Consequential Analysis of the Greenhouse Gas Emissions Impacts of Actions That Influence the Electric Grid: The Theory and Practice of Using Marginal Emissions Rates

Pieter Gagnon[1] and Eric O'Shaughnessy[2]

1 National Renewable Energy Laboratory
2 Clean Kilowatts

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-6A40-91580
December 2024

Consequential Analysis of the Greenhouse Gas Emissions Impacts of Actions That Influence the Electric Grid: The Theory and Practice of Using Marginal Emissions Rates

Pieter Gagnon[1] and Eric O'Shaughnessy[2]

1 National Renewable Energy Laboratory
2 Clean Kilowatts

Suggested Citation
Gagnon, Pieter and Eric O'Shaughnessy. 2024. *Consequential Analysis of the Greenhouse Gas Emissions Impacts of Actions That Influence the Electric Grid: The Theory and Practice of Using Marginal Emissions Rates*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A40-91580. https://www.nrel.gov/docs/fy25osti/91580.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-6A40-91580
December 2024

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by Amazon.com Services LLC under agreement number FIA-20-17145. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

Preface

This paper discusses consequential greenhouse gas emissions analysis for actions that influence the electric grid—that is, the analytical approach of estimating how such actions change greenhouse gas emissions, relative to what they otherwise would have been without the action. Although the paper starts with general considerations, it ultimately focuses on consequential analysis with marginal emissions rates as an option likely more scalable than others.

Consequential analysis has immense potential to guide efficient, impactful decisions. In theory, selecting actions based on their impact is the optimal way for an organization to maximize its objectives. In practice, however, consequential analysis is constrained by the quality of the estimates it can bring to bear. Although there are many areas where this approach can add value to a decision-making process, there are also many shortcomings in currently available data and methods, which can result in inaccurate impact estimates. Depending on the objectives of a decision maker, this can make the approach unsuitable in some contexts.

Consequently, to help guide future research and facilitate discussions about the suitability of this analytical approach for particular applications, this paper focuses to a great extent on the limitations, uncertainties, and unknowns of the current state of the art of consequential emissions analysis when applied to electric sector actions. This paper's focus on limitations should not be misunderstood: Consequential analysis is a valuable tool for many applications and may defensibly be applied to new applications in the coming years. Nonetheless, focusing on current shortcomings can guide research into improvements that increase the likelihood that decisions made with consequential analysis lead to the intended results. Future advances in the practicing of consequential analysis may mitigate many of the points raised here, expanding the domain where it can defensibly be applied.

Acknowledgments

We thank Matthew Brander, Michael Gillenwater, Lissy Langer, Matt Mowers, Wilson Ricks, Lee Taylor, and Colby Tucker for providing comments during the writing of this paper. We thank the Rocky Mountain Institute for coordinating discussions among practitioners in this topical area. This work was funded by Amazon.com Services LLC, under agreement number FIA-20-17145. The National Renewable Energy Laboratory retained full editorial control over the contents of the report. The authors are solely responsible for errors and omissions.

Table of Contents

List of Figures

1 Introduction

1.1 Electricity Interventions and Emissions Analysis

Actors of all types—such as individual consumers, businesses, governments, and non-governmental institutions—make decisions that influence greenhouse gas emissions from the bulk electric grid. Many of these actors have become increasingly interested in estimating the greenhouse gas impacts of their electricity use and electricity interventions (Schaltegger et al. 2015; Bjørn et al. 2022).[1] This document discusses an analytical approach for making such estimates: **consequential emissions analysis. Consequential emissions analysis seeks to estimate how an action changes emissions, relative to a counterfactual where the action had not taken place.** The greenhouse gas impact estimates from a consequential emissions analysis can be combined with other considerations (such as costs) to rank-order different alternatives available to an actor, ultimately selecting the interventions that maximize their objectives.

Consequential analysis has seen widespread use and elaboration across varied disciplines. As a few select examples from the numerous possible: life cycle analysis (Tomas Ekvall 2019; Earles and Halog 2011; European Commission 2010), benefit-cost analysis (Woolf et al. 2020), policy analysis (Bistline et al. 2023; Steinberg et al. 2023; Xu et al. 2024), and philosophy (Hansson 2007).

In theory, consequential emissions analysis can improve decision making (e.g., identify the highest impact action from a suite of possible actions, maximizing impact per dollar spent, minimizing the cost of a given impact), when compared to decisions guided by non-consequential methods (Brander 2021; Tomas Ekvall 2019). While non-consequential methods are useful for many purposes, they are theoretically suboptimal for selecting between alternatives because they do not seek to directly estimate the impact of the interventions being considered—and therefore any resulting rank-ordering could misallocate resources, potentially significantly.

While acknowledging the theoretical value of consequential analysis, it is also helpful to acknowledge that in practice, consequential emissions analysis can have material uncertainty (Gagnon and Cole 2022; Ekvall 2019; Bistline et al. 2023). Although there are many situations where consequential analysis is generally recognized as adding value to a decision-making process, there are also many situations where data or methods do not currently exist to fulfill the theoretical promise of the approach. Shortcomings in the methods and data of consequential analysis can result in erroneous estimates—research has shown that, in some situations, incompletely constructed consequential analysis can result in less effective decisions than those guided by heuristics or non-consequential metrics (Gagnon and Cole 2022; Xu et al. 2024). Much of this document is focused on describing current methods and their limitations, to help practitioners decide where the application of consequential analysis is appropriate, and to

[1] We use the term "intervention" to describe any action taken by an actor—generally one meant to influence the demand for grid-purchased electricity, the supply of electricity, or any action that otherwise changes how electricity is generated. Interventions can take many forms: increasing or reducing load, choosing to maintain an existing load, shifting ("shaping") load, building electric generators, signing offtaker agreements, purchasing renewable energy certificates (RECs or their equivalents), and so on. This paper focuses primarily on estimating the impact of interventions that change the demand on the bulk electric sector, such as changes in load or injections of clean generation—although that is not to diminish the potential impact of other interventions, such as building transmission lines or facilitating interconnection processes.

provide direction for research activities that could increase the accuracy and utility of this analytical approach.

Key terminology: Consequential emissions analysis
Consequential emissions analysis is an analytical approach that seeks to estimate how an action or set of actions impacts emissions (i.e., the emissions induced or avoided by the action). As described in Section 2, consequential analysis defines the emissions induced or avoided by an action as being the difference in emissions between a scenario with the action and a counterfactual scenario without the action.

1.2 Comparison to Attributional Emissions Analysis

Consequential emissions analysis can be further understood by contrasting the method with another common framework for analyzing emissions known as attributional accounting (Figure 1). Attributional analysis seeks to allocate environmental burdens to specific products, organizations, or countries for a given state of the world, whereas consequential analysis seeks to estimate how specific actions change environmental burdens, relative to the state of the world if the action was not taken (Brander 2021; Tomas Ekvall 2019).

Attributional analyses generally draw boundaries in both time and space when allocating emissions: for example, Scope 1 emissions inventories are attributional and are an accounting of emissions within an organization's boundaries for a specific historical period of time. In contrast, consequential analysis of an action by that organization would require, in theory, a comprehensive estimate of the consequences of the action across all time and space (Brander 2016)—although in practice, consequential analysis generally does draw boundaries meant to encompass the majority of the impacts of an action.[2]

[2] Note that, while *impacts* would in theory be comprehensively reflected in the analysis, the *metric itself* does not need to be, depending on the specific question being asked. For example, it is valid to perform a consequential analysis of an action's impact on greenhouse gas emissions from the electric sector of a particular country (i.e., the metric is defined with boundaries)—but for such an analysis to be theoretically complete, it would still have to reflect any relevant dynamics from other countries (i.e., phenomena impacting the metric do not have boundaries) and in domains beyond the electric sector. The same goes for phenomena across time: It is technically valid to perform a consequential analysis of an action's impact on a metric that has been defined for a constrained period of time (e.g., estimating how an action would influence greenhouse gas emissions from the combustion of fuels during the calendar year 2025, even if the action would also influence emissions later). Despite this being a technical valid option, caution is warranted because such a temporally restricted metric may not be the appropriate metric for the question at hand when greenhouse gas emissions are being analyzed. For example, if rank-ordering actions based on a comprehensive estimate of their environmental impacts, it would generally be preferable to include as expansive a time horizon as is practicable, possibly in combination with temporal discounting.

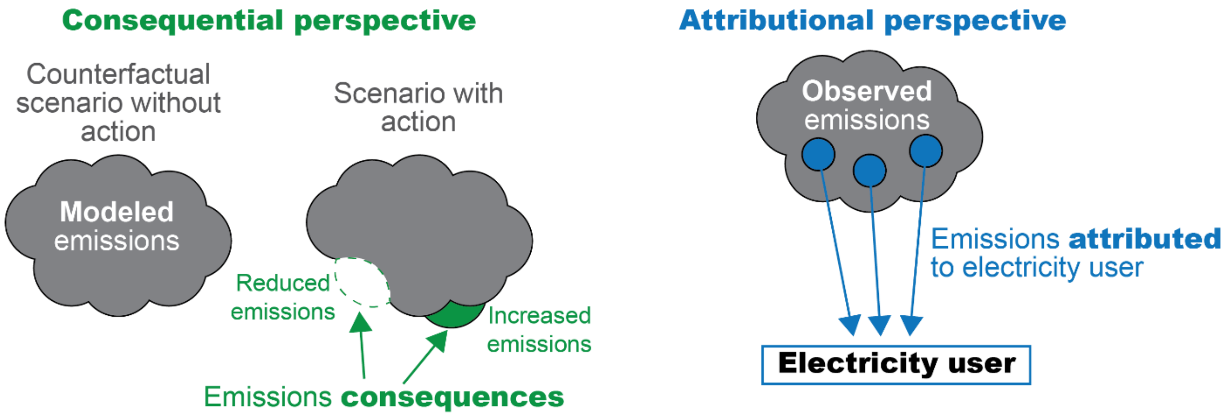

Figure 1. Illustration of consequential and attributional perspectives

Consequential and attributional analysis are both useful emissions analysis tools to apply in different contexts or in conjunction (Anex and Lifset 2014; Brander, Burritt, and Christ 2019). The ultimate selection of which is more suitable depends on the specific application, the quality of the data that can be brought to bear, and numerous other contextual factors. Some scholars suggest that attributional accounting is more theoretically appropriate for assigning "ownership" of emissions, such as in emissions inventories, whereas consequential analysis is more theoretically appropriate for analyzing and guiding specific decisions (Plevin, Delucchi, and Creutzig 2014; Brander and Ascui 2015; Brander, Burritt, and Christ 2019; Brander 2021; Miller 2022).

Parallel analyses separately employing the two methods may sometimes be beneficial (Brander, Burritt, and Christ 2019). For instance, an organization could use attributional accounting to create an inventory of emissions associated with its supply chain, while separately using consequential analysis to report the estimated greenhouse gas emissions impacts of specific projects they are undertaking (recognizing that many organizations have expressed goals in terms of their greenhouse gas inventories, and therefore in practice may predominately or exclusively consider the emissions impacts of an intervention to be how it influences their inventory reporting).

For more complete discussions of the trade-offs and different applications of consequential and attributional analysis, see Brander and Ascui (2016), Brander et al. (2019), Ekvall (2019), and Gillenwater (2023).

2 Stylized Models of Prospective, Retrospective, and Real-Time Consequential Emissions Analysis

2.1 Foundations of Consequential Emissions Analysis

Consequential analysis estimates how an action impacts a metric of interest—for greenhouse gas emissions analysis, that would generally (but not always) be total global emissions. Fundamentally, consequential emissions analysis estimates the difference between emissions in two scenarios: a scenario where the action takes place and a counterfactual scenario where the action does not take place. This fundamental principle can be expressed as:

Equation 1

$$\begin{matrix} \textit{consequential} \\ \textit{emissions} \\ \textit{impact} \end{matrix} = \begin{matrix} \textit{emissions in} \\ \textit{scenario with} \\ \textit{action} \end{matrix} - \begin{matrix} \textit{emissions in} \\ \textit{scenario without} \\ \textit{action} \end{matrix}$$

To be theoretically complete, consequential emissions analysis requires impacts to the metric of interest to be estimated comprehensively—i.e., all relevant phenomena reflected across all space and time, and without sectorial boundaries. In practice, bounds are generally set such that the significant majority of material impacts are captured, balancing completeness with the availability of data and the complexity of the analysis.[3]

Operational and Structural Impacts

For electric sector analysis, it can often be helpful to define two categories of impacts: First, actions can affect the operations of electricity generators. For instance, increased electricity demand from a new manufacturing facility could cause a natural gas plant to have a higher output, relative to what it would have been in absence of the facility's load. These effects are known as **operational impacts.** Second, actions can affect the structure of the electricity grid (i.e., the capital assets of the grid, such as generators or transmission lines). For instance, the manufacturing facility could induce investment in a new wind farm, that would not otherwise have existed in absence of the facility's load. These effects are known as **structural impacts.**

These two categories of impacts map onto another pair of terms: **Short-run** refers to an estimate that assumes capital assets are fixed, and therefore includes only operational impacts, whereas **long-run** refers to an estimate where both operations and capital assets can vary. In practice, this generally means that short-run estimates are useful for reflecting a period of time when the structural impacts of an action have not yet occurred, and therefore are strictly operational. Long-run impacts, in contrast, are useful for reflecting the period of time after structural impacts have

[3] With respect to boundaries, different bodies of literature have at times defined different terms for boundaries, such as "project boundaries" and "leakage" (Aukland, Costa, and Brown 2003), "foreground" and "background" (European Commission 2010), as well as "direct" and "indirect" (U.S. Department of Energy 2023). These can at times be useful categorizations for describing discrete components of impact, but in general do not play a role in the remainder of this document, as consequential impact analysis is understood to be comprehensive, and therefore it is not necessary to draw such boundaries.

occurred, reflecting the combined effects of both operational and structural impacts (Weidema, Frees, and Nielsen 1999; A. D. Hawkes 2014; Gagnon and Cole 2022).

Key terminology: Operational, structural, short-run, and long-run

Operational impacts: Differences in the electric generation from generators, between a scenario with an action and a counterfactual without the action.

Structural impacts: Differences in the capital assets of the electric grid (e.g., generators and transmission lines), between a scenario with an action and a counterfactual without the action.

Short-run analysis: Estimates or analysis performed assuming that capital assets are fixed, and therefore impacts are purely operational.

Long-run analysis: Estimates or analysis performed assuming that capital assets can change in response to an action, and therefore impacts are a combination of both operational and structural.

Prospective, Retrospective, and Real-Time Analyses

It can also be useful to describe three prototypical categories of consequential emissions analysis, which differ in the relation between the timing of the intervention and the performance of the analysis:

- **Prospective analysis (ex-ante):** Analyzing the impact of a future, generally long-lived, action.
- **Retrospective analysis (ex-post):** Analyzing the impact of a past action or set of actions.
- **Real-time estimates:** A signal available in real time that conveys an estimate of the emissions consequences of an action that influences the electric sector, such as increasing or decreasing electric demand at that location and time.

The three categories are suited for different purposes. For example, prospective analysis could guide forward-looking investment decisions for long-lived assets, retrospective analysis could be used as a basis for a hypothetical consequential reporting system, and real-time estimates could in theory be used as a signal to inform operational decisions such as deciding when to charge an electric vehicle fleet.

To be clear, all three analysis categories are considered here to be seeking the same goal (to identify the emissions impact defined in Equation 1). Nonetheless, it can be colloquially helpful to define categories such as this because different models and different data may be more suited for one application than others. For example, many models are useful for prospective analysis, but are not built to inform the other two applications.

