

Economy-Wide Modeling: Uncertainty, Verification, and Validation

June 17, 2016

Prepared for the U.S. EPA Science Advisory Board Panel (SAB) on Economy-Wide
Modeling of the Benefits and Costs of Environmental Regulation

This paper has been developed to inform the deliberations of the SAB Panel on the technical merits and challenges of economy-wide modeling for an air regulation. It is not an official EPA report nor does it necessarily represent the official policies or views of the U.S. EPA.

[This page intentionally left blank.]

Table of Contents

1	Introduction	4
2	Uncertainty in Cost Estimation	5
2.1	Uncertainty in Regulatory Cost Estimates	5
2.2	Uncertainty in implementation of Regulatory Costs in Economic Models.....	8
3	Uncertainty in Benefits Estimation	9
3.1	Uncertainty in Regulatory Benefits Estimates	9
3.2	Uncertainty in Introducing Air Quality Improvement Benefits in Economic Models	11
4	Uncertainty in CGE Modeling.....	11
4.1	Parametric Uncertainty	11
4.2	Structural Uncertainty	13
5	Analytical Approaches.....	13
5.1	Sensitivity Analysis	14
5.2	Formal Uncertainty Analysis.....	15
5.3	Inter-Model Comparison Exercises	20
6	Presentation of Uncertainty Analysis.....	20
7	Verification and Validation Exercises	24
7.1	Model Verification.....	25
7.2	Model validation	26
7.2.1	Back of the Envelope Models	26
7.2.2	Statistical Validation of Model Parameters	26
7.2.3	CGE Model Validation Exercises	27
7.2.4	Challenges of Model Validation.....	28
8	Concluding Remarks.....	28
	References.....	30

1 Introduction

Benefit cost and economic impact analyses of national regulations designed to improve air quality require information from numerous scientific disciplines and sources to be linked and assessed. Each component of these analyses (e.g., engineering assessment, air quality modeling, risk assessment, economic analysis) may be subject to uncertainty and the linkages between these components may act to compound the uncertainty. A recent National Academy of Sciences study on making Environmental Decisions Under Uncertainty noted that “informed identification and use of uncertainties in the process is an essential feature of environmental decision making” (NAS, 2013). In 2002, the NAS also noted that:

“Even great uncertainty does not imply that action to promote or protect public health should be delayed. Decisions about whether to act, when to act, and how aggressively to act can only be made with some understanding of the likelihood and consequences of alternative courses of action. The potential for improving decisions through research must be balanced against the public health costs incurred because of a delay in the implementation of controls. Complete certainty is an unattainable ideal.”

The NAS Committee in 2013 classified uncertainties according to three types: (1) statistical variability and heterogeneity, (2) deep uncertainty, and (3) model and parameter uncertainty. Statistical variability and heterogeneity refers to exogenous sources of uncertainty inherent in the system or process under study. These are sources of uncertainty that, *ex ante*, cannot be reduced through additional research or data collection, for example stochasticity in dynamic processes whose fundamentals are well understood. Deep uncertainty refers to cases where there is a lack of understanding, or notable disagreement, about the fundamentals of underlying processes important to understanding the impact of an intervention in the system of interest. These sources of uncertainty are characterized by cases where additional research or data collection are unlikely to reduce the uncertainty in the time before a policy maker needs to select a course of action. Model and parameter uncertainty refers to uncertainty in models that are used to represent underlying processes relevant to understanding the impact of an intervention in the system of interest, where a model is defined as a “simplification of reality that is constructed to gain insights into select attributes of a particular physical, biologic, economic, or social system” (NRC, 2009). These sources of uncertainty arise due to current limitations of scientific and economic processes, now and in the future, and in theory could be reduced by additional research prior to making a decision.

While all these sources of uncertainty may be relevant for decision making regarding environmental regulations designed to improve air quality, this paper focuses primarily on model and parameter uncertainty related to benefit cost analysis and the potential use of CGE models for economy-wide analysis.¹ This white paper introduces some of the key sources of uncertainty associated with benefit cost

¹ Just as uncertainty may affect decisions about environmental regulations, uncertainty may affect other decisions in the economy. While outside the scope of this paper, dynamic stochastic general equilibrium (DSGE) models can capture some aspects of decision making under uncertainty by economic agents. For examples of DSGE models applied to environmental issues see the white paper on Economy-Wide Modeling: Evaluating the Economic Impacts of Air Regulations.

analyses of air regulations and additional sources of uncertainty that may arise when extending those analyses to include an economy-wide assessment using CGE models. Given these sources of uncertainty, this white paper discusses a number of analytical approaches that may be used to provide decision makers and stakeholders with additional information about the robustness of model results. These approaches and other methods of model validation and verification can help provide confidence in the qualitative conclusions of policy analysis based on modeling results. While the focus of the discussion is on the estimation of benefits and costs, the same sources of uncertainty can also affect estimates of economic impacts.

Specifically, the white paper begins with a brief description on uncertainties inherent in traditional engineering-based cost assessments followed by uncertainties that may be associated with incorporating this information in a CGE model. Subsequently, the paper provides a brief discussion of some of the uncertainties associated with traditional benefits assessment for regulations designed to improve air quality. The white paper then considers sources of uncertainty associated with CGE modeling. Following the introductory discussions on sources of uncertainty, the paper considers approaches for quantitatively assessing the impact of uncertainty on the results and methods for presenting those results in a manner that appropriately conveys the information relevant for decision makers and stakeholders. Finally, the paper considers approaches of model validation and verification for CGE models to further increase confidence in modeling results and reduce potential sources of uncertainty.

2 Uncertainty in Cost Estimation

Economy-wide analysis of air regulations will in part be informed by traditional engineering cost estimates of pollution abatement. Uncertainty in the results of these engineering analyses will influence the overall degree of uncertainty in the results of economy-wide analysis. This section examines uncertainties arising from the estimation of engineering costs that may be used as inputs to CGE models, when CGE models are used to perform regulatory analysis. This section begins by examining uncertainty in the cost inputs from the perspective of the cost analyst or engineer who typically develops the regulatory cost estimates used in regulatory actions. Subsequently, the section discusses uncertainties that may arise using these estimates to introduce a regulatory shock into a CGE model.

2.1 Uncertainty in Regulatory Cost Estimates

Entities affected by a regulation may incur compliance costs in the process of mitigating pollution to comply with the regulation. The largest components of compliance costs are typically the capital and operating costs associated with pollution control equipment. Capital costs are often one-time expenditures related to the installation or retrofit of structures or equipment to reduce emissions. Operating costs are recurring annual expenditures associated with the operation and maintenance of the pollution control equipment and will often include monitoring, reporting and recordkeeping expenditures. Administrative and enforcement costs may also accrue to local, state, and federal regulatory agencies.

Analysts typically estimate compliance costs expected to be incurred by regulated entities *before* the rule is implemented. In many cases, these ex ante estimates of compliance costs are estimated by cost

engineers and policy analysts to fulfill statutory obligations outside of benefit cost analysis or purposes of evaluating the national economic impacts of the regulation. As has been discussed by the National Academies of Science (NAS, 2013), these ex ante estimates of compliance costs are often based on engineering models in which there is uncertainty over facility characteristics, the number of affected facilities, and the degree of regulatory compliance. This uncertainty often leads to cases in which the ex ante estimates of compliance costs differ from the ex post (realized) compliance costs.

