies.

Policy Shifts and Refocusing of Studies.

All evaluation experts have experienced the gut-wrenching experience of having evaluation policy shifts in the middle of a study. The study that was carefully planned no longer fills the policy needs associated with the reasons for conducting the study. Example of these shifts are many: the commission who wanted kWh impacts now wants gas and kW impacts; the stakeholders who did not need estimated demand impacts now needs multi-year verified demand impacts accurate for a selected set of participants; unneeded carbon reduction estimates from coal or natural gas combustion are now needed for both gas and electricity – separately – by measure; the client now wants market segment impacts in addition to program and portfolio level impacts; the evaluation that did not need interactive commission staff review, now finds that the commission has hired oversight contractors who have to be educated about the research methods associated with the studies in which they now are providing oversight.

It is all okay. This is not unusual. Expect it and plan for it. Be ready with alternative approaches, scope changes and revised budgets. Portfolio evaluation planning is hard and re-planning is harder because you may need to move money, change tasks or alter the study focus; and this may mean changing staff and staff assignments. Policy shifts should be expected in today's evaluation environment and viewed as a normal process associated with managing, directing or overseeing large portfolio evaluation efforts. Policies are going to change. These shifts allow us to use our skills in ways that benefit not only the clients and key stakeholders, but make our research more valuable.

A Plan for Capturing Errors and Omissions

Evaluation experts are indeed human, and portfolio level evaluation efforts are not immune to errors and omissions, as noted in the paper by William Saxonis included in this 2015 IEPEC proceeding (Saxonis 2015). In New York, the evaluation oversight team consisting of managers from the Department of Public Service and a team of contracted evaluation experts, has found that many of the evaluations provided for Commission review and

approval needed some level of correction. Errors ranged from minor calculation errors that minimally impacted the overall findings, to major errors in the estimation process. In several of the reports the evaluation contractors made errors in assumptions, errors in methods development, or errors in calculations that in each case lead to inaccuracy in the presented findings. In each of these cases the internal quality control mechanism of the study team failed to catch the errors before they were presented to New York Commission staff for review.

Fortunately, in the case of the New York studies, a team of experts were hired to review all draft reports, and to confirm that the studies and the approaches used were reliable. This external review proved valuable to the New York process and lead to corrections prior to the acceptance of the evaluation reports. The purpose of including this discussion in this paper is to make evaluation managers aware that there is a need to review both the methods and estimation approaches across all evaluation components of the portfolio evaluation, as well as the findings prior to approving an evaluation deliverable. While the errors or omissions for a program evaluation are important, they become especially important when they cut across programmatic evaluations within a portfolio evaluation effort.

Ability to Foresee Regulatory Needs and Testimonial Efforts.

Lastly, this paper discusses the need for evaluation managers to be ready to engage in the regulatory process to the extent needed for those managers to understand the regulatory environment in the states in which the studies are conducted. This is particularly important when those evaluations are portfolio based and cover a number of programs cutting across all market sectors. This means that the study designs and the development of the deliverables need to understand the information needs of the stakeholders. Evaluation reports must convey the information needed in easy-to-understand simple presentations that focus on the key findings needed to support the regulatory process for the portfolio of programs.

It is becoming more common for evaluation professionals to be called upon to provide testimony or testimonial support that helps assure the regulatory staff and policy makers that the program specific and portfolio wide findings in the studies are accurate, that the methods are sound, and that the studies are conducted by skilled evaluation experts who can be trusted. These managers should expect to be called upon to provide testimonial support, and evaluation budgets should be structured to be flexible enough to be able to provide that support within a very short period of time without impacting the resources need to conduct the studies. This support can be expensive and time consuming. In one such portfolio level support effort the author of this paper needed to budget for over \$70,000 in testimonial support efforts for a single interaction with the commission staff and their contracted oversight agent, in another effort in a large state with a multi-billion dollar portfolio, the evaluation support to the regulatory commission required the equivalent of 3 full time staff for over 12 months to address issues associated with the portfolio of evaluations for a single program cycle.

Conclusion

The move away from contracting single program evaluations to contracting for larger portfolio scale evaluations has changed and focused the operations and management of the evaluation function. Evaluation professionals are exposed to new pressures with which they have to deal. The timelines associated with the evaluation efforts for a portfolio are often less flexible with regulated reporting dates. Evaluation experts must be able to prove that their evaluations were conducted using up-to-date methods that follow accepted protocols or present the reasons for not following those approaches. Evaluation experts are experiencing new pressures from stakeholders who are not shy about expressing their opinions on what the evaluation function should be doing for them or their firm. Yet, it is still possible to design, plan and conduct objective, high-quality evaluations that present reliable objective findings if the evaluation function and the regulatory oversight approach is well established. Evaluation managers must become familiar with the state-specific evaluation market and operational environment and be ready to defend their choices within the evaluation oversight process. The author of this paper hopes this short discussion has provided some help in this task.

References

Michals, Julie and Arlis Reynolds. "What's in Your EM&V Genome? In proceedings of the 2015 International Energy Program Evaluation Conference. Forthcoming. Long Beach CA. August 2015"

Saxonis, William. "Evaluation Quality Is Job One" In proceedings of the 2015 International Energy Program Evaluation Conference. Forthcoming. Long Beach CA. August 2015.