Note that the three categories listed previously are not exhaustive, and they can have overlapping features: For example, retrospective analysis could have a prospective component if all impacts of a past action have not materialized when the analysis is conducted, and real-time estimates can be based on short-term forecasts (therefore being a special class of prospective analysis that occurs at a much-reduced timescale).

To gain some further intuition about these three analytical categories—and unique challenges associated with each—three stylized examples are shown below, illustrating possible impacts over time for each of the three analysis types. Each impact trajectory is drawn from a simple model, described in more detail in Appendix A.2. The key assumptions of the model, for the purpose of this section, are 1) any increases in load induce an increase in variable renewable generation that is half the size of the load increase (e.g., a 100 MWh/year load increase induces 50 MWh/year of variable renewable generation), 2) there is a 3-year time lag between a change in load and its influence on the structure of the grid (generator capacities, in this case), 3) any remaining generation needs are met by natural gas generators, 4) load increases during the relevant period of time, and 5) structural decisions are made annually. This simple model is meant to illustrate several key dynamics and provide a starting point for discussing how reality may differ materially.

Note that the discussions that follow—and largely throughout this document—are focused on interyear phenomena, as that is an area that appears to merit more conceptual development. Intrayear phenomena are also important, in the sense that there can be meaningful hour-to-hour variations in the relationship between actions and emissions. Intrayear considerations are less discussed in this document because they are already receiving much consideration in existing literature.

2.2 Stylized Prospective Analysis

First, consider the most straightforward and analytically mature category: a prospective analysis of a long-lived intervention. As an example, consider an analysis being conducted in 2023, for a hypothetical load addition in 2024, that will persist for 20 years (see Figure 2). As mentioned previously, the responses shown are derived from the simple model described in Appendix A.2, where load is served by a mixture of natural gas and variable renewables and there is a 3-year time lag for structural change.

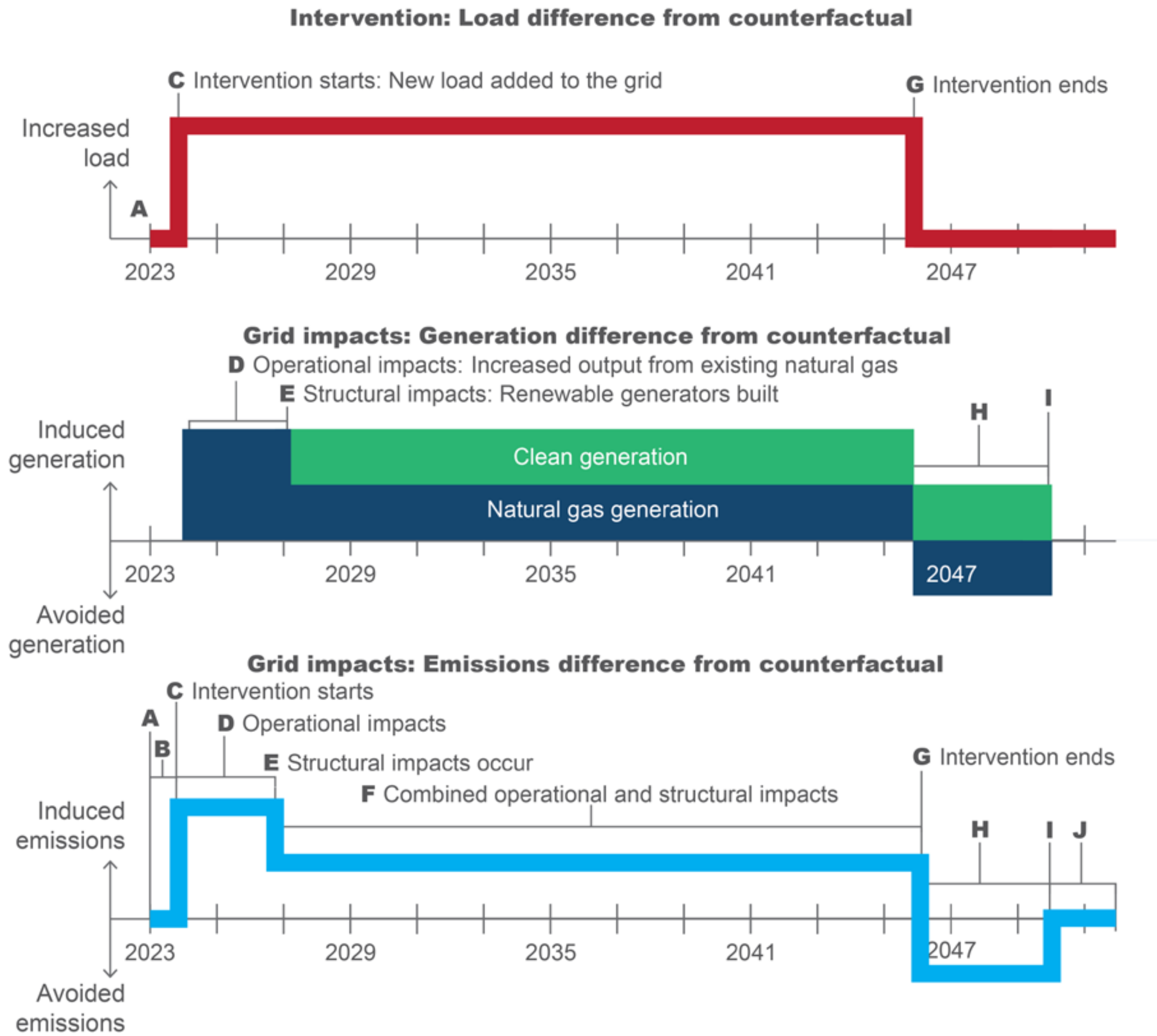

Figure 2. Stylized prospective analysis

Notes:

A. At the beginning of 2023, the prospective emissions analysis is performed, seeking to estimate the emissions impact of the upcoming intervention (a new load).

B. Throughout 2023, no impacts occur, as the intervention has not begun.

C. In 2024, the intervention starts; that is, the load is added to the grid.

D. From the initiation of the intervention through 2026, the load induces emissions at a relatively high rate from the grid because the demand induces generation from existing natural gas generators (under the assumptions of our example). This corresponds to purely operational impacts, in the language defined in Section 2.1.[4]

[4] This stylized example is a load addition and therefore, under the assumptions of this model, induces emissions. These dynamics are relevant for interventions that reduce the demand on the grid as well, however, such as load reductions or generation injections. In such cases, the emissions impacts could be inverted; that is, there would be avoided emissions instead of induced emissions.

E. At the beginning of 2027, 3 years after the load was first added to the grid, the structural response manifests: Given the model's assumptions, variable renewable generators (that would not have otherwise been built) become operational and start to inject clean generation into the grid at a level that is half of the MWh demand from the intervention.[5]

F. From 2027 through 2043—the bulk of the intervention's lifetime—induced generation (i.e., the difference between the grid's generation mixture with the intervention relative to the counterfactual without the generation) is a mixture of variable renewable and natural gas generation, and is therefore a relatively lower emissions rate than the first 3 years.[6]

G. At the end of 2043—20 years after initiation—the intervention is projected to cease (i.e., the load is projected to be removed from the grid).

H. Under the assumption of a 3-year time lag for structural impacts, the grid persists with a greater quantity of variable renewable generation than in the counterfactual for that time, until the resource planners adapt to the removal of the load. During this time, the induced emissions rate is negative relative to what would have occurred at that time in the counterfactual without the intervention because the renewable generators (that were induced by the intervention) would continue to inject electricity into the grid.

I. In 2047, 3 years after the intervention has ceased, the grid returns to the same state that it would have been in absence of the intervention.[7] Under the assumptions here, where there is positive load growth, this adaptation takes the form of slowing down the rate of building new generators (not retiring the induced generators).

J. From 2047 onward, the induced emissions from the intervention are zero because there are no differences between the with-intervention and counterfactual scenarios.

The action being analyzed could be a new activity (such as in the example above)—but it also could be the analysis of the future of an activity that already exists when the analysis is performed (e.g., a load that already exists on the grid, where a decision maker wishes to estimate the consequences of the load continuing to exist). In such a situation, the analysis could be phrased as the action being either the continuation of the activity (where the counterfactual is therefore the situation in which the activity ceases), or the action could be the termination of the activity (where the counterfactual is the continuation of the activity).[8]

The inclusion of structural impacts in the example above raises an issue that challenges electric-sector consequential analysis: There are large step-change impacts (e.g., a power plant investment or retirement, or the unit commitment process for large generators)—and it is often

[5] There are a variety of possible causal chains for this response, depending on the specifics of the actors involved in structural decisions for the local grid region: For example, the load intervention could have lifted energy prices above what they otherwise would have been, inducing independent power producers to complete projects in the interconnection queue that otherwise would have been withdrawn. Another example: A utility regulated to minimize costs may have increased its purchase or investment in renewable generation relative to what it otherwise would have been.

[6] Note that there could be many changes to the grid during this time that would affect the induced emissions rate. For example, if the characteristics of the natural gas fleet were to change over this time, this may impact the induced emissions rate during 2027 through 2043. In addition, there may be piecemeal structural responses during this time period, not just a single response at the 3-year point. This stylized example shows a constant induced emissions rate for simplicity, recognizing that, in practice, it would be likely to vary over time.

[7] In the stylized example given here, the with-intervention and counterfactual scenarios become identical. In practice, at least minor path-dependent differences would likely have occurred, meaning that the two scenarios never become fully identical.

[8] The authors of this paper often receive questions about whether consequential analysis is applicable to both "new" as well as "existing" activities. As described here, it can be applied to both, although analysts should always reflect on whether consequential analysis is the right framework for the question they are asking.

not feasible to assess with precision a particular action's contribution to a particular step-change impact.[9] This paper follows other researchers in considering actions as contributing proportionally to projected step changes (Brander 2016; Hertel et al. 2010; Hsu et al. 2010) while recognizing that this is, strictly speaking, a deviation from the theoretical ideal of exactly determining the consequences of a specific action being analyzed.[10]

The stylized example here assumes there is a time lag for structural impacts. There may, however, be situations where it would be defensible to assume no time lag. For example, if the action being analyzed is the connection of a renewable generator to the bulk power grid, in a situation where there is a bottleneck in the interconnection queue, the processing of the given generator through that queue may be directly offsetting another generator that would have otherwise been deployed at approximately the same time. This would represent an immediate structural impact (i.e., the suppression of a generator that would have otherwise existed). As another example, if an action is known and prepared for by resource planners, there may be a planned structural response that occurs more or less immediately (for example, if a load is anticipated by local resource planners who build generators they would not have otherwise built that come online at approximately the same time as the load).

Lastly, note that the discussion above for Figure 2—and in the following sections—shows load impacts caused by actions. While the load impact of an action can often be straightforward to determine, in many cases it may not be: This is particularly relevant for actions meant to support the deployment of clean generators, where the impact of an action may often not be the generation output of the associated generator (for example, if the project would have gone forward in absence of the action). Because this section is focused on theoretical impacts, this challenge is not discussed further here, but the point is raised in more detail in Section 3 (which focuses on the practical implementation of the theory), and in particular in Section 3.4.

To reemphasize: The previous example is a stylized response to an action, meant to illustrate several key aspects and difficulties, and to set up the discussions in Section 3. The actual response would vary materially between interventions, in practice.

[9] Considerations of time sometimes help mute the severity of this complication—that is, even if a particular action does not induce structural change in the first relevant time step, it may move up an investment during a later time step. Such framing does not make the problem linear, but at least breaks the nonlinear steps into smaller pieces.
[10] Note that such an assumption is reflected in widespread consequential electric sector analyses that use linear programs, a common framework for capacity expansion models (Ho et al. 2021).

2.3 Stylized Retrospective Analysis

The prior subsection started with a stylized prospective analysis because it is the most straightforward. In this section, a stylized retrospective analysis is examined, which is a less developed application for consequential analysis. The chosen example has two particularly difficult features: First, it is assumed that the intervention being analyzed is short enough that a significant portion of its impacts manifest after the action ceases; second, it is assumed that the analysis is being conducted at a point in time when not all of the impacts of the intervention have manifested.

While these features would not always be present (retrospective evaluations can be made of long-lived interventions in the distant past, of course), the example here includes these features because they may reflect challenges in an application such as a hypothetical corporate reporting system (i.e., a regularly updated estimate of impacts for a recently concluded calendar year). For example, if an organization wished to describe the impact of all the activities it undertook in 2024, and was performing the analysis in early 2025, it would likely face these challenges.

Figure 3 shows the stylized example for this retrospective analysis. This example shows an analysis of load that existed during 2024, being analyzed in the first quarter of 2025. As with the prior example for prospective analysis, the response shown in Figure 3 comes from the model described in Appendix A.2, where load increases are projected to induce generation equally from both variable renewable generators and natural gas generators and where there is a time lag of 3 years for the grid to structurally respond to actions. For ease of comprehension, this example reflects a situation where the load came online in 2024 and would persist indefinitely—and where the counterfactual situation is that the same load would instead come online in 2025. However, note that the response shown in Figure 3 could also reflect other situations, such as where the load existed for only 1 year (with the counterfactual being no load).[11]

[11] See Appendix A.2 for more discussion of this point.

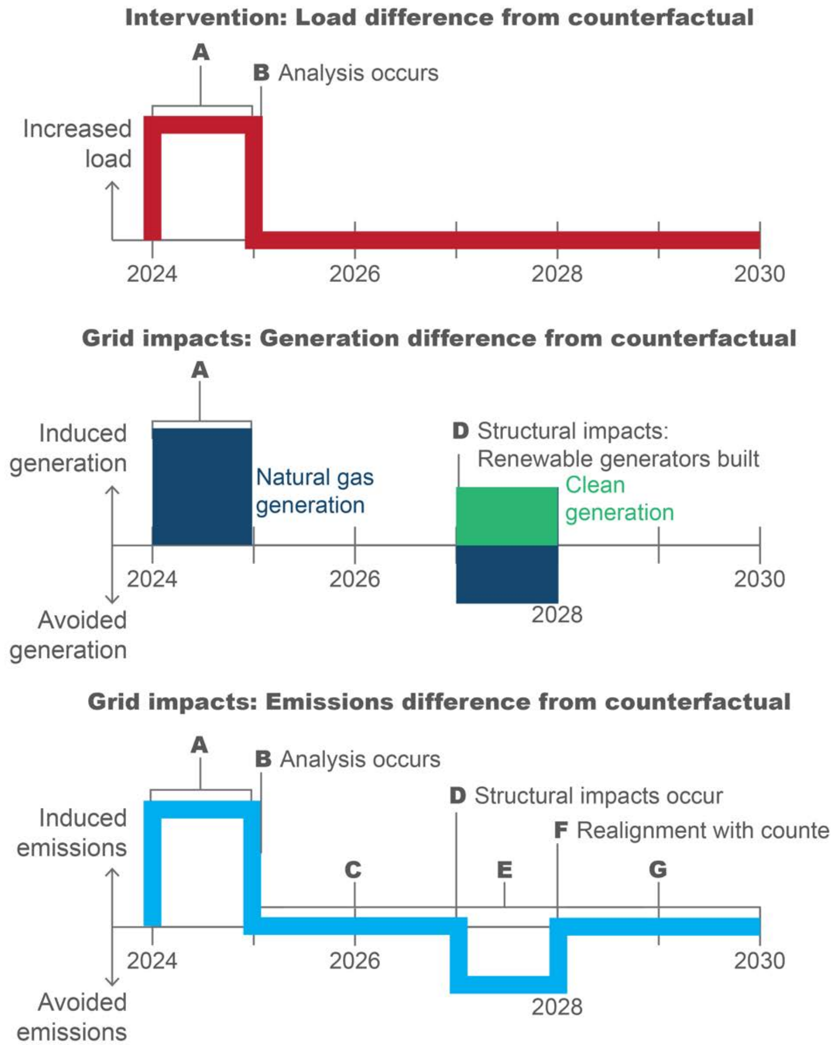

Figure 3. Stylized retrospective analysis

Notes:

A. Load comes online in 2024, drawing electricity from the grid. There is an immediate operational response from the grid, where there is increased generation from natural gas generators, relative to what would have occurred in the counterfactual (here assumed to be the load instead coming online in 2025).