The “Retrospective Study of the Costs of EPA Regulation” highlighted a number of reasons that ex ante and ex post compliance costs may differ (Kopits et al., 2014). For example, technological innovation and unforeseen compliance options might also cause ex ante costs to diverge from ex post compliance costs. Unanticipated changes in other exogenous factors, such as changes in energy prices or demand shifts independent of regulation, may also cause ex post compliance costs to diverge from the ex ante estimates. Furthermore, strategic behavior by firms can influence ex ante compliance cost estimates, as much of the information used in estimating compliance costs comes directly from industry.

Since EPA compliance cost estimation is primarily performed by cost analysts and engineers, it is useful to review the framework developed by those professions to articulate the accuracy of engineering cost estimates. For example, the EPA Air Pollution Control Cost Manual (2002) is described as a comprehensive set of “procedures and data for sizing and costing control equipment” for VOCs, PM, SO₂, NO_x, and some acid gases. This manual as well as the ACE International Cost Estimating Classification System² have been used to provide context for the uncertainty associated with pollution control cost estimates for compliance with National Ambient Air Quality Standards (NAAQS).

In the 2015 Final Ozone NAAQS Regulatory Impact Analysis (RIA), for example, EPA stated that there is a range of ± 30 percent for non-electrical generating unit point source control costs, citing the EPA Air Pollution Control Cost Manual (2002). This level of accuracy is described in the EPA Air Pollution Control Cost Manual as a “study estimate.” According to the Manual, the study estimate is well suited for use in regulatory development because it does not require detailed site-specific information necessary for industry level analyses. They also can be prepared at a relatively low cost with relatively minimal data. While information that is more detailed may be available to the EPA, it is often proprietary and considered confidential business information.

In addition to the study estimate, the EPA Air Pollution Control Cost Manual discusses other types of estimates:

- **Order-of-magnitude:** This estimate provides a rule-of-thumb procedure applied only to repetitive types of plant installations for which there exists good cost history. Its error bounds are greater than ± 30 percent. The sole input required for making this level of estimate is the control system’s capacity.
- **Scope, Budget Authorization, or Preliminary:** This estimate, nominally of ± 20% accuracy, requires more detailed knowledge than the study estimate regarding the site, flow sheet,

² ACE International. Recommended Practice No. 17R-97. Cost Estimate Classification System – Cost Estimating and Budgeting. Revised on November 29, 2011. Available at <http://www.acei.org/>.

equipment, buildings, etc. In addition, rough specifications for the insulation and instrumentation are needed.

- **Project Control or Definitive:** These estimates, accurate to within $\pm 10\%$, yet require more information than the scope estimates, especially information concerning the site, equipment, and electrical requirements.
- **Firm, Contractor, or Detailed:** This is the most accurate type of estimate ($\pm 5\%$) and requires complete drawings, specifications, and site surveys. Consequently, detailed cost estimates are typically not available until right before construction is commenced.

These levels of accuracy also correspond to those reported in Perry's Chemical Engineers' Handbook, another important reference for costs engineers and analysts.

The type of ex ante compliance cost estimate calculated by EPA prior to a rulemaking will be regulation specific, given the heterogeneity in available information and scope of regulations. For example, some rulemakings may affect relatively few facilities, and detailed information on facility characteristics may be available publically or through formal Information Collection Requests. In other cases, regulation may only affect new facilities where heterogeneity in facility characteristics that would affect retrofit costs are not a confounding factor. Compliance cost estimates for these actions may be more precise than for regulations like a NAAQS, where the illustrative controls strategies examined in NAAQS RIAs may require reducing emissions at a large number of facilities (potentially in the thousands) across multiple sectors over relatively long time horizon.

In its air-related RIAs, the EPA has limited experience with evaluating uncertainty in compliance cost estimates. However, some recent examples of where the Agency has examined the implication of uncertain factors on compliance costs include:

- The level of voluntary emissions reductions versus regulatory emissions reductions for oil and natural gas emissions sources (EPA 2012);
- whether states would adopt a rate-based or mass-based compliance approaches for meeting requirements of the Clean Power Plan (EPA 2015a);
- alternative assumptions on the costs of emissions controls for NAAQS-related cost analysis (EPA 2015b);
- and, the influence of important cost components such as the value of natural gas that is captured by emissions controls and routed to sales lines or used in processes to offset other fuel purchases (EPA 2012 and EPA 2016).

As noted in the "White Paper on Using CGE Models to Evaluate Social Cost of Air Regulations", regulatory analysis may also differ due to the available information on potential pollution control options. For example, in some cases pollution control equipment may already be in use within parts of the industry or closely related industries. In other cases, there may be uncertainty in how states or local governments may implement regulations and therefore in the facilities affected and the pollution control equipment that might be adopted. For example, in illustrative attainment analyses conducted for some NAAQS, once all identified control technologies have been applied some areas of the country may still be modeled as

out of compliance with the air quality standard. In some cases, EPA has had to assume the per ton compliance costs for “unidentified controls” that reduce emissions sufficiently to bring areas into attainment of alternative air quality standards. When considering the cost of unidentified controls, the EPA has traditionally offered sensitivity analyses with a less expensive and a more expensive cost per ton estimate. However, there is not a firm analytical basis for the values chosen for the sensitivity analyses; rather, they roughly correspond to the same +/- 30 percent range assumed for identified controls.

2.2 Uncertainty in implementation of Regulatory Costs in Economic Models

As discussed in “Economy-Wide Modeling: Social Cost and Welfare White Paper”, there are few examples in which a CGE model has been used to estimate the costs of regulations designed to improve air quality. In the cases where a CGE model has been used to analyze non-price regulations, such as the emissions limit and air quality regulations promulgated by the EPA, they are often modeled as a productivity shock. Specifically, compliance with the regulation creates a need for additional inputs to produce goods in the regulated sector along with pollution abatement. While the total cost of these additional inputs can be derived from detailed compliance cost estimates from an engineering or partial equilibrium model, it is not always clear how to allocate the total cost among the inputs specified in the CGE model, because CGE models are by their nature an aggregated, parsimonious representation of the economy. One frequently used approach is to allocate the abatement costs in direct proportion to the inputs (i.e., capital, labor, and intermediate goods) used in the regulated sector of the CGE model. In other words, regulatory requirements do not change the proportion of labor, capital, or other inputs in the firm’s production function. This “Hicks-neutral” allocation is the approach taken by Hazilla and Kopp (1990) and Jorgenson and Wilcoxon (1990). Ballard and Medema (1993) allocate all of the abatement costs to capital and labor inputs only. There is inherent uncertainty associated with the decision to model the production of pollution abatement as having the same production function as the affected industry. Even in cases where detailed compliance cost estimates provide additional information as to the shares of inputs expected to be used in pollution abatement there may be uncertainty as to how the inputs from these engineering cost models map to the goods within a CGE model. Furthermore, there may be uncertainty as to how the substitution possibilities between inputs to production in the regulated sector in the CGE model map to substitution possibilities in pollution control options that may be adopted under non-price based regulations, or how future technology development might be expected to change those substitution possibilities.

It is also important to consider both the spatial and temporal allocation of cost. There may be a mismatch between affected facility locations within input and output markets and the scale of the economy-wide model, which is typically regional or national in scale. That is, it may be difficult to precisely model changes that affect specific areas of the country in an aggregated model. As a result, this may lead to some uncertainty in the results of the aggregate model.

Likewise, costs may be incurred at different points during the time horizon of the economy-wide model. With a long time horizon, it is necessary to consider the role of technological change and its impact on the prices of inputs over time. Recent NAAQS RIAs include discussions of how technological change (specifically, innovation in control technologies) may reduce costs over time, and the empirical literature

also has noted that variable costs of production or environmental abatement tend to decline over time with cumulative experience. While the study on the costs and benefits of the Clean Air Act from 1990 to 2020 (EPA 2011) accounted for “learning curve” effects (or the extent to which the costs of a technology decline as experience with that technology increases), regulatory analyses often do not attempt to adjust costs to account for different assumptions about technological change or the effect of experience on control costs.³ As a result, when incorporating compliance costs in a CGE model, any impact of technological change will be determined by the representation of technological change that exists in the model, rather than assumptions that are incorporated in the engineering cost estimates. As noted by Pindyck (2007), characterizing the uncertainty over technological change is in itself difficult.

3 Uncertainty in Benefits Estimation

As would be the case with entering costs into a CGE model, uncertainty in any inputs related to the benefits of a regulation will propagate through the analysis and lead to some uncertainty with respect to results. While a complete discussion of the sources of uncertainty in benefits estimates is beyond the scope of this paper, this section provides a brief summary of EPA’s efforts to characterize uncertainty in the benefits estimates in recent RIAs.

3.1 Uncertainty in Regulatory Benefits Estimates

In any complex analysis using estimated parameters and inputs from numerous models, there are likely to be many sources of uncertainty. EPA benefits analyses include inputs from many data sources, including emissions inventories, air quality data from models (with their associated parameters and inputs), population data, population estimates, health effect estimates from epidemiology studies, economic data and parameters for valuing benefits, and assumptions regarding the future state of the world (e.g., regulations, technology, emissions, and human behavior). There may be uncertainty associated with each of these inputs. Understanding key uncertainties in each stage of the analysis, and how they might interact, can be important for understanding the information contained in the total quantified benefit estimates.

While the National Research Council (NRC, 2002, 2009) reviewed EPA’s methodology for calculating the benefits of reducing air pollution and found it to be reasonable and informative, they also highlighted the need to conduct rigorous quantitative analyses of uncertainty and to present estimates in a way that recognizes their inherent uncertainty. In response, EPA has continued to work to improve the characterization of uncertainty in health incidence and benefits estimates. Quantitative approaches to uncertainty analysis for benefits assessment in RIAs have included:

- Monte Carlo assessments that account for random sampling error and between study variability in epidemiological and economic valuation studies (for pollutants PM_{2.5} and O₃);

³ An exception to this is recent greenhouse gas emission standards for light-duty and for medium- and heavy-duty vehicles. See <https://www3.epa.gov/otag/climate/documents/420r12901.pdf> for one example of how learning curves have been applied in these contexts.

- alternative concentration-response functions (PM_{2.5} and O₃);
- alternative income elasticities in the specification of willingness-to-pay functions used for mortality and morbidity endpoints (O₃);
- inclusion of thresholds in some concentration-response functions (O₃);
- alternative cessation lags for long-term mortality (O₃);
- a concentration benchmark assessment characterizing the distribution of avoided deaths relative to specific concentrations in the long-term epidemiological studies used to estimate PM_{2.5}-related mortality (PM_{2.5}); and
- when monetizing climate benefits, the use of four possible measures for the social cost of carbon (SC-CO₂) and social cost of methane (SC-CH₄), reflecting a lack of consensus on the appropriate discount rate and to account for the possibility of higher-than-expected impacts from climate change.

While such techniques can provide valuable information about key uncertainties and how they influence benefits estimates, these approaches may still face challenges in accounting for the role of uncertainty in other input variables, including emissions and air quality modeling, baseline incidence rates, and population exposure estimates. Challenges also remain in addressing correlations between input parameters and fully characterizing input distributions. As a result, reported confidence intervals and the range of estimates may present only a partial picture of the overall uncertainty in the final estimates.

In 2009, EPA undertook a project to identify the input parameters and assumptions that have the potential to be significant contributors to the uncertainty in the benefits estimates produced by BenMAP, the primary tool EPA uses to estimate the human health impacts and economic value of air quality changes. To assess the impact of uncertainty in these parameters and assumptions, sensitivity analysis was conducted where a range of values for key parameters was defined, and their effects on results was calculated while holding all other parameters at their mid values. The study found that the components of a PM_{2.5} benefits analysis contributing most to uncertainty of the monetized benefits and mortality incidence are the estimate of the value of a statistical life, the choice of concentration-response function for mortality, and the change in PM_{2.5} concentration (Mansfield et al., 2009).

As noted in the “Economy-Wide Modeling: Benefits of Air Quality Improvements White Paper”, the benefits of regulations can vary significantly over space due to a number of factors, including the spatial heterogeneity in air quality impacts and the spatial variation in population and baseline incidence rates. A number of studies have analyzed the sensitivity of benefits estimates to the resolution of the air quality inputs and/or underlying population and incidence rate data. They have found that differences in spatial resolution, operationalized in the studies by varying the size of the grid cell within which air quality is assumed to be homogenous, can lead to substantially different benefits estimates (De Ridder et al. 2014, Fann et al. 2011, Kheirbek et al. 2013, Li et al. 2015, Pungler and West 2013, Thompson et al. 2014). However, the studies do not find a consistent bias in the results. Pungler and West (2013) found that very coarse grid resolutions (>250km) produce mortality estimates that are biased substantially low for PM_{2.5},

but they find that the mortality estimates for both PM_{2.5} and O₃ at 36km resolution are slightly higher than the results modeled at 12km resolution. Li et al. (2015) also found a negative bias in mortality estimates at very coarse resolutions. De Ridder et al. (2014) found a similar negative bias in exposure to NO₂ when moving from a finely resolved grid (1km) to a coarser resolution (64km). Thompson et al. (2014) found O₃ health impacts to be sensitive to resolution, but did not find a similar pattern with PM_{2.5}, at least when comparing 36km results to 12km results. Fann et al. (2011) and Kheirbek et al. (2013) found that very spatially resolved data allowed for better identification of areas of high vulnerability within cities.

3.2 Uncertainty in Introducing Air Quality Improvement Benefits in Economic Models

Setting aside uncertainties in the upstream elements affecting benefits estimates, any benefits analysis will produce a set of estimates that may include confidence intervals for some endpoints, may not completely account for all of the benefits (or disbenefits) of an action, and which may not be directly translatable to the inputs needed for an economy-wide model. As discussed in the “Economy-Wide Modeling: Benefits of Air Quality Improvements White Paper,” challenges remain in fully implementing the effect of air quality on the behavior and well-being of economic agents in a CGE model. To date, there has been some limited inclusion of air quality impacts in CGE models including changes in medical expenditures and effects on the labor force through a change in the time endowment. Expanding those categories of benefits or introducing a more complete treatment of air quality may represent notable challenges and be associated with uncertainty surrounding the implementing methodology. In these cases, there may be uncertainty over the structure of the model, or at least the components used to incorporate the additional benefit categories.

As was the case when considering the inclusion of costs in an economic model, it is important to consider both the spatial and temporal allocation of benefits. While more spatially resolved data may lead to different benefits estimates, the spatial resolution of the benefits estimate is typically much finer than the spatial resolution of any CGE model into which it may serve as an input. As a result, some aggregation of benefits estimates and/or air quality modeling will be necessary. The degree to which this aggregation influences CGE modeling results is an important question. Likewise, the timing of benefits affects the budget constraint and time endowment available to households, which may influence the results from a CGE model.

4 Uncertainty in CGE Modeling

Applied GE analysis, including with CGE models, has long been known to be sensitive to both structural and parametric assumptions. In some cases, these uncertainties may be important for the interpretation of policy analysis conducted using CGE models. Before discussing approaches to quantitatively modeling and presenting uncertainty over policy results, this section discusses types of uncertainty related to CGE models.