B. In Q1 of 2025, the retrospective analysis of the 2024 load is conducted. Note that the load has continued to exist beyond 2024, so it is assumed that the analyst is specifically seeking to understand the impact of the load that existed in 2024, not the lifetime impacts of whatever is causing the load.

C. From 2025 through 2026, there are no impacts from the intervention because there are no differences between the with-intervention and the counterfactual scenarios (because the intervention is considered to be only the 2024 load, and the structural response from it has not yet occurred).

D. At the beginning of 2027, the structural response to the 2024 load occurs, 3 years after the load appeared. This takes the form of induced renewable generator builds, injecting clean generation into the grid.

E. During 2027, the emissions impact of the intervention is negative because the only impact occurring at that point in time, from the 2024 load, is the injection of clean generation into the grid from the induced renewable generators—which is assumed here to reduce generation from natural gas generators.[12]

F. At the beginning of 2028, the impact from the 2024 load ends because the with-intervention and counterfactual scenarios become identical (i.e., because the counterfactual has load appearing in 2025 instead of 2024, the appearance of the renewable generators is 1 year later, after which there is no difference between the scenarios).

G. From 2028 onward, for this stylized example, the intervention has no further impacts, because the grid has returned to a state that is identical to the counterfactual.

This example was chosen to analyze the initiation of a new load, with a counterfactual of the same load starting 1 year later, because the impacts in such a situation are relatively clean and therefore an easier starting point for comprehension. Other similar retrospective assessments would face significantly greater conceptual challenges, for example, the retrospective analysis of load in 2024, if it was part of an ongoing load that has started before 2024, and would persist after—or the analysis of load that existed *only* in 2024. Although either situation could be reflected with the model used for Figure 3, the assumptions given above may not be appropriate, as they depend in nuanced ways on the specific process by which operational and structural decisions are made (which vary materially by time, type of intervention, location, and other relevant parameters), and may depend in more pronounced ways on the surrounding context (e.g., whether the load is part of an ongoing trend of transient loads).

There are many situations where the impact of an intervention would differ from what was seen in Figure 3, for example, with transient interventions under asymmetric structural response times.[13] The point of Figure 3 is not to describe a universally applicable model, but to illustrate several key aspects of one form of retrospective analysis.

One attractive feature of the stylization shown in Figure 3: If a retrospective assessment was performed with this model year-over-year, the cumulative reported impact would be identical to a prospective assessment (assuming that the future unfolded as projected). This is illustrated in Figure 4, which shows the effect of adding only 5 annual assessments—but the reader can intuit how the addition of 20 such assessments would replicate the shape of the stylized prospective analysis shown in Figure 2. Whatever stylized pattern of impacts ultimately proves most useful

[12] Note the cumbersome language here: The impact is a reduction of emissions from generators, which is defined as an operational impact above, yet it was caused by the building of renewable generators—a structural impact. Colloquially, we often refer to the time-lagged portion simply as a structural impact, although it would be more accurate to describe it as both.

[13] Analysis may implicitly be reflecting a transient intervention at counterintuitive times. For example, if an analyst wished to understand the impact of load in 2024, which was part of an ongoing load that extended before and after 2024, the intervention could be phrased as the transient removal of the load in 2024—with both the with-intervention and counterfactual scenarios both representing the load outside of 2024. An analyst would, in effect, be asking, "What would have occurred if we had turned off this load, just for 2024?" This naturally raises many difficult conceptual and analytical issues, such as the myriad short-run phenomena that could manifest if an otherwise persistent action was actually terminated for a single year. Ultimately, there may be an argument on practical grounds to stylize this type of analysis (i.e., to use long-run values instead of seeking to precisely estimate the short-run dynamics—a deviation from the theoretical ideal, but one that may be justifiably useful).

for conceptualizing retrospective assessments, it seems possible that an approach that can cumulatively reflect a project's total impact may be desirable.

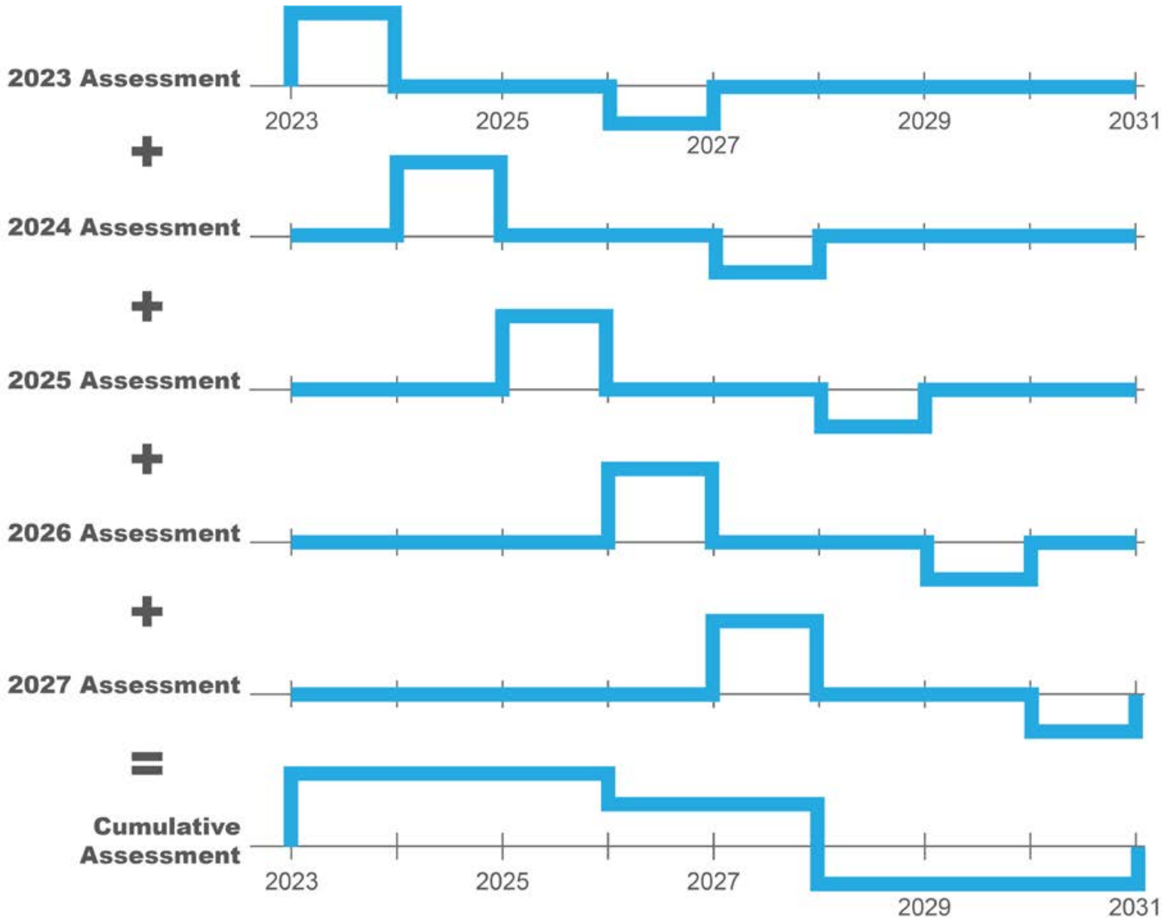

Figure 4. Cumulative retrospective assessments

Figure 3 and Figure 4 help show an important point: To omit structural impacts would be to systematically err in the retrospective estimate of the impact of actions (in the same manner that it would systematically err to ignore structural impacts for any other type of assessment). Given that structural impacts are usually associated with clean generation, whereas operational impacts are usually associated with fossil generation, a systematic exclusion of structural impacts could materially distort estimates (Gagnon and Cole 2022).

The implications of omitting structural impacts are particularly important for estimating the impact of actions meant to support the deployment of clean generators: Structural impacts are one way in which the system-level impact of a particular intervention would be less than the associated project's nameplate output, through phenomena such as economic competition or replacement effects. This is discussed in more depth in Section 3.4.

2.4 Stylized Real-Time Analysis

For completeness, this section briefly discusses a stylized real-time analysis (see Figure 5)—although this class of analysis is even more underdeveloped than the preceding retrospective section. The following example considers a brief load addition (e.g., choosing when to charge an electric vehicle on a particular day) that takes place during 2023, guided by an analysis that occurred in real time or shortly before the decision was made (e.g., based on data from system operators or short-term forecasts of grid conditions). As with the prior two stylized examples, this example is drawn from the simple model described in Appendix A.2, which assumes that load increases are projected to induce generation from both variable renewable generators as well as fossil generators, that there is a time lag of 3 years for the grid to structurally respond to an intervention, and that structural decisions are made annually.

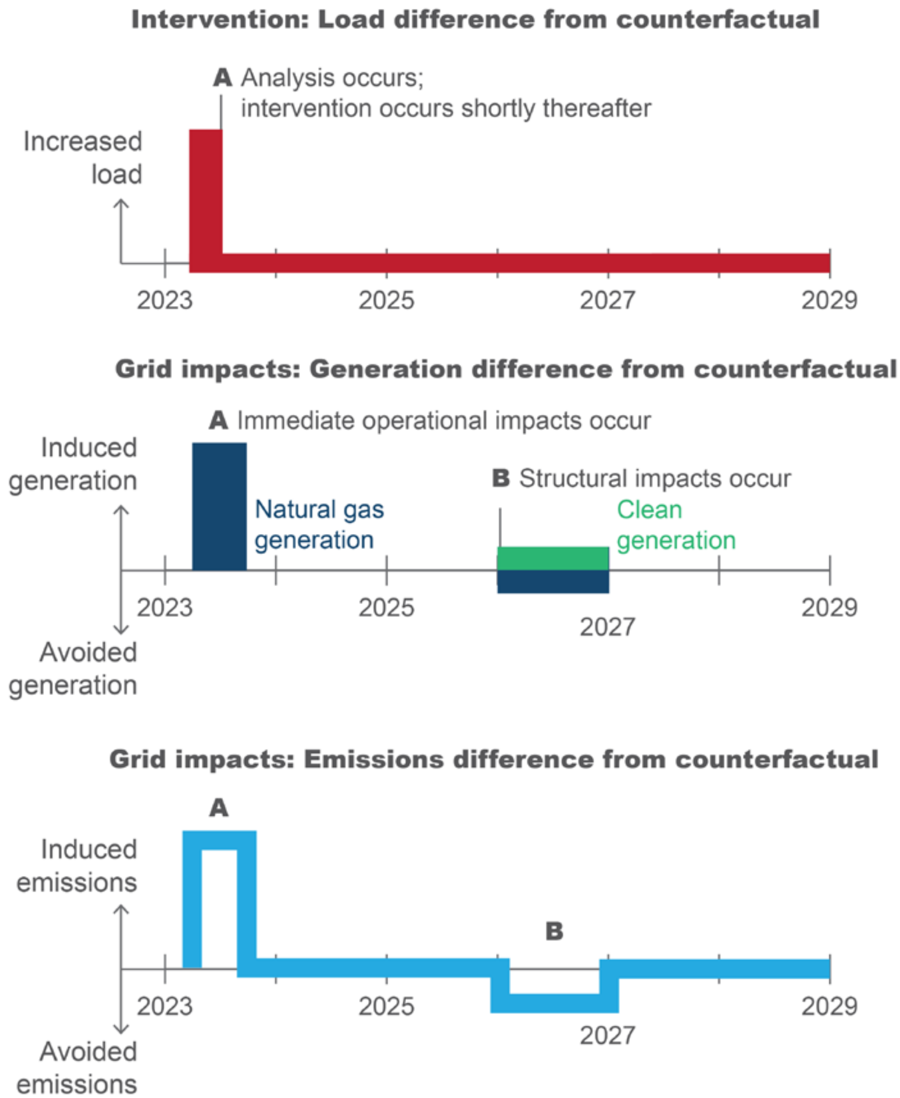

Figure 5. Stylized real-time analysis. Duration and magnitudes of impacts not to scale.

A full walkthrough of Figure 5 is not provided, as the dynamics are largely the same as what has been discussed in the prior section. The only element to clarify is that, in this example, it is assumed that a small load addition (A) in 2023 influences the metrics used for resource planning (e.g., data on load, prices, and so forth), and therefore has a small time-lagged contribution to structural decision making (which manifests during B).[14]

A critical difficulty with this analytical class—and the reason the authors consider it underdeveloped—is the estimation of the relationship between A and B (i.e., between a momentary action and the time-lagged structural impact of that action). This is particularly difficult because the relationship would be expected to change moment-to-moment depending on the conditions of the grid at any point in time (e.g., charging an electric vehicle during a windy period would be expected to influence wind investment to a greater degree than charging an electric vehicle several hours later when the wind has materially decreased in strength).[15] No method for making such an estimate in real time currently exists, and certain important points remain without consensus among researchers (such as whether transient events influence structural decisions and, if some but not all do, what differentiates them).[16]

The current state of the art for this analysis class is discussed further in Section 3, including a brief discussion of the possibility that, given the unique challenges associated with this task, it may ultimately be more defensible to use heuristics for real-time operational decisions, rather than attempting to estimate the type of phenomena shown in Figure 5.

[14] As throughout this document, this assumes that small actions proportionally contribute to structural change. This is a simplification of reality, which would likely involve large step changes.

[15] Another difficulty, not reflected in Figure 5, is that, in contrast to retrospective analyses, these choices may be sufficiently short (e.g., a single action that may persist for minutes or hours) that there is a materially greater likelihood of intertemporal difficulties on the operational scale. For example, an action may influence what generators are committed (i.e., turned on) and therefore the generation mixture in the hours following the intervention—that is, a strictly operational impact that persists beyond the end of the intervention itself. We are not aware of any efforts to estimate such phenomena in real time.

[16] The relationship between transient interventions and structural impacts has been discussed in the literature (Miller 2022). This is a difficult conceptual challenge—plausible arguments can be made that certain transitory interventions would be unlikely to cause structural change, whereas others would. For example, a nuclear power plant shutting down for a month of maintenance may be unlikely to influence structural decisions (because it was a known transient event to planners), whereas each individual vacation to Las Vegas is a transitory action, but the sum total of such events seems virtually certain to have impacted the generation capacity in the region. At a minimum, it does not seem possible to define whether an action contributes to structural change based solely on its duration, as such designations would ultimately depend on the context (e.g., it is part of a pattern of other transient events), and the decision-making workflows of planners (i.e., what exact algorithms they use to process historical timeseries data). Our view is that it is likely generally defensible to consider all actions as proportionally contributing to ongoing structural decision making in a region, unless there is a specific demonstration of why an action would be likely to not do so—but we recognize that this is an area of open discussion.

3 Using Marginal Emissions Rates To Estimate Consequential Impact

In the prior section, several stylized examples were presented to highlight some key theoretical aspects of consequential impact assessments. In this section, two methods are discussed for putting the analysis into practice using marginal emissions rates—as well as the known limitations of those methods.