4.1 Parametric Uncertainty

The CGE models currently used in policy analysis often contain hundreds, if not thousands, of parameters, including those that define production technologies, household preferences, benchmark economic

activity, current policies, and changes in technologies and factor endowments over time. This parametric uncertainty may have implications for the type of analysis conducted and the interpretation of the results.

One well recognized area of uncertainty in CGE models are the estimates of elasticities that help define production technologies and agent preferences. Uncertainty around elasticity parameters has frequently been a focus of analysts as the modeling results are often highly sensitive to these parameters (e.g., Shoven and Whalley, 1984). The sensitivity of results to elasticities has been the subject of much discussion given the common approach of selecting the values through a calibration process as opposed to econometric estimation (Hansen and Heckman, 1996). However, selecting econometrically estimated parameter values from the literature is not without its own concerns due to inconsistencies between the empirical analyses and the structure of the CGE model and a large range of potentially contradictory studies that provide elasticity estimates (Canova, 1995). Beckman et al. (2011) demonstrated that use of current and well-researched substitution elasticities for calibrated models can be crucial for achieving defensible behavior of a CGE model, and demonstrated an approach for validating model behavior/parameterization using historical observations. Given that not all parameters of a model may be necessarily verified in such a process, sensitivity analysis is one means of trying to assess a reasonable range for the modeling results of interest given likely ranges of the input parameters.

The sensitivity of CGE results to assumptions and/or estimates of substitution elasticities has been well demonstrated. For example, Fox and Fullerton (1991) find that estimates of welfare changes associated with tax reform can be more sensitive to assumptions about elasticity parameters than the actual level of detail about the tax system included in the model. Elliot et al. (2012a) find that the results of applied CGE analyses are likely far more sensitive to uncertainty around assumptions about elasticity parameters than other data inputs such as the benchmark social accounting matrix.

In dynamic models, results have been found to be particularly sensitive to additional parameters beyond the substitution elasticities. For example, Webster et al. (2001) find that exogenous temporal projections, such as those defining labor productivity growth and autonomous energy efficiency improvement over the model's time horizon, notably influence results. Additional studies have found model results to be particularly sensitive to assumptions about future technology availability, when such technologies are directly related to compliance with the policy being analyzed (e.g., Clarke et al. (2014)).

Previous economy-wide analyses conducted by EPA have primarily examined the sensitivity of results to key policy parameters. For example, in 2009 members of the United States Congress requested EPA conduct an analysis of House Resolution 2454, which proposed a multi-sector allowance-trading program designed to reduce emissions of greenhouse gases.⁴ In that analysis, EPA examined the sensitivity of results from two CGE models (ADAGE and IGEM) under a number of scenarios in which key policy parameters (e.g., availability of offsets, energy efficiency provisions) were varied (EPA, 2010a). However, in that analysis and a supplementary analysis (EPA, 2010c), EPA also examined the sensitivity of results to key assumptions regarding the future availability of certain technologies (e.g., new nuclear power plants,

⁴ <http://thomas.loc.gov/cgi-bin/bdquery/z?d111:H.R.2454:>

carbon capture and sequestration). The results of the analysis were found to be sensitive to exogenous assumptions regarding the availability of technologies in the future.

4.2 Structural Uncertainty

The development of any CGE model involves many structural choices to address a particular question or set of questions. Common high-level structural choices include static vs. dynamic, myopic vs. forward-looking, the degree of sectoral and regional disaggregation, and closed vs. open economy. Other examples of structural choices include the trade specification (Armington, Heckscher-Ohlin, or Melitz), savings and investment closures, departures from the full employment assumption, CES nesting structure, and the form of the utility function (e.g., Cobb-Douglas or Stone-Geary/Linear Expenditure System). Banse (2013) in a review of nine global CGE models used for agricultural analysis list 16 structural specification options for labor markets and 25 options for capital markets.

In general, structural testing is more resource intensive than parametric sensitivity testing, though this depends upon the type of structural change and the range of parametric uncertainty. There are many examples of single model structural comparison testing in the literature. An illustrative sample of this literature follows. Balistreri and Rutherford (2012) examine the differences between Armington and Melitz treatments of international trade in a study of border carbon adjustments. They find higher leakage rates and more effective border adjustments under the heterogeneous firm structure of a Melitz approach. Jacoby and Sue-Wing (1999) and McFarland et al. (2004) show that the incorporation of vintaged capital in a CGE model raises the costs of reducing greenhouse gas emissions. Babiker et al. (2009) compares forward-looking and recursive-dynamic specifications for climate policy analysis. They note that the macro-economic costs are lower in the forward-looking version because of the ability to shift consumption optimally over time. However, the forward-looking model, due to computational limitations, also drops the full capital vintaging specification and contains fewer explicit emission reduction technologies. These factors likely have countervailing influences on the relative macroeconomic costs. EPA addressed structural uncertainty in the aforementioned economy-wide analysis of House Resolution 2454 by using two models: ADAGE and IGEM (EPA, 2010a). The models differ on several structural fronts. ADAGE is a calibrated model and IGEM is econometrically estimated. ADAGE has a more detailed representation of the energy sector and energy sector technologies; IGEM has more non-energy sector detail.

5 Analytical Approaches

Uncertainty associated with engineering cost and health and environmental benefit inputs, analytical assumptions, and model parameters and structure, may in some cases warrant additional analysis to better explore the expected impacts of a policy. In some cases, the specifics of the regulation in question may inform whether the additional resources required to perform additional uncertainty analysis is of sufficient value. For example, Circular A-4, which “provides the Office of Management and Budget’s (OMB’s) guidance to Federal agencies on the development of regulatory analysis,” suggests a formal quantitative analysis of relevant uncertainties be conducted for major rules that are expected to have “annual economic effects of \$1 billion or more” (OMB, 2003). This section considers different quantitative

approaches for considering uncertainty in policy analyses conducted using CGE models, including sensitivity analysis, formal uncertainty analysis, and inter-model comparisons. For ease of exposition, we delineate sensitivity analysis and formal uncertainty analysis, by defining the former as analyses that consider scenarios in which one or more inputs are varied outside of a well-defined probability space and the latter as analyses that attempt to characterize the distribution of model results.

5.1 Sensitivity Analysis

The most common approach to testing the robustness of modeling results is to vary an exogenous variable or parameter and solve the model to obtain the new set of endogenous state variables under the alternative parameterization. In this paper, such an approach is defined as sensitivity analysis to distinguish it from more formal uncertainty analysis (Section 5.2) and inter-model comparisons (Section 5.3). We use the term sensitivity analysis as inclusive of comparative statics and dynamics, along with situations in which the models allow for the existence of temporary disequilibrium.

It is common for sensitivity analysis to vary a single parameter in the alternative scenarios to isolate the impact it has on the results of interest. This approach allows analysts to gather information about how robust qualitative conclusions, based on modeling results, are with respect to specific assumptions or key parameters. EPA's *Guidelines for Preparing Economic Analysis* (2010b) suggests that "[i]n cases where the data are uncertain, or not easily quantified, but may have a significant influence on the results, the analyst should describe the weaknesses in the data and assumptions, and include some type of sensitivity analysis."

Sensitivity analysis that varies only one, or a few, parameters at a time could potentially provide an incomplete characterization of the uncertainty surrounding the results due to important interactions between parameters within complex and non-linear CGE models (Abler et al., 1999). By its nature, basic sensitivity analysis is limited by the fact that is normally conducted outside of a well-defined probability space that guides which parameters to vary, and the magnitude by which they should be perturbed. However, in cases with incomplete information about the complete distributional parameter space and/or limited resources and time, basic sensitivity analysis may provide useful information as to the robustness of modeling results and provide valuable information for decision makers (Pannell, 1997). Furthermore, as noted by OMB (2003) in Circular A-4:

"In some cases, the level of scientific uncertainty may be so large that you can only present discrete alternative scenarios without assessing the relative likelihood of each scenario quantitatively. For instance, in assessing the potential outcomes of an environmental effect, there may be a limited number of scientific studies with strongly divergent results. In such cases, you might present results from a range of plausible scenarios, together with any available information that might help in qualitatively determining which scenario is most likely to occur."