Note that prospective analysis of long-lived interventions can often be conducted with power sector models such as capacity expansion models (alone or in combination with marginal emissions rate analysis). Model-based approaches can often represent phenomena or dynamics that are not reflected in marginal emissions rates. As one example, marginal emissions rates require impacts to be linearized for a representative perturbation (either the infinitesimal marginal or an incremental perturbation), whereas models can incorporate nonlinear phenomena, resulting in estimates that are generally more defensible for large interventions. Although there is a theoretical basis for model-based predictions (cost minimization, reflecting either a well-functioning market or well-regulated organization) and such models are widely used in applications such as policy analysis (Bistline et al. 2023; Steinberg et al. 2023), it is nonetheless important to note that as a class, the accuracy of such models is not readily validated through empirical scientific means. Therefore, conducting widely defined scenario analysis and maintaining an awareness of uncertainty is generally advisable when using such models.

In theory, real-time and retrospective assessments could also be based at least in part on models similar to capacity expansion models, although no models built for those purposes currently exist to the author's knowledge—and using existing models (that were built for other purposes) for either retrospective or real-time estimates can lead to material errors.[17] For example, a model built primarily for multidecadal forward-looking analysis might perform relatively worse when used for retrospective estimates because of the omission of phenomena that may be highly relevant in the short term but anticipated to be less so in the long term.

At any rate, this section of the document focuses on marginal emissions rates as a possible route to consequential emissions analysis that can scale (i.e., be applied widely at tractable cost)—whereas the complexity and nuance of power system modeling potentially precludes them from widespread use.

By focusing on marginal emissions rates, this section narrows the scope of what actions can be evaluated. Marginal emissions rates can be suitable for estimating interventions whose influence is predominately changing the demand for grid electricity (such as increasing/decreasing load, shaping load, or injecting electricity into the grid), and whose impacts can be reasonably approximated with linearized values. There are many other types of interventions, such as building transmission lines or facilitating administrative improvements to interconnection queue processes. While all interventions ultimately can be reflected in consequential style analysis, as reflected in the fundamental counterfactual equation given at the start of Section 2, marginal emissions rates are not suitable for characterizing interventions whose influences are not

[17] For example, capacity expansion models generally do not endogenously represent the administrative processes of the generator interconnection. Omitting such a phenomenon for a retrospective or near-term short-lived prospective analysis could result in a projection that was not administratively feasible for the given time.

predominately tied to marginal perturbations of the demand for grid electricity. Such non-marginal actions are often possible to evaluate with power sector models.

3.1 Marginal Emissions Rates: General Considerations

Marginal emissions rates refer to estimates of changes in emissions that result from unit changes in the demand for grid electricity.[18] They can be either *a priori* estimates developed to inform upcoming decisions (Gagnon et al. 2024; A. D. Hawkes 2014; IFI 2022; Energy+Environmental Economics 2016), or they can be developed and applied retrospectively (Oates and Spees 2022; Azevedo et al. 2020). The basic principle of this approach can be expressed in the following equation:

Equation 2

$$\begin{array}{c} \textit{consequential} \\ \textit{emissions impact} \\ \textit{estimate} \end{array} = \underbrace{\Delta electricity}_{\substack{\textbf{impact of action on} \\ \textbf{the demand for grid} \\ \textbf{electricity}}} \times \underbrace{\frac{\Delta emissions}{\Delta electricity}}_{\textbf{marginal emissions rate}}$$

Marginal emissions rates may be developed through statistical methods, reported by system operators, or developed with power sector models. Each approach has its own strengths and weaknesses. Methods purpose-built to leverage data from system operators are likely significantly more accurate for estimating the immediate operational impacts of actions, while being unable to reflect structural impacts. In contrast, marginal emissions rates created with power sector models are generally relatively poorer at estimating immediate operational impacts, but can estimate structural impacts.

Geographic resolution can be impactful for marginal emissions rates because of phenomena such as transmission congestion, variations in renewable energy resources, and variation in existing generator fleets. In general, between two hypothetical data options of equal quality, it would be preferable to select the finest geographic resolution available.[19,20] In practice, one would rarely encounter two datasets of identical quality but different resolution, and therefore an analyst may

[18] The term “marginal emissions rate” implies an infinitesimal perturbation. In practice, the term “marginal emissions rate” sometimes reflects an incremental emissions rate, such as the estimated change in emissions resulting from a 1-MWh increase in electricity demand or greater. Given that the magnitude of the response to a given signal can be quite large, the most suitable step size for calculating marginal emissions rates may not be an infinitesimal perturbation, and may vary by situation. This is an ongoing research topic for which general consensus has not been achieved.

[19] For example, transmission congestion can influence short-run marginal emissions rates in a manner that generally makes nodal estimates more accurate than regional estimates.

[20] At risk of misinterpretation, we clarify that this statement is intended to apply only to marginal emissions rates. The appropriate geographic resolution for other metrics, such as attributional metrics, is a balance of factors that does not, to our knowledge, have a generalized answer. Many attributional emissions metrics would not benefit from fine geographic resolution.

be faced with a choice between higher geographic resolution but potentially ambiguous differences in quality.

Similarly, it is typically preferable to use hourly data or finer, given material variations in marginal emissions rates hour-to-hour. If hourly data do not exist, the equation can be used with more coarse data (e.g., annual), but users should recognize that such an approach likely degrades the accuracy of the analysis, given diurnal and seasonal variations in marginal emissions rates. The data can be represented as a timeseries (e.g., hour-over-hour, year-over-year) or as averaged or levelized values for the time period under analysis.[21]

Performing an analysis using marginal emissions rates requires obtaining data on the change in demand for grid electricity for the given intervention as well as the marginal emissions rate. Demand change estimates can be obtained from project developers (because that information is often available as part of the engineering design process), by identifying a similar project that has been conducted previously and using its performance, or by using various tools for the type of intervention at hand (for example, the National Renewable Energy Laboratory's [NREL's] System Advisor Model can give production estimates for different types of renewable generation technologies).

Importantly, note that it is necessary to rigorously develop the counterfactual, to reflect what demand perturbation is caused by the intervention being analyzed (Gillenwater 2012). In many cases this is straightforward—and in many cases, it is not.[22] For example, it is generally not suitable to make an *a priori* assumption that the demand perturbation resulting from actions meant to support the deployment of clean generators (such as the purchase of RECs or signing of long-term offtaker agreements) is the output of the associated generator. This is related to the concept often referred to as "additionality," discussed in more detail in Section 3.4. The discussion below in Sections 3.2 and 3.3 is predicated on having an appropriately designed counterfactual (i.e., having a defensible estimate of the demand perturbation *caused* by the action, not simply *associated* with it).

Sections 3.2 and 3.3 below discuss two approaches to estimating a marginal emissions rate. The two approaches are similar but bring slightly different component data types to bear. Both approaches implicitly recognize the importance of considering both operational and structural impacts, although they use different data to do so. Neither approach (based on data that exist as of publication of this paper) completely reflects any of the three stylized analysis types described previously in Section 2, and therefore this should be understood as a discussion about possibly defensible ways to use the types of data that currently exist—not a method that exactly describes all relevant phenomena.

Which approach is more suitable depends in large part on data availability: Approach #1 relies on an *a priori* effort with power sector models to develop emissions factors and, by virtue of that modeling, can often produce more defensible estimates than the following method. However, the

[21] Averaging or levelization is often defensible, although it can give rise to errors if there is a difference in how the emissions rates and loads evolve over time.
[22] An example of a straightforward case: If an organization is deciding whether to perform an energy efficiency upgrade, the estimated reduction in demand can usually be taken at face value as the consequences of the action, because the business is effectively in full control of the relevant aspects of demand in both sides of the counterfactual.

data that underpin Approach #1 currently (in 2024) are not widely available. Approach #2 has a tractable method for estimating values with sparse data but relies to a great extent on judgment and heuristics, and is therefore likely not suitable for many purposes. As discussed in more detail below, both methods are most developed for long-term prospective analysis, and are less developed for either retrospective or real-time use cases.

3.2 Marginal Emissions Rate Approach #1: Short-Run and Long-Run Margins

The first approach for using marginal emissions rates to estimate the impact of electric sector interventions has been described previously in Gagnon, Hale, and Cole (2022) and Miller (2022):[23]

Equation 3

$$MER = SRMER \times w_t + LRMER \times (1 - w_t)$$

Where:

- *MER* is the composite marginal emissions rate for the action being analyzed
- *SRMER* is an estimated short-run marginal emissions rate (i.e., the marginal emissions rate assuming a fixed set of capital assets)
- *LRMER* is an estimated long-run marginal emissions rate (i.e., the marginal emissions rate from a grid that has structurally responded to an intervention, thereby including both operational and structural effects)
- w_t is a weight designed to reflect the relative influence of the SRMER and LRMER terms, based on how long it is expected to take for the structural response to the intervention to occur.

The resulting composite MER estimated in Equation 3 can then be multiplied against the change in the demand for grid electricity (Equation 2) to gain an estimate of induced or avoided emissions.

Long-Run and Short-Run Approach: Prospective Analysis

The concept behind this approach can be understood by referring to the stylized prospective analysis shown previously in Figure 2, the emissions-impact portion of which is duplicated for convenience below. As shown in Figure 6, there may be an initial period (Segment D) where actions may largely induce purely operational changes, which can often be estimated with SRMERs. After a time lag, structural impacts occur, and the induced emissions can be estimated with LRMERs (Segment F).

[23] As described here, it is assumed that an analysis is looking for a single value or vector of values that represents all of the years covered by an analysis—which is common, because many analyses do not have year-over-year MWh demand projections but simply a single annual value or hourly vector. If an analysis is instead explicitly resolving each year, the SRMER and LRMER values can be employed alone as appropriate for each year in the analysis. This avoids the need to calculate a weight while representing the same concept in the choice of how many years to apply the SRMER versus LRMER.

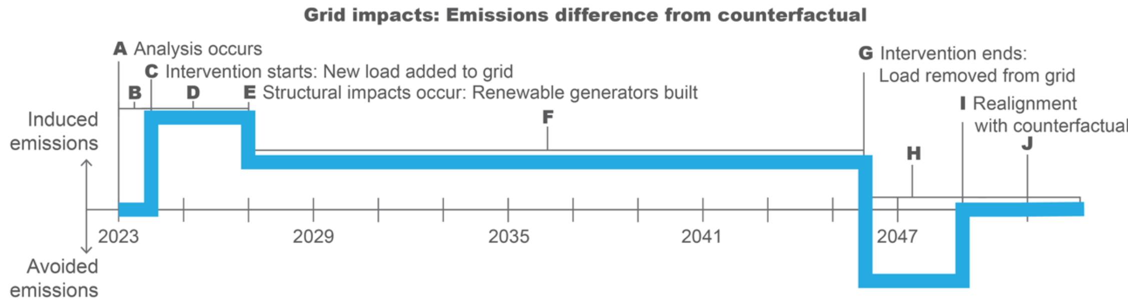

Figure 6: Stylized emissions impact for prospective analysis (duplicated for reference)

Equation 3 uses the w_t term to combine the SRMER and LRMER estimates; w_t can reflect the estimated share of time before material structural impacts start to materialize. For example, in Figure 6, the SRMER portion is 3 years, whereas the LRMER portion is 17 years. If the analyst is indifferent to the timing of emissions, this would equate to a w_t of 0.15.[24]

Note that, as described, this approach omits the impacts that occur after the intervention has ended (Segment H). For sufficiently long-lived interventions, this may be justifiable. For short-lived interventions, where the Segment H impacts are nontrivial, it may not be acceptable to ignore them—although no method has been described, based on data that exist today, to capture Segment H through marginal emissions rates.[25]

A key feature of this method is that, as defined previously, LRMERs reflect the induced or avoided emissions from all generators—those whose capacity is influenced by the intervention (built or retired), and those whose capacity is not influenced. Therefore, LRMERs are not equivalent to the build margin described in the following section, which describes only the emissions rate of generators whose existence is modified by the intervention. This is the primary reason why this document recommends this method as preferable over the method described in the following section: LRMERs leverage a model-based workflow to estimate the combined result of both operational and structural impacts in Segment F, whereas the following method relies on heuristics to do so.

Using SRMERs and LRMERs and blending them with a single transition point, as described here, is a simplification. For example, real electricity systems are never in true economic long-run equilibrium (an assumption underlying the LRMER estimations known to the authors), and in reality, there would not be a single transition point where structural change takes hold but more likely a stepwise response over time. In addition, the w_t value would likely vary year-to-year as well as by intervention and region. To the author's knowledge, no research exists providing guidance on appropriate values for w_t.

[24] If the analyst is not indifferent to the timing of emissions, they may wish to use time-value-of-damages equations to weigh time steps differently. For an example, see Gagnon (2024).

[25] We note that, if the structural impacts are assumed to occur without a time lag (either because that is a defensible assumption on its own merits, or possibly as an analytical convenience as a means to approximate the impact of Segment H), then the entire impact could be described with a LRMER. This is appealing for its simplicity and its ability to approximate the influence of Segment H, although we also note that LRMERs are currently generated with capacity expansion models, which tend to be relatively poor at reflecting near-term conditions—so the use of only a LRMER may lose an opportunity to leverage higher-quality near-term SRMER data based on empirical observations. Which way the balance of issues tilts here is not immediately apparent to us.

To reiterate a point raised in Section 3.1: This approach is predicated on accurately estimating the demand perturbation caused by the action being analyzed, not merely associated with it. In addition, it is predicated on employing a marginal emissions rate that adequately captures relevant phenomena. Section 3.4 discusses some of the current shortcomings of marginal emissions rate data products when used to estimate the avoided emissions from actions meant to support the deployment of clean generators as well as heuristics that can help identify situations where their application may lead to erroneous impact estimates.

Long-Run and Short-Run Approach: Retrospective and Real-Time Analysis

Previously, Approach #1 was described for a long-lived prospective analysis. A reader, recalling the patterns of a stylized retrospective (Figure 3) and real-time (Figure 5) analysis will note a key difficulty: For both of those applications, material structural impacts may occur after the end point of the action being analyzed (the emissions impact pattern from Figure 3 is duplicated below for convenience). This is the same phenomenon as mentioned above for Segment H of prospective analysis, but where its importance is greatly enhanced because of how short the duration may be for certain retrospective and real-time analyses.

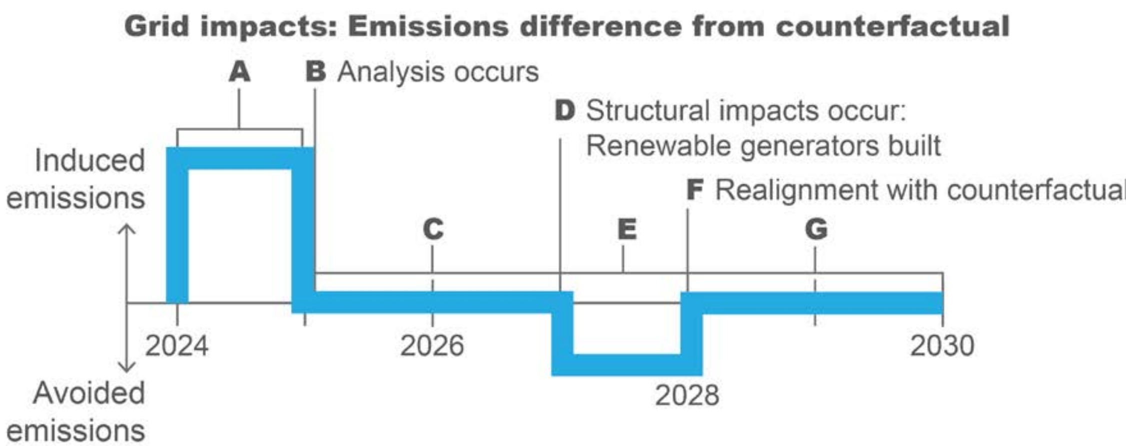

Figure 7: Stylized emissions impact for retrospective analysis (duplicated for reference)

No marginal emissions rate data product is currently built to reflect Segment E in Figure 7, or its equivalent (Segment B) in real-time analysis (because they are both impacts that extend beyond the temporal bounds of the action itself). These segments are equivalent to Segment H that was shown previously for prospective analysis. Segment F in prospective analysis is not equivalent to Segment E in retrospective analysis, because Segment F represents a period when the intervention was still occurring, and therefore could be estimated with LRMERs.