While this description reinforces the benefits of sensitivity analysis, it also highlights the challenges present in conducting sensitivity analysis, as concepts such as a range of plausible scenarios are not formally defined. However, it is possible to bring some structure to sensitivity analysis. For example, the Congressional Budget Office (CBO) approaches sensitivity analysis in its macroeconomic estimates using

a defined process to address the lack of structure. Specifically, in its dynamic scoring approach CBO determines the two parameters in the model to which the results are most sensitive. Then potential estimates are generated by examining each case in which those two parameters are at the ends of their ranges and other parameters are equal to central estimates, where CBO ultimately reports the cases that show the most and least favorable budgetary outcomes.⁵ It is also worth noting that CBO does not conduct a quantitative uncertainty analysis for each bill considered if the range derived from the analysis could potentially be misinterpreted. For example, in cases where the underlying conventional cost estimate is associated with significant uncertainty relative to the uncertainty in the macroeconomic model, conducting sensitivity analysis over only the parameters in the macroeconomic model could produce a range of estimates that should not be interpreted as characterizing the full range, given significant additional uncertainty in the conventional cost estimate (CBO, 2015).

Sensitivity analysis may also be used to examine uncertainties with respect to the structure of CGE models and how that may affect the results of policy analysis. Unlike parametric sensitivity analysis, structural sensitivity analysis examines the effect of changes to the underlying equations in the model. This type of analysis typically takes place within a single model. In general, structural testing is more resource intensive than parametric sensitivity analysis, though this depends upon the type of structural change and the range of parametric uncertainty.

5.2 Formal Uncertainty Analysis

Sensitivity analysis of a few select parameters, as described in Section **Error! Reference source not found.**, may allow an analyst to get a general sense of how the model's results may depend on key parameter values. However, due to the lack of a methodological structure, which leads the approach to ignore information contained in the covariance of uncertain parameters, this approach is necessarily imprecise (Bernheim et al., 1989). As a more structured alternative, modelers have considered formal uncertainty analysis that takes into consideration additional information about the uncertain parameters.⁶

The current class of applied general equilibrium models used for policy analysis have hundreds of input parameters. Even for a relatively small model with eight parameters that each have five potential values, if the model took one minute to solve and save the results a complete factorial experiment design that considered every possible combination of parameters would require nine months of processing time (Pannell, 1997). It is likely that with modern computing resources models that solve in one minute likely have far more than eight uncertain parameters making Pannell's example even more apt. It is also likely that parameters are more appropriately characterized by distributions than by a discrete set, as in the example set forth by Pannell (1997). Therefore, Gaussian quadrature (e.g., Hertel et al., 2007), Monte Carlo (e.g., Selin et al. 2009), and linearization (e.g., Jorgeson et al., 2013) methods are more commonly used to conduct forms of probabilistic analysis. These techniques provide ways of approximating integrals over uncertain model parameters to obtain expected values and confidence intervals for model results.

⁵ <http://www.slideshare.net/cbo/dynamic-scoring-at-cbo-51635156>

⁶ Approaches to more formal and structured uncertainty analysis have also been described as “systematic sensitivity analysis.”

These benefits come at the cost of additional computation and need to define probability distributions over the uncertain models inputs and/or parameters. This section describes these various methods of uncertainty analysis and examples in which they have been applied.

The endogenous variables, Y , in a CGE model may be represented as the result of the implicit function

$$Y = H(X, \beta, \theta), \quad (1)$$

where X are exogenous variables and known parameters, including policy variables, β are uncertain parameters of the model, excluding calibration parameters which are defined by θ . The calibration parameters, θ , are chosen such that the model replicates benchmark conditions, Y_0 , such that

$$\theta = h(X_0, \beta, Y_0), \quad (2)$$

where

$$Y_0 = H(X_0, \beta, h(X_0, \beta, Y_0)). \quad (3)$$

In most cases, the goal of the analyst is to determine the expected change in endogenous variables of interest given a change in policy parameters, represented as a movement from X_0 to X_1 , and uncertainty in the model's parameters β ,

$$\begin{aligned} E[Y_1 - Y_0] &= E[H(X_1, \beta, \theta) - H(X_0, \beta, \theta)] \\ &= E[H(X_1, \beta, \theta)] - Y_0 \end{aligned} \quad (4)$$

Given the nonlinear nature of most CGE models it will be the case that

$$E[H(X_1, \beta, \theta)] \neq H(X_1, E[\beta], \theta), \quad (5)$$

such that evaluating the change in model output under the policy at the expected values of the parameters will not be equal to expected change when fully considering the parametric uncertainty. Probabilistic analysis allows analysts to provide more robust assessments of the expected change in economic variables by integrating over the joint distribution of the uncertain parameters, $g(\beta)$,

$$E[Y_1] = \int_{\Omega} H(X_1, \beta, \theta) g(\beta) d\beta \quad (6)$$

and provide a variance associated with model's results by evaluating

$$E[(Y_1 - E[Y_1])^2] = \int_{\Omega} \{H(X_1, \beta, \theta) - E[Y_1]\}^2 g(\beta) d\beta. \quad (7)$$

Given the nature of applied CGE models, analytical solutions for (6) and (7) do not exist and numerical approximations are required. A common approach for numerically approximating a sampling distribution of Y_1 from which sample moments can be calculated is a Monte Carlo simulation. The basic process

involves conducting N simulations of Y_1 , where each time a new set of parameters, $\hat{\beta}_n$, is drawn for the joint probability distribution $g(\beta)$. The expected values of the endogenous variables in (6) may then be estimated by the sample mean

$$E[Y_1] \approx \frac{1}{N} \sum_{n=1}^N H(X_1, \hat{\beta}_n, \theta), \quad (8)$$

with the variance calculated analogously. The accuracy of the approximation will, in part, depend on the number of simulations conducted, with the error being of order $1/\sqrt{N}$ and therefore requiring a potentially large number of simulations.

In cases of highly complex models, or relatively scarce computing resources, the number of simulations required to get reliable estimates of the moments around the variables of interest can be of concern. To address the computational burdens, Harrison and Vinod (1992) suggest approximating the underlying probability distribution with a discrete set of points and probabilities, allowing for faster convergence. However, their approach of choosing the points for each parameter by dividing the support of the marginal probability distribution into a finite number of equi-probable regions, is known to understate the variance (and all other higher order even numbered moments) of the parameter distribution (Miller and Rice, 1983). This behavior is inherited by the approximate joint distribution of the parameters and when used in the Monte Carlo approach will bias the results (De Vuyst and Preckel, 1997).

As an alternative to Monte Carlo simulations with CGE models, Arndt (1996) and DeVuyst and Preckel (1997) introduce an approach based on the Gaussian Quadrature to approximate the integrals in (6) and (7). The general principle is to carefully choose a series of J weights, w_j , each associated with a set of parameters, β_j , from the joint probability distribution, such that the expectation may be approximated as

$$E[Y_1] \approx \sum_{j=1}^J w_j H(X_1, \beta_j, \theta), \quad (9)$$

where $\sum_{j=1}^J w_j = 1$ (with similar approximations holding for higher order moments). By carefully selecting the sets of parameters and corresponding weights, the analyst may, in some cases, be able to approximate the moments around the model outputs with far fewer model runs than under the Monte Carlo approach, $J < N$. In general, the basis for selecting the weights and corresponding parameters is to ensure that the moments about zero, up to a given order, would be adequately integrated using the selected values. By using higher order quadrature techniques greater accuracy may be achieved when integrating complicated systems of non-linear equations, such as a CGE model, though the ability to approximate the integral depends on the ability of a polynomial to approximate the curvature of the model.