If an analyst wishes to make the best effort possible at a retrospective evaluation with marginal emissions rates, despite the above difficulty, one option is to rephrase the question being asked and use the result as an approximation of the actual question. Specifically, instead of analyzing just the single year in question (i.e., trying to use marginal emissions rates to construct the pattern seen in Figure 7), an analyst could use marginal emissions rates to estimate the consequence of the historical action *if it were to persist for a long period of time.* For example, building from the example shown in Figure 7, the analyst could estimate the consequences of the 2024 load persisting for, say, 10 or 20 years, under a continuation of the grid conditions in

2024.[26] Such a rephrasing would give the retrospective analysis the same characteristics (i.e., the shape of the emissions impact) as the long-term prospective analysis, where a combination of SRMERs and LRMERs could be used. The impact could be annualized and used as an approximation of the historical action's impact.

There are many caveats associated with this approach, a key one being that no LRMER dataset has been produced with this application in mind, and therefore the accuracy of the approach is uncharacterized—and given that existing LRMERs were built for long-term prospective analysis, the error in applying them to single-year retrospective analysis could be significant. In addition, the assumed duration would be important (the shorter the extension, the greater weight given to the short-run period), but no guidance exists for a defensible duration to assume in this situation.

Another approximation method would be to employ a LRMER only, potentially justified by noting that the Segments E and A combine to approximately a LRMER, if the analyst is indifferent about the timing of emissions. As with the above approximation, the lack of any LRMER calculated for this purpose is a significant obstacle.

Given the above discussions, it would generally be advisable to consider either of the above approximations to be estimates of unknown accuracy—which may be suitable for some, but not all, retrospective applications.

In theory the same approximations could be applied to real-time analysis. In practice, however, no real-time estimates of LRMER exist, nor are the authors aware of any publicly stated efforts to create such a dataset. Given that real-time analysis is typically interested in hour-to-hour variations, and that LRMERs would vary materially hour-to-hour based on the actual conditions of the grid (e.g., weather and demand patterns), using *a priori* calculated LRMERs would lead to estimation errors of unknown, but potentially significant, magnitude. As emphasized previously, omitting structural impacts (i.e., resolving only Segment A from the stylized real-time response) would generally systematically err in estimation—particularly because structural impacts convey diurnal and seasonal trends that can differ from short-run trends. On balance, given the difficulty of real-time estimation of induced structural change and the fact that such estimates are not readily amenable to scientific validation, it is possible that many users will ultimately find it more suitable to their purposes to use heuristics to guide real-time operational decisions, rather than hypothetical future attempts to produce real-time composite marginal emissions rates.[27]

Long-Run and Short-Run Approach: Data Limitations

As emphasized throughout this paper, a material shortcoming of any model-based approach to estimating induced structural change (as is implied in the creation of LRMERs using capacity expansion models) is that they, as a class, have not undergone empirical validation. While uncertainties can generally be expressed through scenario analysis, not all decision-making workflows are amenable to impacts being presented as ranges of possible values.

[26] An argument could be made that the duration should be infinite.
[27] Research is ongoing at NREL to study the performance of heuristics (e.g., renewable energy fractions, energy prices, and so forth) in their efficacy at guiding real-time operational decisions in light of minimizing greenhouse gas emissions. Publication forthcoming.

In addition, current methodologies for estimating sub-annual (e.g., hourly) LRMERs have known deficiencies. For example, Gagnon and Cole (2022) saw that a method that uses a scalar perturbation across all hours as a basis for calculating hourly LRMER values, while being able to directionally reflect differences in differently shaped interventions, nonetheless was not perfect at anticipating the induced emissions when the interventions were modeled directly (for the suite of interventions analyzed, there was a root-mean-square error of 131 kg/MWh).[28] The method studied in Gagnon and Cole (2022) is similar to the method used in the creation of LRMER data in Gagnon et al. (2024) and is therefore likely present in the LRMER data in those datasets as well. While not a fatal flaw for many purposes, methodological shortcomings such as these degrade the accuracy of estimates using these types of marginal emissions rates, and may make them unsuitable for some purposes.

Finally, a major drawback of the approach described in this section is that LRMER estimates are not widely available globally. Because of this, the following section describes a second approach that relies to a greater extent on heuristics but whose application is often tractable even when very limited data are available.

3.3 Marginal Emissions Rate Approach #2: Build and Operating Margins

The second marginal emissions rate approach extends the framework from the Greenhouse Gas Protocol's (GHGP) Guidelines for Quantifying GHG Reductions from Grid-connected Electricity Projects (Broekhoff 2007).

The GHGP's guidance defined the concepts of *operating margins* and *build margins*. The framework was originally defined for long-term projects that reduced greenhouse gas emissions, and therefore is defined in terms of reductions. The operating margin was defined as the weighted average emissions intensity of existing generators whose operations would be reduced in response to the action being analyzed, whereas the build margin was defined as the weighted average emissions rate of avoided generation from new generators that would not be constructed because of the intervention. To extend the framework into other applications (such as the prospective, real-time, and retrospective categories described in Section 2), this paper amends this guidance with two minor adjustments.

The first adjustment is simply to note that the concepts can apply to actions that either increase or decrease emissions, by either increasing or decreasing the demand for grid electricity relative to the counterfactual—and to modify the definitions accordingly.

Next, the definition of the operating margin is slightly changed to be the weighted average emissions rate of the change of generation from any generator whose capacity is not expected to

[28] This is caused by interhour phenomena (i.e., actions during one hour of the year can influence emissions during other hours). It is expected that interventions whose demand perturbation varies materially over the year (e.g., demand response; electric vehicle charging) are more susceptible to this type of error than relatively flatter interventions (e.g., flat blocks of load; energy efficiency measures). Notably, wind and solar generators—whose generation patterns can vary materially from hour to hour—are likely more susceptible to this type of deficiency than other, flatter interventions. In practice, the use of a LRMER (derived from a capacity expansion model using a scalar perturbation method) to calculate the avoided emissions from a wind or solar project may often produce higher estimates of avoided emissions, relative to what would have been estimated if the project was directly modeled in the capacity expansion used to create the LRMER.

be modified by the intervention being considered, whereas the build margin is the weighted average emissions rate of the change in generation from any generator whose capacity is expected to be modified by the intervention being considered (inclusive of new builds, retirements, uprates, and so forth).[29]

With these slight modifications, the definitions used in this paper are as follows:

Equation 4

$$MER = operating\ margin \times (1 - w_b) + build\ margin \times w_b$$

Where:

- *MER* is the composite marginal emissions rate for the action being analyzed
- *operating margin* is the generation-weighted average emissions rate of generation changes from the counterfactual, for generators whose nameplate capacity would not be impacted by the action being analyzed
- *build margin* is the generation-weighted average emissions rate of differences in generation from the counterfactual, from generators whose nameplate capacity would be impacted by the action being analyzed
- w_b is the weight (varying from 0 to 1) of the build margin. A value of 0 would imply that the intervention does not impact the nameplate capacity of any generators, whereas 1 would imply that induced or avoided generation would come only from plants whose nameplate capacity would be impacted by the intervention, and anything between would imply a mixture of both.

The resulting composite marginal emissions rate estimated in Equation 4 can then be multiplied against the change in demand for grid electricity caused by the action (Equation 2) to gain an estimate of induced or avoided emissions.

Build and Operating Margins Approach: Prospective Analysis

Because Equation 4 is defined as the total effect of the perturbations of all generators, it can at least in theory, be interpreted in a way that reflects the phenomena shown in all three stylized analysis types shown previously in Section 2. Start by considering the emissions impact of the prospective analysis shown previously in Figure 2, duplicated below in Figure 8 for convenience.

[29] These generalizations capture a phenomenon whose categorization was potentially ambiguous in the original GHGP's definitions, such as the change in generation from existing generators that would be projected to retire in response to the intervention being considered (i.e., the absence of a generator in the with-intervention scenario, that is present in the counterfactual scenario).

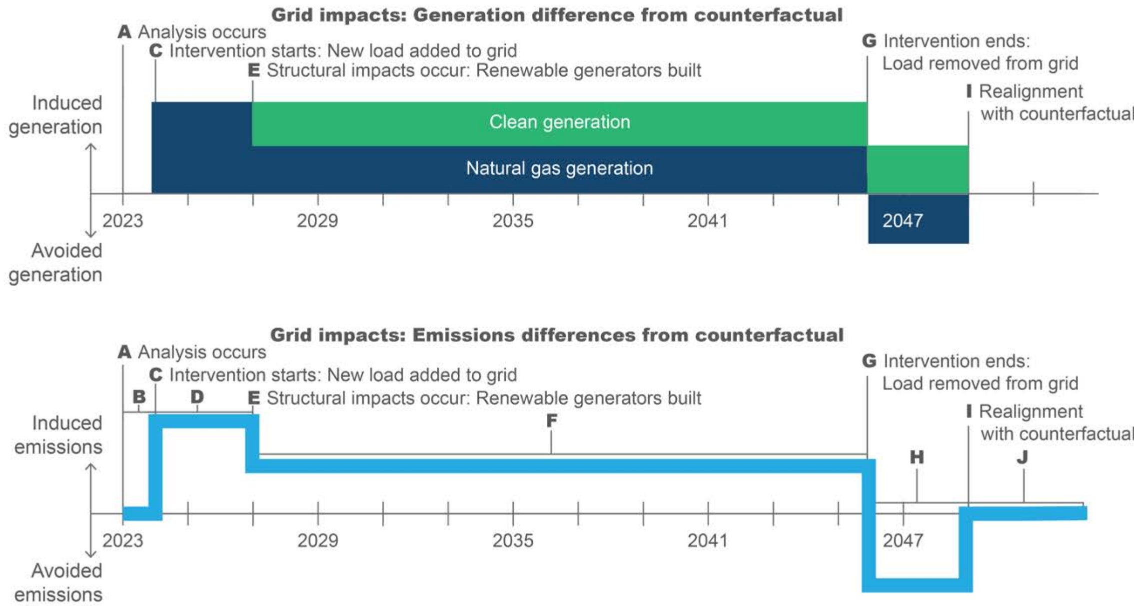

Figure 8: Stylized impacts for prospective analysis (duplicated for reference)

Given the (simple) assumptions used to create this stylized example, the operating margin is known to be exclusively the fossil generation (400 kg/MWh), whereas the build margin is known to be exclusively the clean generation (0 kg/MWh)—and that the net generation quantities of the two are identical (therefore, w_b is 0.5).[30] From this, the composite marginal emissions rate for the duration of the intervention is 200 kg/MWh.

Note that the operating margin of Approach #2 and SRMER of Approach #1 often (but not always) manifest as similar phenomena, but that the build margin and LRMER are not equivalent. The build margin is generation *only* from induced or avoided generators. Examining Segment F of Figure 8: Approach #1 (Section 3.2) uses a LRMER to convey the combination of both the blue *clean generation* and purple *fossil generation* elements, whereas this method separately uses the build margin to convey the *clean generation* and the operating margin to convey the *fossil generation.*

Build and Operating Margins Approach: Retrospective and Real-Time Analysis

The above discussion applied the build and operating margin approach to long-lived prospective analysis. As throughout this document, the extension of this method to retrospective and real-time analysis is possible in theory but faces challenges in practice.

To illustrate, consider the stylized retrospective analysis previously shown in Figure 3, duplicated below in Figure 9 for convenience. The responses underlying the emissions impact shown are all generator perturbations that fit into the definition given above for build and operating margins: Segment A is an increase in existing generator output (operating margin), whereas Segment E is the combined effect of an induced generator's output (build margin) and a simultaneous decrease in existing generator output (operating margin).

[30] The build and operating margins being entirely clean and fossil, respectively, are a product of the intentionally simple assumptions underlying this example—in practice, both the build and operating margin could have a mixture of emitting and nonemitting generation.

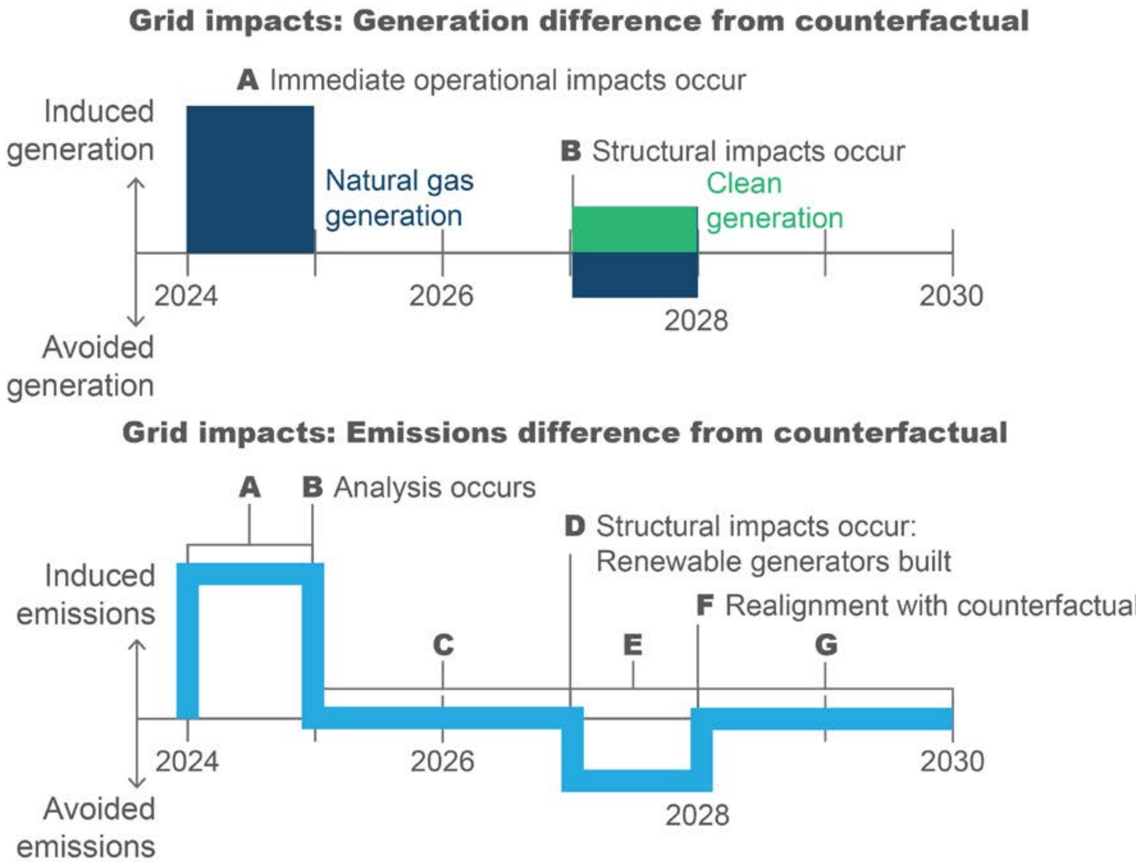

Figure 9: Stylized impacts for retrospective analysis (duplicated for reference)

Note that attempting to construct the shape shown above may be cumbersome—for example, having an analyst combine both positive and negative operational perturbations, with a time lag.[31] In response to the complexity, the same simplifications mentioned for the prior SRMER/LRMER method could be adopted: interpreting retrospective and real-time interventions as long-lived, thereby minimizing the negative components. A disadvantage of this approach is the same as for the prior method: It degrades the accuracy of the method to a largely unquantified degree, and, to the authors' knowledge, no datasets have been created for this purpose.