In their original application, DeVuyst and Preckel (1997) use a global, energy-explicit CGE model with uncertain fossil fuel supply elasticities to evaluate the impacts of a carbon tax. The authors demonstrate

the approach with six uncertain parameters from independent distributions, and find that they can obtain a reasonable approximation of the mean and standard deviation of the model's endogenous variables with relatively few model runs. Since their original demonstration, Gaussian Quadrature has been used to estimate expected values and confidence intervals for key results from CGE models in more complex contexts. For example, Hertel et al. (2007) use the technique to derive means and confidence intervals from a global CGE model with dozens of uncertain parameters, when studying the impact of reducing international trade barriers. However, to apply the method they assume that all uncertain parameters are normally distributed and independent.

For common distributions and in the case of independent uncertain parameters, techniques for selecting weights and parametric values are readily available (Miranda and Fackler, 2002). The original application using this technique with CGE models focused on cases with symmetric and independent distributions for the uncertain parameters. For the case of the non-symmetric distributions, non-diagonal covariance matrices, and/or large parameter spaces techniques for Gaussian Quadrature may not be readily available or may not provide the same computational improvement over Monte Carlo simulations. Such limitations may be important, as Jorgenson et al. (2013) demonstrates that assuming independence of uncertain parameters in a CGE model can significantly bias confidence intervals around endogenous variables downward. However, Horridge and Pearson (2011) demonstrate an approach to applying Gaussian Quadrature in the context of CGE models in cases with non-diagonal covariance matrices across the uncertain parameters. The authors demonstrate the feasibility of the method using an example from the Global Trade and Analysis Project.

Another alternative to Monte Carlo simulations is based on the linearization of the CGE model (e.g., Pagan and Shannon, 1985; Wigle, 1991). This approach is sometimes referred to as the Delta method. For example, Bernheim et al. (1989) and Tuladhar and Wilcoxon (1998) use this approach to derive confidence intervals around endogenous state variables in a CGE model. The method seeks to provide the analyst with a confidence interval around the endogenous variable evaluated at the central tendency of the uncertain parameters. The method is based on linearizing the model using a first order Taylor series expansion around the endogenous variables with respect to the model's parameters.⁷ Under this approximation, a covariance for model's point estimate is derived as the product of the model's Jacobian with respect to its parameters and the covariance matrix for the parameters. Using the notation above, the model is linearized as

$$H(X_1, E[\beta], \theta) \approx H(X_1, \beta, \theta) + J_\beta \Sigma_\beta J'_\beta, \quad (10)$$

where J_β is the Jacobian of the CGE model with respect to the uncertain parameters and Σ_β is the covariance matrix associated with the uncertain parameters. The second term in (10) provides the analyst

⁷ Higher order approximations may also be applied to cases where there is significant curvature in the implicit function relating the level or change in the endogenous variable to the uncertain parameters (Bernheim et al., 1989).

with a covariance for the endogenous variables based on the model's sensitivity to those parameters and their degree of uncertainty.

Given the complexity of most applied CGE models, the Jacobian is computed by means of numerical differentiation and in most cases modelers have used a basic forward finite difference to minimize the number of model solves required to derive the Jacobian. Jorgenson et al. (2013) have shown that, even in the case of complex dynamic CGE models, the method can provide standard errors as a percentage of the model's endogenous variables that are reasonably close to the standard errors as a percentage of the expected endogenous variables in the case of Monte Carlo simulations. Their results suggest that in some cases, depending on the model, the Delta method may be able to derive relative standard errors with fewer simulations than a Monte Carlo approach.

With the increased availability of computing resources and simulation techniques, Monte Carlo simulations have been used in numerous recent studies to explore the sensitivity of modeling results (e.g., Webster et al., 2001; Sokolov et al., 2009; Elliot et al. 2012a; Elliot et al. 2012b). For example, Elliot et al. (2012a) take advantage of large-scale parallel processing to conduct Monte Carlo simulations that explore the sensitivity of CGE model results to parametric uncertainty in the case of a global CGE model.⁸ Notably, they find that the bias associated with model results based on running the CGE model at the central tendency of all parameters versus the case with full uncertainty in parameters is small. However, they find the variance around the CGE results was still significant. In a bootstrap analysis of their results, Elliot et al. (2012a) find that the number of simulations required to obtain information about the shape of the distribution around endogenous variables might vary considerably depending on the uncertain parameters considered and the endogenous variables of interest. For example, Elliot et al. (2012a) find that endogenous variables with lower coefficients of variation require fewer simulations to adequately capture the mean and variance of the result, compared to variable with relatively higher coefficients of variation.

Basic Monte Carlo analysis relies on the strong law of large numbers and the central limit theorem for its convergence properties. In terms of estimating the mean of the endogenous variables the approximation error tends to zero as the number of simulations tends to infinity; however, in practice the approximation error is random due to the probabilistic nature of the analysis and can be large even when the number of simulations is large. In recent Monte Carlo studies, researchers have used quasi-Monte Carlo techniques to more efficiently sample the parameter space and therefore reduce the number simulations required. The goal of these approaches is to impose additional structure on the method used to draw random parameters for the analysis, such that the approximation error is reduced beyond what it would be in a basic Monte Carlo analysis.⁹ For example, Webster et al. (2001) use Latin Hypercube sampling to reduce the number of runs required for a desired degree of accuracy.

⁸ Similar to other studies, Elliot et al. (2012a) assume the uncertain parameters are normally distributed and independent.

⁹ See Judd (1998) for details on quasi-Monte Carlo techniques.

Building upon Monte Carlo techniques, variance decomposition methods, such as the Sobol method (Sobol, 1993), produce sensitivity indices that estimate the fractional contribution of an input parameter (and parameter interactions) to the variance of an output. These methods allow the modeler to trace the dependencies between outputs and inputs. However, they can be computationally intensive (Iooss and Lemaitre, 2015). Although variance decomposition methods have been employed in environmental modeling and chemical engineering, they have not been widely used in conjunction with CGE models. However, in one example Mohora (2006) employs the Sobol method in a study of the Romanian economy and finds that the variance decomposition identified that indirect effects and database structure contribute most to output variance. Interestingly, Mohora found these results to be largely consistent with a simpler parameter screening method used in the study.

5.3 Inter-Model Comparison Exercises

Inter-model comparison exercises are a means to examine the effect of both structural and parametric differences between models. These exercises convene several modeling teams to examine policy and modeling issues by comparing model results for a standardized set of scenarios. Inter-model comparison exercises can serve many purposes. The exercises bring together modelers from academia, industry, government, and other organizations to examine an issue from different perspectives and using different tools. The section will discuss the basics of inter-model comparisons projects, past examples, and important considerations to take into account when designing such exercises.

Inter-model comparison exercises are structured around a set of policy-relevant research questions. Each of the modeling teams produce results for a common set of scenarios (e.g., with and without a policy or a technology). The scenarios and results are developed and analyzed typically over a series of two to three workshops held several months apart. These exercises serve several valuable functions. First, a multi-model approach to a topic may highlight areas of robust agreement as well as identify and the key factors driving disagreement across models. Second, the exercise helps to explain the strengths and limitations of alternative modeling approaches. Third, the exercises provide a forum for modelers to receive peer feedback on modeling issues that range from data and parameter selection to structural specifications and modeling techniques. Finally, the exercises identify future research and model development needs.

The longest running and a well-known series of inter-model comparisons is Stanford University's Energy Modeling Forum (EMF) co-directed by John Weyant and Hillard Huntington¹⁰. Begun in the mid-1970's EMF has completed nearly 30 projects on issues related to energy and its relationship to the economy and environment. Past projects have explored the oil, gas, coal, and electric power markets. EMF exercises within the last fifteen years have emphasized the interaction between energy markets, climate change, and climate change policy. Several recent inter-model comparison exercises have also focused on energy-economy-climate interactions (see e.g., AMPERE (Kriegler, et al. 2015) and LIMITS (Kriegler et al., 2013) and agriculture-economy-climate interactions (see e.g., AgMIP (Rosenzweig, 2013)).

¹⁰ See <https://emf.stanford.edu/>.

Although inter-model comparison exercises are valuable to both the policymaking and modeling communities, there are important challenges. For example, it can be difficult to discern the key factors that lead to the range of results across models because the models may differ by type (e.g., general vs. partial equilibrium) in addition to structural and parametric differences. One notable exception to this is EMF 29: The Role of Border Carbon Adjustment in Unilateral Climate Policy (Bohringer et al., 2012). EMF 29 was entirely comprised of CGE models that used the same underlying dataset (GTAP 7.1). In addition, caution should be taken when interpreting the range of results from comparison exercises because the range typically does not represent the full range of uncertainties. Instead, the range of results represent the uncertainties across the participating models and possibly over a handful of parameters or assumptions in the study. Finally, a study may not necessarily reflect all models as participation is voluntary and may not be supported by external funds.

6 Presentation of Uncertainty Analysis

While additional sensitivity or formal uncertainty analysis can provide useful information as previously described, it also presents a challenge to analysts who must summarize this additional information in a format that is readily understood by policy makers and stakeholders. When uncertainty analysis is conducted, Circular A-4 guides analysts to try to “provide some estimate of the probability distribution of regulatory benefits and costs” (OMB, 2003). However, there are different formats in which such information can be conveyed. Information about the probability distribution of benefits and costs can be presented numerically, verbally or graphically (NAS, 2013).

A numerical representation may include summary statistics, such the central tendency (e.g., mean and median), variances, confidence intervals, or other potentially relevant characteristics. A numerical presentation of the results can provide a significant amount of information, however, the approach may require that the audience have the appropriate expertise to understand and interpret the information. Verbal representations convey information using cues such as likely or unlikely to differentiate potential outcomes and characterize the probability distribution of benefits and costs. While verbal representations of uncertainty may be more approachable to a broader audience, there is the potential for the interpretation to differ across individuals. The United Nations Intergovernmental Panel on Climate Change (IPCC) has attempted to address this concern by establishing specific definitions and use guidelines for the verbal characterization in their assessment reports (IPCC, 2010). Graphical representations of probabilistic information have the potential to convey more information than basic verbal cues while remaining accessible to broader audiences. Graphical representations can take different forms such as histograms or color wheels. For example, Webster et al. (2008) present some results of their probabilistic CGE analysis as probability distributions. Figure 1, which is from their analysis, presents the loss in global consumption in a given simulation year across different policy stringencies (i.e., Level 1 through 4).

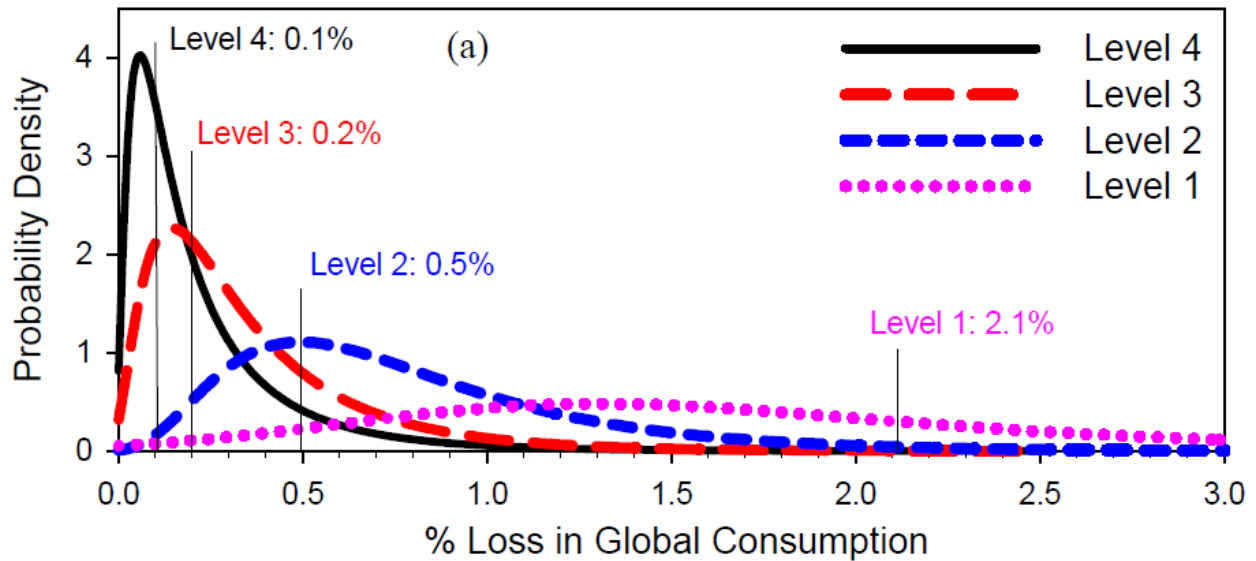


Figure 1: Distribution of Consumption Loss across Stringencies in EPPA CGE Model (Webster et al., 2008)

In the case of dynamic analyses, it can be useful to depict the results in a way that provides information as to how the distribution of results changes over time. For example, in their CGE analysis Elliott et al. (2012a) present the mean estimate of a key output variable, CO₂ emissions intensity, over time along with the different confidence intervals that are differentiated by their shading (Figure 2).

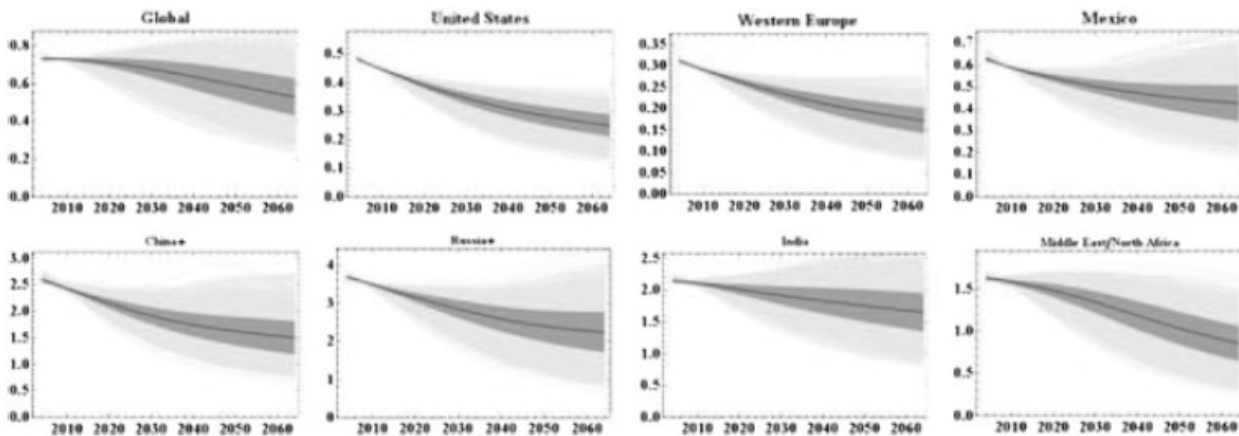


Figure 2: Carbon Intensity [kg CO₂ / 2004\$ of GDP] from CIM-EARTH CGE Model (Elliott et al., 2012a)

A well-known example of color wheels to convey probabilistic information is the work by MIT's Joint Program on the Science and Policy of Global Change on communicating uncertainty surrounding climate change (Figure 3).¹¹ This plot presents the probability of different climate outcomes with and without policy as two different pie charts where the area of the slices corresponds to the probability of the outcome conditional on the policy assumption associated with the pie chart. The graphic eases

¹¹ See <http://globalchange.mit.edu/focus-areas/uncertainty/gamble>.

interpretation of the outcomes and comparison across the two scenarios by coordinating the color-coding of the pie slices.