Build and Operating Margins Approach: Data Limitations

The two subsections above largely obscure the primary critical limitation of this method: The only current sources and methods for obtaining w_b and build margin values are heuristic methods of unquantified accuracy. In addition, existing datasets were created primarily for forward-

[31] Note that the combination of both positive and negative demand perturbations at different points in time may counterintuitively result in marginal emissions rates greater than the emissions intensity of any existing generator. For example, if Segment A of Figure 9 had an intensity of 1,000 kg/MWh and the fossil generation of Segment E had an intensity of 500 kg/MWh, an attempt to describe the combination of those two effects would result in an operating margin of 1,500 kg/MWh. In an edge case where there is an equal MWh of both positive and negative perturbations for either build or operating changes, with different emissions intensities for the two portions, the build or operating margin would become infinite.

looking prospective analysis—to the authors' knowledge there is no purpose-built build margin or w_b datasets or guidance intended for retrospective or real-time applications.

Considering w_b values: Because operating margins are often predominately fossil fuels and build margins are often predominately clean generation, the relative weight of the two values drives the magnitude of an action's impact to a great extent.[32] The authors are aware of only two sources that give values or guidance for w_b:

1. Chapter 5 of the GHGP's Guidelines for Quantifying GHG Reductions from Grid-connected Electricity Projects (Broekhoff 2007) gives a heuristic method for estimating w_b. The method described focuses on whether the project is expected to provide firm capacity.[33]
2. The United Nations Framework Convention on Climate Change (UNFCCC), through technical working groups with international financial institutions, has published guidance to assign variable renewable generators (e.g., wind and solar) a w_b of 25%, "firm" renewable generators (e.g., hydropower, geothermal, biomass) a w_b of 66%, and both load reductions (e.g., energy efficiency measures) and increases (e.g., electrification) a w_b of 66% (IFI TWG 2023; 2019).

The treatment of w_b in both of these sources is relatively simple compared to a structured modeling method. A challenge that is not reflected in the above sources is that w_b would be expected to vary materially between actions with different temporal profiles and in different regions, and may vary meaningfully for actions with different magnitudes and durations. For example, the w_b of a wind generator in a region would be expected to vary from the w_b of a solar generator in the same region, the w_b of two wind generators in different regions would be expected to vary, the w_b of two different energy efficiency measures with different diurnal trends in energy savings would be expected to vary, and so forth.[34]

Similarly, existing data and methods for build margins are relatively simple. The GHGP provides multiple heuristic methods for estimating build margins, relying either on selecting a representative or "proxy" plant, or by examining recent capacity investment trends in the region being considered. The UNFCCC has released a dataset that contains country-level build margin estimates (as well as operating margins and, in combination with their w_b value recommendations mentioned above, composite marginal emissions rates for representative interventions). The UNFCCC build margins are drawn from modeling with the World Energy

[32] For instance, the impact of procuring energy from a wind generator in a region where wind is already being built in large amounts may have a w_b of approximately 1 (because it may be directly offsetting another wind generator that would otherwise have been built, as a result of, for example, competition or direct replacement in a rate-limited interconnection process), whereas a wind generator in a region where renewables are not being built may have a w_b of approximately 0. Even if the build and operating margins are identical in the two regions, the difference in w_b would be exceptionally important.

[33] Because much relevant structural change is the construction of variable renewable generators, which are often built in large part because of their energy value, we note that methods based on firm capacity provision have relatively lesser applicability in current contexts, compared to the context when the GHGP guidance was originally developed.

[34] For example, Gagnon and Cole (2022) contains data on the generation mixtures from load perturbations of various shapes, using a capacity expansion model. The w_b implicit in those data is seen to vary from nearly 0 to nearly 1, depending on how closely the load addition aligned with variable renewable energy production profiles (i.e., the more closely the action aligned with the shape of wind and solar generation, the greater the w_b).

Model, where the "new" electricity generation that the model deploys over an 8-year period is averaged and interpreted as the build margin, after further regression-informed adjustments to develop country-level values (IFI TWG 2022).[35]

The UNFCCC database is an example of how a forward-looking power sector model can be used to estimate build margins, which were then combined with heuristic w_b values and empirical/heuristic operating margins to arrive at a single composite marginal emissions rate for four different representative interventions. Note that, depending on the capabilities of the model at hand, it may also be possible to use it to calculate w_b values—although in doing so, it would be likely that an analyst could simply directly calculate a LRMER instead (thereby packaging operational and structural impacts into one self-consistent dataset).

In addition, as with the SRMER/LRMER method described in Section 3.2, naive applications of this operating/build margin approach can fail to capture the type of phenomena that can degrade (or entirely negate) the impact of actions meant to support the deployment of clean generation. It is still necessary to establish whether the project would be present in the counterfactual (i.e., whether the project is "additional"). It would also generally be useful to apply the screening tests discussed in Section 3.4 to help identify situations where phenomena are present that would tend to make the avoided emissions estimates from marginal emissions rates materially erroneous.

No method or dataset exists, or has been proposed, for calculating a time-varying w_b in real time. This is a particular difficulty for implementing real-time consequential analysis (i.e., the attempt to create a composite marginal emissions rate in real time) with build and operating margins because w_b would be expected to vary materially hour-to-hour based on the current conditions of the grid.[36] Similar to the prior discussion about real-time LRMER estimates, estimating w_b in real time appears to be unusually difficult and not readily amenable to scientific validation. Given this, it is possible that many users will ultimately find it more suitable to their purposes to use heuristics to guide real-time operational decisions, rather than hypothetical future attempts to produce real-time composite marginal emissions rates.[37]

The strength in this Approach #2 is its simplicity (the heuristic methods for estimating the values can be put into practice even in data-poor regions). The limitations, however, are significant. It can be said that this method gives a practicable approach to estimating impacts that is often at least directionally useful, but which can also materially err in its accuracy in ways that likely make it unsuitable for many purposes. Careful consideration of the manner in which this method could err is advised when deciding whether to adopt it for any specific purpose.

---

[35] While drawing from World Energy Model projections is more sophisticated than some other methods of estimating the build margin, it is still an approximation, as the method used implicitly assumes that further demand perturbations will induce or avoid the same build mixture projected to occur over the 8-year period.

[36] For example, the composite marginal emissions rate for charging a vehicle when wind is blowing strongly during one hour would be expected to have a higher w_b (from future induced wind investment) then a point several hours later in the day when wind is at a lull.

[37] Research is ongoing at NREL to study the performance of heuristics (e.g., renewable energy fractions, energy prices, and so forth) in their efficacy at guiding real-time operational decisions in light of minimizing greenhouse gas emissions. Publication forthcoming.

3.4 Evaluating the Impact of Actions Meant To Support the Deployment of Clean Generators

Estimating the avoided greenhouse gas emissions from actions intended to support the development of clean generators (such as the purchase of RECs or the signing of long-term offtaker agreements) is a topic of current importance, given the desire of many organizations to engage in activities that support defensible avoided emissions claims.[38]

Two key issues, however, are 1) the application of marginal emissions rates depends on accurately estimating how an action changes the demand for electricity from all generators beyond those explicitly considered within the intervention itself and 2) currently available marginal emissions rate data and methods are all imperfect, and to varying degrees omit important phenomena or depend on uncertain assumptions that can result in materially erroneous estimates of the impact of the actions.[39] In some contexts, naive application of marginal emissions rates can systematically overestimate the impact of actions meant to support the deployment of clean generators.

These challenges can influence the accuracy of any marginal emissions rates impact analysis, but actions meant to support clean generator deployment are particularly susceptible, given complicated interactions with the deployment of other clean generators that can significantly decrease impact—for example, inducing the construction of a wind generator may negatively influence the investment in other wind generators in the broader region.

Consequently, when seeking to estimate the impact of actions meant to support clean generator deployment with marginal emissions rates, analysts may wish to apply a three-step process, which includes two heuristic filters designed to help identify situations where the marginal emissions rates may be producing materially erroneous estimates.[40]

The first heuristic filter is meant to address the question, "would this project have gone forward without this specific action supporting it?" (i.e., additionality). The second heuristic filter is meant to address the question, "are the marginal emissions rates being used for this analysis suitable for the specific context under consideration, or are there relevant system-level phenomena not being well represented through the marginal emissions rates being used?".

[38] Note that there are many actions that would reduce induced emissions (such as load siting decisions or investing in energy efficiency upgrades) or increase avoided emissions (such as facilitating the construction of transmission lines). This section focuses on actions intended to support the deployment of clean generation for ease of exposition, but many of the concepts here can be extended to other activities as well.

[39] For example, capacity expansion models (used to create LRMERs) are generally built for long-term prospective analysis, not retrospective evaluations, and can err in their impact estimates through a failure to properly reflect relevant near-term phenomena (such as the administrative processing capacity of an interconnection process, as one of many possible examples). Or, as another example, consider methods for estimating short-run marginal emissions rates, which by nature omit any reflection of induced structural change. As a final example, many analysis methods *a priori* assume that the project under consideration would not have moved forward in absence of the action being analyzed—not always an appropriate assumption.

[40] The concept of a multistep screening process is not new, although it takes other forms in prior literature. For example, ACR (2023) defines a "three-pronged additionality test," and Gillenwater (2012) presents additionality as a binary screening after which a consequential analysis would need to still be applied. This section seeks to add to existing discussion by focusing, to a greater extent, on the shortcomings of marginal emissions rates in estimating the avoided emissions of actions meant to support the deployment of clean generators.

The following three-step process results:

1. **Heuristic Screen #1: Additionality.** Establish whether the action being taken has defensibly caused (or will defensibly cause) the associated project to move forward.
2. **Heuristic Screen #2: Marginal Emissions Rate Suitability.** Apply a set of screening heuristics designed to indicate situations where currently available marginal emissions rates may be missing important phenomena that would tend to result in their application producing materially erroneous estimates.
3. **Perform Analysis.** If the above two screens are passed, perform the marginal emissions rate analysis (i.e., an approach like the ones described in Sections 3.2 and 3.3).

Both heuristic filters are, in a sense, redundant, given the general requirement for all consequential analysis to appropriately define a counterfactual and use data that is suitable for the given purpose. They are raised here explicitly, however, due to the recent occasions of marginal emissions rates being used to estimate the avoided emissions of actions meant to support the development of clean generation, where the conclusions of the analyses are not being meaningfully restrained by known limitations of the data being employed. For example, Arne Olson et al. (2023) and Hua He et al. (2023) are both analyses that *a priori* assumed the projects being studied are additional, and solely reflect short-run operational impacts while neglecting the potential for induced structural change. Such approaches have the potential to systematically overestimate the impact of actions meant to support the deployment of clean generators (Gagnon and Cole 2022; Xu et al. 2024) and do not follow best practices in consequential analysis (Broekhoff 2007; Woolf et al. 2020).

Additionality

The first heuristic filter—commonly referred to as an additionality test—is much discussed in communities such as those focused on carbon offsets. This document has largely avoided the term "additionality" because it does not perfectly align with consequential impact analysis (e.g., a nonadditional project can still have an impact, and an additional project can have no impact).[41] However, as a practical matter, additionality tests can be employed as heuristics as part of a multistep method, when employing marginal emissions rates, for estimating the avoided emissions from actions meant to support the deployment of clean generators.

If an additionality test is failed, an analyst may wish to remove the project from consideration, as an indication that the action has ambiguous and potentially low impact—or at any rate, even if the project does have an impact, that impact cannot be estimated using marginal emissions rates combined with the associated project's nameplate output (because the lack of additionality means the associated project's nameplate output does not represent a defensible difference from

[41] For example, consider an organization signing a long-term offtaker agreement for a project that would have otherwise still gone ahead. Although not additional in the strict binary sense used in this paper, the action may have secondary impacts on the capital available for other clean generation projects and may ultimately therefore still have a systemwide emissions impact. While such a causal chain is theoretically possible—and indeed the authors consider it likely prevalent to at least some degree—there is no theoretical expectation that the ultimate impact would be well-represented by the given project's nameplate output and local marginal emissions rates. The authors are not aware of research to develop methods of estimating the impact of actions that support nonadditional projects, but such research may be useful.

the counterfactual, given that a failed additionality test indicates the potential that the project would have become operational even in absence of the action being analyzed).

This document has little to add to the existing literature on this topic and so does not dwell on it at length, beyond noting 1) the increasing frequency of renewable generators with positive net values (Wiser et al. 2024) enhances the difficulty of robustly demonstrating additionality for such projects and 2) there is not general consensus on workable additionality tests.[42] The current lack of generally accepted additionality tests challenges the use of marginal emissions rates as a basis for claims of avoided emissions from actions meant to support clean generation projects.[43]

Marginal Emissions Rate Suitability

The section above discussed heuristic tests that focused on the project associated with the action. This section discusses heuristic tests that focus on how the project interacts with the system at large.[44] The motivation for this second heuristic screen is that marginal emissions rate data and models can, due to either known or unknown missing phenomena, omit important system-level impacts—so combining existing models and data with screens meant to help avoid the most likely misfiring of the use of their outputs would often be helpful. Said differently: The second screen is meant to help identify situations where a particular marginal emissions rate data product may be likely to produce erroneous estimates (and conversely, where the application of marginal emissions rates may be likely to produce accurate estimates).

Documenting all possible tests and describing their interaction and relevance with the breadth of possible interventions and contexts is a large task that goes beyond the scope of this document. The most useful tests would vary by data product and context. This section presents several possible heuristics, to build intuition, recognizing that widespread adoption of avoided emissions claims would benefit from a more thorough collection and articulation of these tests.

Following is a list of possible heuristics:

1. Is the project that the action is supporting substantially similar to other projects that have been deployed in the region (i.e., a wind generator where wind is actively being deployed)? This is sometimes referred to as a common practice test.

---

[42] To be clear, even with generally low renewable costs and subsidies, many potential clean generation projects would not be expected to be competitive without additional revenues. Establishing additionality may be relatively straightforward and unambiguous in such cases. In addition, even where expected revenues exceed expected costs, considerations of risk and capital availability may still allow for defensible claims of additionality. The note in the body of the text is drawing attention to the fact that increasing prevalence of competitive projects increases the frequency of situations where additionality is less clear, and that tests that incorporate phenomena such as risk and capital availability have not yet achieved consensus.

[43] For example, financial tests are popular, but financial models are complicated with many highly uncertain input assumptions—meaning that an analyst can generally adjust values within plausible ranges and achieve a desired outcome.

[44] Note that, in some literature, tests such as presented here are also categorized as additionality tests—for example, ACR (2023). This paper seeks to differentiate between project-level tests and system-level tests, while recognizing that there are likely multiple useful ways to classify the different tests.

2. Is there a bottleneck associated with an administrative or engineering process (such as approvals in an interconnection queue), and are there other clean generators behind the bottleneck?[45]
3. Does the project support reaching a stated clean energy or decarbonization target, goal, legislative requirement, or similar, where the relevant entity has the capacity to see that the target is fulfilled without the action under consideration? This is sometimes referred to as a regulatory test.
4. If a market exists for the energy attribute certificates produced by the project (e.g., renewable energy credits), is the price of the certificates low?
5. Does power sector modeling of the region where the project is located indicate investment in the same type of technology, without explicit representation of the specific action being analyzed?

To varying degrees, the phenomenon implied by these tests (such as competition with other clean generators, or interactions with policy requirements) can be represented in the workflow being used to create a marginal emissions rate. If a heuristic indicates that a phenomenon is potentially present, and it is not defensibly represented within the marginal emissions rate being used, the use of that marginal emissions rate could lead to erroneous impact estimates.

The use and interpretation of such heuristics would be situation-specific (partially because different issues can be of varying importance depending on the context, and partially because of the requirement for normative decisions by whoever is defining the analytical workflow), and therefore this paper does not give specific guidance on how such a filter should be employed in any particular situation. To build intuition on possible lines of reasoning, however, consider two prototypical contexts: guiding internal decision making and reporting impacts externally.

These two example contexts differ in two important ways: First, internal decision making often allows for nuanced decision-making processes that can, for example, fold in qualitative assessments of data quality by analysts or estimates expressed as ranges instead of point values. External reporting, in contrast, is often stripped of nuance. Second, analysts performing consequential analysis that is purely for internal consumption may have the desire for accurate estimates aligned with the incentives of the organization. In contrast, when reporting results externally, an organization may not have strong incentives for accurate impact assessments.

Considering the first example context: When consequential analysis is used internally to guide decisions, analysts may apply heuristics like the above, and then incorporate the results of the evaluation into their decision making in a variety of flexible ways. For example, if analysts are evaluating 10 projects, they may wish to reject the bottom 5 that score worst on the screen, and proceed with analyzing the remaining 5. Or they may wish to set a numerical criterion, such as only considering projects that have a "yes" on, say, 4 out of 5 of the criteria (and if no projects

[45] To avoid confusion, this is not intended to be an implementation barriers test, which would be part of the additionality screening according to the classification used in this paper. Instead, this test is seeking to identify situations where there is a limit on the rate of projects that can be deployed, where the limit is not expected to be significantly impacted by the action being analyzed. Unless such situations are explicitly represented in the methods used to create the marginal emissions rates being used, such rate limits may produce material errors. For example, a congested interconnection queue may result in the deployment of one generator directly offsetting another that would otherwise have proceeded, likely largely negating the impact of the first generator—a situation not commonly captured by marginal emissions rates.

meet that threshold, perhaps directing their resources toward efforts with less ambiguous impacts).

Consider, in contrast, the second example context: If used externally (such as a component within a hypothetical consequential corporate reporting framework), the ambiguity of the heuristics and potential for misaligned incentives could justify a more defined approach, depending on the objectives of the standard-setting organization. For example, instead of a list of heuristics, the reporting system may define a set of heuristic-based pathways that achieve broad consensus as being indicative of impactful projects, where an organization can claim an offset only if they demonstrate compliance with at least one of the pathways.[46] The stringency of the pathways would reflect the normative position of the standard-setting organization (i.e., the organization's stance toward the desired prevalence of "false positives" versus "false negatives"—taking the form of overestimates of impact and exclusions of impactful projects, respectively).

These heuristics are being suggested here only for practical reasons, not theoretically satisfying ones. In the long term, it would likely be preferrable to build a model that has a sufficiently adequate representation of all relevant phenomena, to be able to natively output reliable estimates of impacts. In the meantime, however, the author's view is that currently-existing marginal emissions rates data products would all benefit from supplemental screening like the above to minimize the likelihood of claiming an action as impactful when it is not.

3.5 Summary of General Limitations of Using Marginal Emissions Rates for Consequential Analysis

Throughout this document, there have been various discussions of some specific limitations of using marginal emissions rates when used for consequential analysis. This section steps back to discuss some of the general limitations in broader terms, building from larger discussions from the literature (Plevins et al. 2013; Anex and Lifset 2014; Brandão et al. 2014; Dale and Kim 2014; Jones et al. 2017; Ekvall 2019). This, in combination with the more focused discussions in Sections 3.1 through 3.4, is provided to help potential users reflect on the suitability of this analytical approach for a specific application.

- **The accuracy of consequential emissions analysis performed with marginal emissions rates has not been empirically quantified.** The accuracy of analytical methods can be evaluated through empirical validation, where analytical outputs are compared to observed outputs. Empirical validation is unusually challenging in the case of consequential emissions analysis which, by definition, depends on unobservable counterfactual scenarios. This challenge is compounded when analyzing electric-sector interventions, where scientific techniques for establishing causality, such as randomized controlled trials, are generally not practicable, particularly when structural change is present. As a result, although work has been done to increase confidence in the

[46] An example pathway: If an action to support a clean generator being deployed can show 1) it passes an externally verifiable test to demonstrate that the project would not move forward in absence of the action being evaluated and 2) the generator is of a type where no other generators of that class (e.g., wind, solar, geothermal) have been previously deployed in the region (i.e., a common practice test). If those two conditions are met, an organization could claim the project as an offset. While not a guarantee of impact, such a pathway may achieve consensus as generally indicative of impact.

reasonableness of various components of consequential emissions analysis (such as SRMERs), the uncertainty associated with comprehensively estimating the consequences of specific actions remains unquantified.

- **Consequential emissions analysis is sensitive to highly uncertain key assumptions.** All analytical methods are sensitive to underlying assumptions, but consequential analysis of power sector impacts is exceptionally exposed to impactful yet uncertain assumptions. The uncertainty can have various causes, such as a deficit of data (e.g. the financial state of developers in the interconnection queue), difficult-to-model phenomena (e.g. the decision-making processes of resource planners), or the inherent uncertainty about the future (e.g. future changes in state or federal policies). These uncertainties must be appreciated when a certain threshold of accuracy is desired (or necessary). Best practices for reflecting the uncertainty (e.g., producing a range of plausible values through sensitivity analysis) are not practicable for all decision-making workflows, which may have the ability to use only a single value without an expression of uncertainty.
- **Important phenomena may be missing from models and methods.** Consequential analysis is highly sensitive to the types of phenomena that are represented in the models employed, or implicitly reflected in the design of an empirical methodology. A model that omits an important phenomenon may produce a materially inaccurate impact estimate. Transmission congestion, elasticity of demand for various services, capital supply curves, interconnection queues, supply chain constraints and dynamics, and induced structural change are just some of many possible examples of frequently omitted phenomena that may materially affect the accuracy of the marginal emissions rates produced by a workflow, and therefore the accuracy of the consequential analysis using those rates. Some errors may be roughly symmetrical in their over- or under-estimation, while others may be materially and systematically biased in one direction: For example, section 3.4 discussed ways in which marginal emissions rate could systematically over-estimate the impact of actions meant to support the deployment of clean generation.
- **The effects of intervention size and duration on impact have not been rigorously characterized.** Interventions vary in magnitude and duration. The impact of the variations has not been rigorously characterized. A conceptual framework for the relationship between an intervention's size and duration, and structural decisions, is lacking. Marginal emissions rates are generally applied equally to interventions of varying size and duration, but the degree to which this errs is not known.
- **Consequential emissions analysis can be sensitive to the spatiotemporal boundaries of the underlying models.** Consequential analysis seeks to comprehensively estimate impacts, but in practice, resource and data constraints force analysts to define geographic and temporal boundaries. Drawing these boundaries is often of significant importance to consequential analysis, as a poorly drawn boundary can materially impact the results of an analysis.
- **The models used to generate marginal emissions rates are often complex, complicating external critique.** The models used to generate marginal emissions rates can be complex, and this complexity can make it difficult for the model and assumptions to be externally examined, particularly if the model, inputs, or intermediate analytical steps are not made publicly available.
- **Marginal emissions rate data availability and quality varies by location.** A method that works for a certain region of a certain country may not work elsewhere in the

country, let alone internationally. Comparing point estimates of impact between projects from different workflows should be done cautiously.

While the above considerations, as well as those expressed earlier in the document, are important (and in many instances may have sufficient impact that a decision maker would be better served by a different analytical approach), this discussion should not be interpreted as suggesting that consequential emissions analysis is fatally flawed for all applications or functionally impossible to practice. A well-conducted consequential emissions analysis whose results are interpreted in light of their underlying uncertainty can often be a useful contribution to a decision-making process.

In addition, active work is ongoing from diverse groups, to improve various of the above considerations. The authors' views on potential future research activities are discussed in Section 4. In some cases, certain issues may be partially or entirely mitigated. In others, downstream decision-making workflows and guidance may be developed to mitigate the issue's negative impact on decisions. This is an active field of research, with many improvements likely in the coming years.

4 Looking Forward

This paper has discussed the state of the art of consequential emissions analysis and has referenced the concept of the approach being "suitable" for particular applications throughout. Such designations are unusually challenging, however, for three reasons: First, many of the shortcomings and uncertainties are unquantified, meaning that any conclusions require some degree of subjectivity. Second, the suitability of a particular method for a particular purpose depends on normative positions, which are unlikely to be shared by all practitioners. And third, the variety of different interventions, contexts, and objectives calls for nuanced, case-by-case treatment.

Given these difficulties, this paper does not make broad concluding statements about the suitability of current methods. Some uses of consequential emissions analysis seem, to the authors, to likely be generally accepted as suitable (e.g., long-lived prospective analysis of policies or loads, given its current widespread use for that purpose already) or generally viewed as unsuitable (e.g., real-time estimates, given the lack of any real-time estimates of induced structural change)—however, many potential applications of consequential analysis merit nuanced discussions before a decision about suitability can be made.

Looking forward, however, numerous possible research activities could help increase the accuracy of consequential analysis, thereby generally increasing its usefulness as well as extending its suitability into new domains. Some possible research activities are discussed here.

Future Research: Near Term

As discussed in Section 3.4, two critical shortcomings of the current state of the art are 1) the lack of generally accepted additionality tests and 2) deficits in the methods for calculating marginal emissions rates that create the potential for large, systematic errors in their estimates of impact, even when additionality is demonstrated. Therefore, two useful near-term research activities could focus on assembling researchers and practitioners to develop tests and heuristics to mitigate the worst of those shortcomings:

- Within an inclusive working group with broad participation, develop externally verifiable additionality tests for actions meant to support the deployment of clean generators (i.e., tests meant to establish whether a project is in the counterfactual).
- Within an inclusive working group with broad participation, develop a list of heuristics that indicate situations where currently available marginal emissions rates would tend to result in erroneous estimates because of the omission of countervailing system-level phenomena (i.e., an expanded and elaborated list of the manner described in Section 3.4), or alternatively, identify activities whose impacts can defensibly be estimated with currently available data products. Develop guidance on the interpretation of the heuristics, and develop tractable pathways that tend to indicate impactful projects.

Future Research: Long Term

In the longer term, fundamental theory and practice could be advanced, to improve the quality of consequential impact estimates and reduce the reliance on heuristics and subjectivity. Several possible longer-term activities:

- Develop either a global available, regularly updated database of long-run marginal emissions rates, or improve the sophistication of the w_b estimates that are currently available. Pair such databases with guidance on identifying interventions where the application of the data would tend to result in erroneous estimates (e.g., an elaboration of the additionality and system-level considerations discussed in Section 3.4).
- Expand the methods and guidance for defensible avoided emission claims associated with forward-looking prospective analysis for specific projects. For example, develop scalable methods for estimating the impact of interventions that are not demand perturbations (e.g., building transmission lines), potentially through standardization of power sector modeling for specific types of questions, or by developing reduced-form models.
- Develop a method for calculating a composite marginal emissions rate that folds all cascading impacts over time (both inter- and intrayear effects), back to the original point of action.
- Research the impacts of electric sector transmission congestion in the long-run.
- Research methods of estimating the impact of actions meant to support clean generator deployment for nonadditional projects.
- Develop a purpose-built model, or combination of models, for calculating retrospective impact estimates. Explore the possibility of combining short-run marginal emissions rates developed from operator data with structural impacts developed through purpose-built near-term capacity expansion models (i.e., develop a combined approach that leverages the relative strengths of empirical and model-based methods).

5 References

A. D. Hawkes. 2014. "Long-Run Marginal CO2 Emissions Factors in National Electricity Systems." *Applied Energy* 125:197–205.

ACR. 2023. "The ACR Standard: Requirements and Specifications for the Quantification, Monitoring, Reporting, Verification, and Registration of Project-Based GHG Emissions Reductions and Removals." ACR. https://acrcarbon.org/wp-content/uploads/2023/10/ACR-Standard-v8.0.pdf.

Anex, Robert, and Reid Lifset. 2014. "Life Cycle Assessment." *Journal of Industrial Ecology* 18 (3): 321–23. https://doi.org/10.1111/jiec.12157.

Arne Olson, Gregory Gangelhoff, Anthony Fratto, Hadiza Felicien, and Karl Walter. 2023. "Analysis of Hourly And Annual Greenhouse Gas Emissions." E3.

Aukland, Louise, Pedro Moura Costa, and Sandra Brown. 2003. "A Conceptual Framework and Its Application for Addressing Leakage: The Case of Avoided Deforestation." *Climate Policy*.

Azevedo, Inês, Priya Donti, Nathaniel Horner, Greg Schivley, Kyle Siler-Evans, and Parth Vaishnav. 2020. "Electricity Marginal Factor Estimates." Center For Climate and Energy Decision Making. https://cedm.shinyapps.io/MarginalFactors/.

Bistline, John, Geoffrey Blanford, Maxwell Brown, Dallas Burtraw, Maya Domeshek, Jamil Farbes, Allen Fawcett, et al. 2023. "Emissions and Energy Impacts of the Inflation Reduction Act." *Science* 380 (6652): 1324–27. https://doi.org/10.1126/science.adg3781.

Bjørn, Anders, Joachim Peter Tilsted, Amr Addas, and Shannon M. Lloyd. 2022. "Can Science-Based Targets Make the Private Sector Paris-Aligned? A Review of the Emerging Evidence." *Current Climate Change Reports* 8 (2): 53–69. https://doi.org/10.1007/s40641-022-00182-w.

Brander, Matthew. 2016. "Transposing Lessons between Different Forms of Consequential Greenhouse Gas Accounting: Lessons for Consequential Life Cycle Assessment, Project-Level Accounting, and Policy-Level Accounting." *Journal of Cleaner Production* 112 (January):4247–56. https://doi.org/10.1016/j.jclepro.2015.05.101.

———. 2021. "The Most Important GHG Accounting Concept You May Not Have Heard of: The Attributional-Consequential Distinction." *GHG and Carbon Accounting, Auditing, Management & Training | Greenhouse Gas Management Institute* (blog). April 21, 2021. https://ghginstitute.org/2021/04/21/the-most-important-ghg-accounting-concept-you-may-not-have-heard-of-the-attributional-consequential-distinction/.

Brander, Matthew, and Francisco Ascui. 2015. "The Attributional-Consequential Distinction and Its Applicability to Corporate Carbon Accounting." In *Corporate Carbon and Climate Accounting*, edited by Stefan Schaltegger, Dimitar Zvezdov, Igor Alvarez Etxeberria, Maria Csutora, and Edeltraud Günther, 99–120. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-27718-9_5.

Brander, Matthew, Roger L. Burritt, and Katherine L. Christ. 2019. "Coupling Attributional and Consequential Life Cycle Assessment: A Matter of Social Responsibility." *Journal of Cleaner Production* 215 (April):514–21. https://doi.org/10.1016/j.jclepro.2019.01.066.

Broekhoff, Derik. 2007. "Guidelines for Quantifying GHG Reductions from Grid-Connected Electricity Projects." World resources institute.

Earles, J. Mason, and Anthony Halog. 2011. "Consequential Life Cycle Assessment: A Review." *The International Journal of Life Cycle Assessment* 16 (5): 445–53. https://doi.org/10.1007/s11367-011-0275-9.

Energy+Environmental Economics. 2016. “Avoided Cost Calculator User Manual.” Energy+Environmental Economics.
European Commission. 2010. *International Reference Life Cycle Data System (ILCD) Handbook*. LU: Publications Office. https://data.europa.eu/doi/10.2788/38479.
Gagnon, Pieter. 2024. “Long-Run Marginal Emission Rates for Electricity - Workbooks for 2023 Cambium Data.” NREL Data Catalog. https://doi.org/10.7799/2305481.
Gagnon, Pieter, and Wesley Cole. 2022. “Planning for the Evolution of the Electric Grid with a Long-Run Marginal Emission Rate.” *IScience* 25 (3). https://doi.org/10.1016/j.isci.2022.103915.
Gagnon, Pieter, Elaine Hale, and Wesley Cole. 2022. “Long-Run Marginal Emission Rates for Electricity - Workbooks for 2021 Cambium Data.” NREL Data Catalog. https://doi.org/10.7799/1838370.
Gagnon, Pieter, Pedro Andres Sanchez Perez, Kodi Obika, Marty Schwarz, James Morris, Jianli Gu, and Jordan Eisenman. 2024. “Cambium 2023 Scenario Descriptions and Documentation.” 88507. National Renewable Energy Laboratory.
Gillenwater, Michael. 2012. “What Is Additionality.” What Is Additionality. Greenhouse Gas Management Institute.
Gregory Miller. 2022. “Applying the Consequential Emissions Framework for Emissions-Optimized Decision-Making for Energy Procurement and Management.” Clean Energy Buyers Institute.
Hansson, Sven Ove. 2007. “PHILOSOPHICAL PROBLEMS IN COST–BENEFIT ANALYSIS.” *Economics and Philosophy* 23 (2): 163–83. https://doi.org/10.1017/S0266267107001356.
Hertel, Thomas W., Alla A. Golub, Andrew D. Jones, Michael O’Hare, Richard J. Plevin, and Daniel M. Kammen. 2010. “Effects of US Maize Ethanol on Global Land Use and Greenhouse Gas Emissions: Estimating Market-Mediated Responses.” *BioScience* 60 (3): 223–31. https://doi.org/10.1525/bio.2010.60.3.8.
Ho, Jonathan, Jonathon Becker, Maxwell Brown, Patrick Brown, Ilya Chernyakhovskiy, Stuart Cohen, Wesley Cole, et al. 2021. “Regional Energy Deployment System (ReEDS) Model Documentation: Version 2020.” Golden, CO: National Renewable Energy Laboratory. https://www.nrel.gov/docs/fy21osti/78195.pdf.
Hsu, David D., Daniel Inman, Garvin A. Heath, Edward J. Wolfrum, Margaret K. Mann, and Andy Aden. 2010. “Life Cycle Environmental Impacts of Selected U.S. Ethanol Production and Use Pathways in 2022.” *Environmental Science & Technology* 44 (13): 5289–97. https://doi.org/10.1021/es100186h.
Hua He, Alex Derenchuk, Aleksandr Rudkevich, and Richard Tabors. 2023. “Paths to Carbon Neutrality.” Tabors Caramanis Rudkevich.
IFI. 2022. “The IFI Dataset of Default Grid Factors.” International Financial Institution.
IFI TWG. 2019. “GHG Accounting for Grid Connected Renewable Energy Projects.” International Financial Institutions. https://unfccc.int/sites/default/files/resource/Renewable%20Energy_GHG%20accounting%20approach.pdf.
———. 2022. “Methodological Approach for the Common Default Grid Emission Factor Dataset.” International Financial Institutions. https://unfccc.int/sites/default/files/resource/IFITWG_Methodological_approach_to_common_dataset.pdf.

———. 2023. “IFI Approach to GHG Accounting for Energy Efficiency Projects.” International Financial Institutions. https://unfccc.int/sites/default/files/resource/Energy%20Efficiency_GHG%20accounting%20approach.pdf.

Michael Gillenwater. 2023. “What Is Greenhouse Gas Accounting? Turning Away from LCA.” December 2023. https://ghginstitute.org/2023/12/19/what-is-greenhouse-gas-accounting-turning-away-from-lca/.

Oates, DL, and Kathleen Spees. 2022. “Locational Marginal Emissions.” REsurety.

Plevin, Richard J., Mark A. Delucchi, and Felix Creutzig. 2014. “Using Attributional Life Cycle Assessment to Estimate Climate-Change Mitigation Benefits Misleads Policy Makers.” *Journal of Industrial Ecology* 18 (1): 73–83. https://doi.org/10.1111/jiec.12074.

Schaltegger, Stefan, Dimitar Zvezdov, Igor Alvarez Etxeberria, Maria Csutora, and Edeltraud Günther, eds. 2015. *Corporate Carbon and Climate Accounting*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-27718-9.

Steinberg, Daniel, Maxwell Brown, Ryan Wiser, Paul Donohoo-Vallett, Pieter Gagnon, Anne Hamilton, Matthew Mowers, Caitlin Murphy, and Ashreeta Prasanna. 2023. “Evaluating Impacts of the Inflation Reduction Act and Bipartisan Infrastructure Law on the U.S. Power System.” NREL/TP-6A20-85242. Golden, CO: National Renewable Energy Laboratory. https://doi.org/10.2172/1962552.

Tomas Ekvall. 2019. “Attributional and Consequential Life Cycle Assessment.” In *Sustainability Assessment at the 21st Century*. Intech Open. https://www.intechopen.com/chapters/69212.

U.S. Department of Energy. 2023. “Assessing Lifecycle Greenhouse Gas Emissiosn Associated with Electricity Use for the Section 45V Clean Hydrogen Production Tax Credit.” U. S. Department of Energy.

Weidema, Bo P., Niels Frees, and Anne-Merete Nielsen. 1999. “Marginal Production Technologies for Life Cycle Inventories.” *The International Journal of Life Cycle Assessment* 4 (1): 48–56. https://doi.org/10.1007/BF02979395.

Wiser, Ryan, Mark Bolinger, Dev Millstein, and Joachim Seel. 2024. “Grid Value and Cost of Utility- Scale Wind and Solar: Potential Implications for Consumer Electricity Bills,” June.

Woolf, Tim, Chris Neme, Mike Alter, Steve Fine, Karl Rabago, Steven Schiller, Kate Strickland, and Brenda Chew. 2020. “National Standard Practice Manual.” National Energy Screening Project. https://www.nationalenergyscreeningproject.org/wp-content/uploads/2020/08/NSPM-DERs_08-24-2020.pdf.

Xu, Qingyu, Wilson Ricks, Aneesha Manocha, Neha Patankar, and Jesse D. Jenkins. 2024. “System-Level Impacts of Voluntary Carbon-Free Electricity Procurement Strategies.” *Joule*, January, S2542435123004993. https://doi.org/10.1016/j.joule.2023.12.007.

Appendix

A.1 Stylized Model for Emissions Impacts

Sections 2 and 3 contain figures showing the impacts of several stylized interventions. Those figures were generated with an intentionally simple model that was designed to help illustrate operational and structural impacts. The model assumes the following:

- In a given year, annual generation from variable renewable generators (i.e., wind and solar generators) is half of the average electric demand from 3 years prior (e.g., if demand averaged 100 MWh in 2024, then throughout 2027 the quantity of variable renewable generation is assumed as 50 MWh). Variable renewable generation is assumed to not decrease.
- Any demand not met by the variable renewable generators is met by natural gas generators. It is assumed that there is sufficient natural gas generation capacity to meet all remaining demand, and that any changes in natural gas capacity do not impact the generation mixture.
- Load growth is linear and, on an annual basis, is greater than half the size of the intervention being analyzed.
- The variable renewable generators have a CO_2 emissions rate of 0 kg/MWh and the natural gas generators have a CO_2 emissions rate of 400 kg/MWh.

The time delay for the level of renewable energy is meant to represent the time lag of investment decisions because of myriad factors, such as the time it takes for actions to appear in the data used for planning, the time to construction capital assets, administrative procedures prior to construction, and so forth. The quantity of load growth was selected to be large enough to allow the system to return to a state identical to the counterfactual within 1 year, for ease of explanation.

To help build intuition about how this model produced the impact patterns shown in the body of the paper, Figure 10 and Figure 11 show two example outputs from the model. Both examples reflect the assumptions made for the retrospective assessment class from Section 2.3, because feedback during the writing of this paper indicated that it warranted additional explanation. The assumptions in the figures below (e.g., the starting point and growth of demand) were selected to facilitate visual examination of the dynamics shown in the figures—the values are not intended to be realistic.

Figure 10 shows the same intervention pattern described in Section 2.3, where the intervention is a load addition that starts in 2024, and the counterfactual is the same load addition instead starting in 2025.

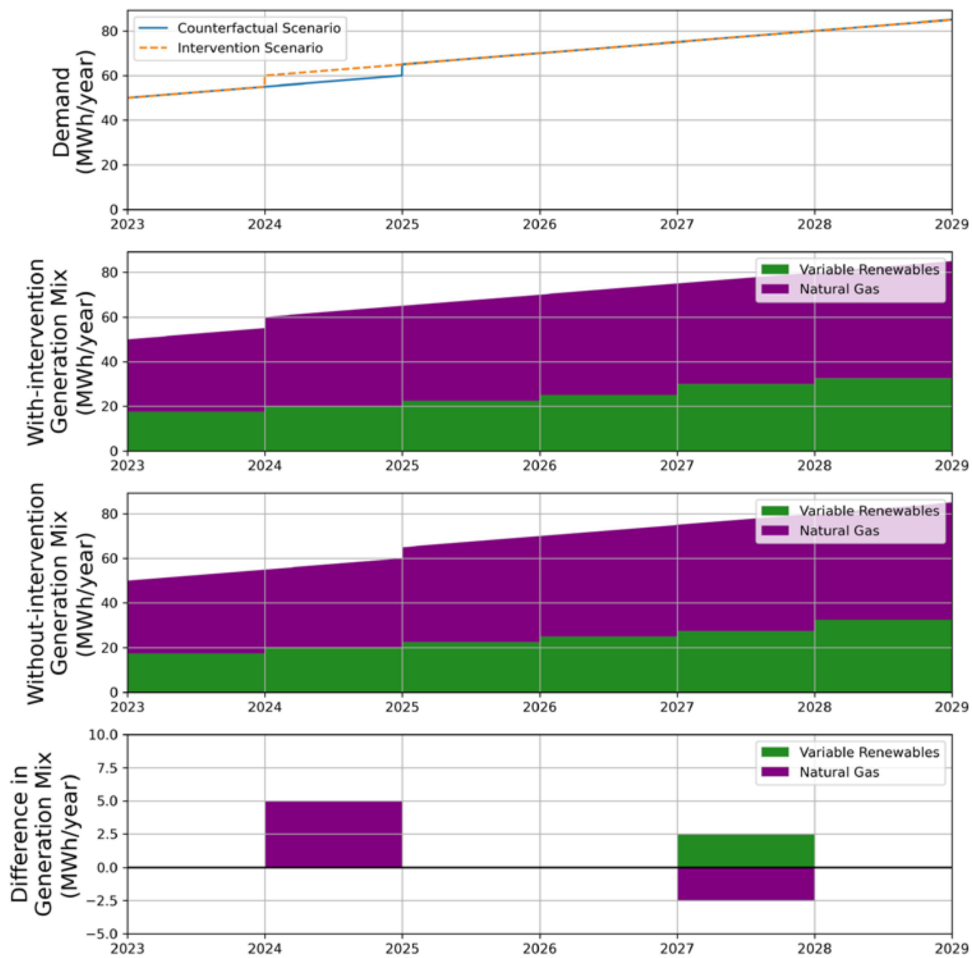

Figure 10. Stylized model outputs: Example #1

Figure 11 shows the results of a different situation, where the intervention is instead just a single-year load increase and where the counterfactual is the load increase not occurring at any point.

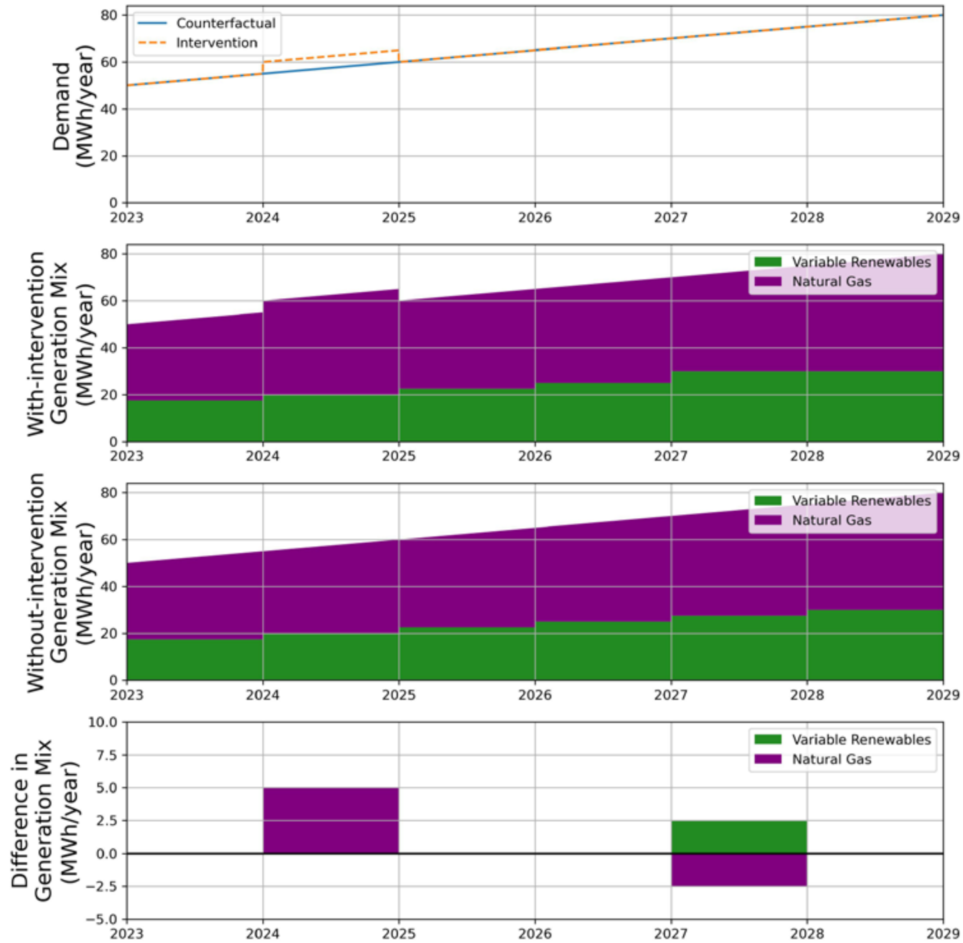

Figure 11. Stylized model outputs: Example #2

The interventions in Figure 10 and Figure 11 were selected, in part, to illustrate how, under certain stylized situations, multiple types of interventions can produce the type of pattern that was previously seen in Section 2.3. We see how symmetrically time-lagged structural decisions can produce both operational and structural responses for short-lived decisions, for example.

In addition, however, they were selected as a basis to discuss a key uncertainty, which is the exact nature of how the entities making investment decisions would respond to the interventions. In the stylized example here, the response is simple and mechanical, where renewable generator investment is mechanically defined to achieve a fraction of total load, with a 3-year time lag. In practice, the decision could be different, and a great deal more complex—and in many plausible circumstances, not result in the outcomes shown above.

Decisions would likely be made based on more years of data, potentially reducing in magnitude but increasing in duration the time-lagged structural response. Many interventions would be sufficiently small that they would not be directly "observed" by resource planners (in the sense that any specific person would discern a specific event), and therefore the manner in which they would influence planning would potentially be a mechanical result of an analysis process.

Any effort to develop sophisticated representations of how resource planners would react to specific interventions faces many obstacles, including the diversity of approaches across administrative regions, the fact that multiple actors may be involved in structural decisions and may be using different planning methods (e.g., a regulated utility and independent power producers in the same region), the fact that methodological approaches of specific institutions vary over time, and so forth. Ultimately, it appears likely that methods will face material counteracting challenges in tractability and accuracy without the benefit of scientific validation—and the suitability of any stylized methodology for a particular purpose may not be clear but ultimately depend on nuanced normative considerations.