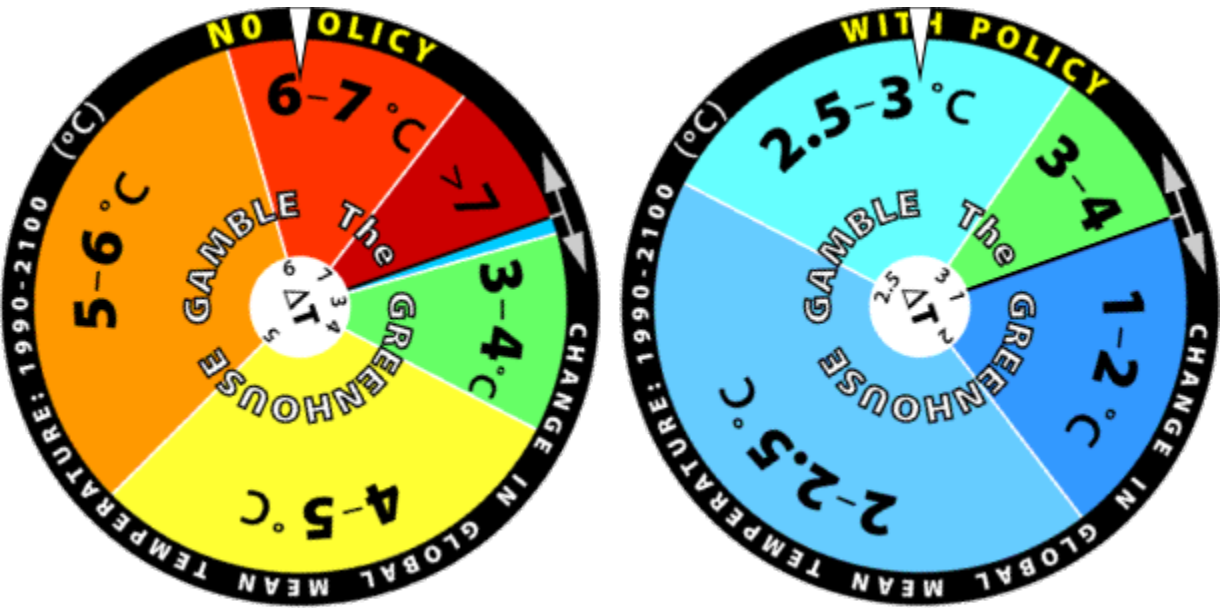


Figure 3: MIT Greenhouse Gamble Representation of Climate Change Uncertainty¹¹

In addition to presenting the results of uncertainty analysis, analysts may choose to provide additional information as to the sources of uncertainty most relevant to the results. For example, Tornado plots may also be useful in communicating the results of sensitivity analysis to show the parameters of greatest influence over the results of interest. Jacoby et al. (2006) conduct a study with the EPPA CGE model in which they vary key parameters, one by one, by plus and minus one standard deviation. They then plot the effect of these experiments on the welfare change of a given policy in order of greatest to least influence (Figure 4).

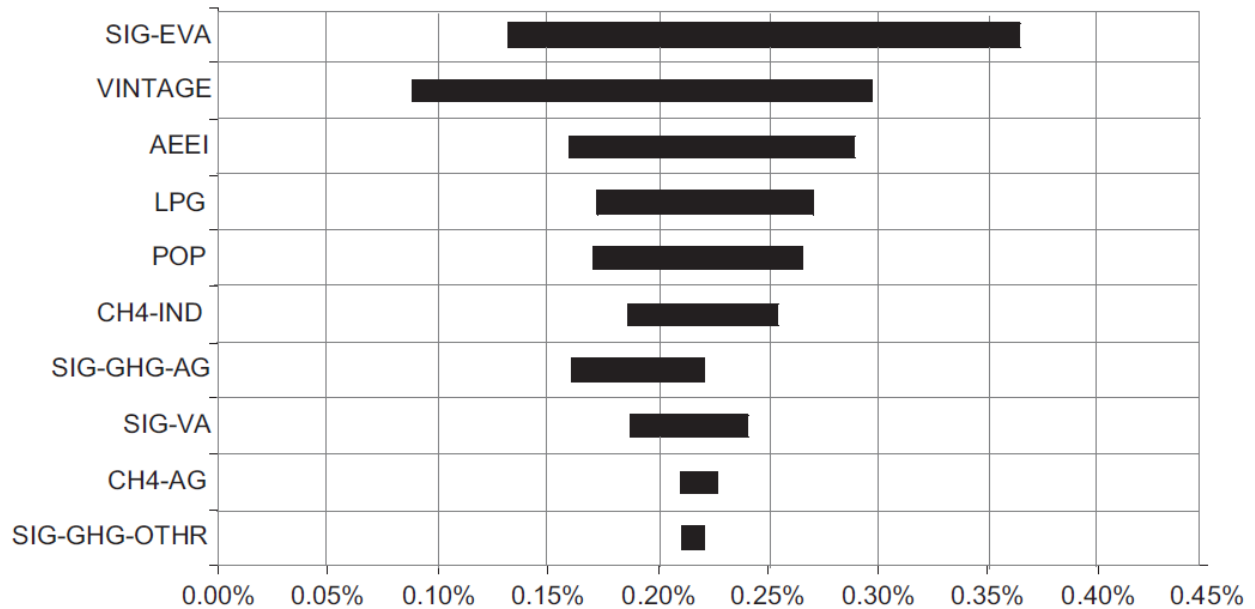


Figure 4: Change in Welfare for a Policy with Plus and Minus One Standard Deviation Change in the Parameter with the EPPA CGE Model (Jacoby et al., 2006)

However, tornado plots can also be used to present the results of probabilistic analyses. For example, standardized regressions may be used to analyze the results of Monte Carlo simulations, as described in Section 5.2, to determine the influence of different parameters on the results of the analysis. After standardizing the draws for the input parameters, they may be regressed on the output variable of interest to obtain standardized regression coefficients that represent the impact of a one standard deviation change in the parameters on the output variable of interest. The results of such additional analysis are often presented in the form of tornado plots.

7 Verification and Validation Exercises

The methods of conducting and presenting uncertainty analysis discussed in Sections 5 and 0 are approaches to providing decision makers and stakeholders with additional information about the robustness of model results to help instill confidence in the qualitative conclusions of policy analysis based on modeling results. Additional confidence in the model results may be provided through formal approaches to model verification and validation. Model verification and validation represent two important, interrelated activities in the CGE model building process. Carson (2002) defines model verification as a “process and technique that the model developer uses to assure that his or her model is correct and matches any agreed-upon specification and assumption.” Sargent (2013) defines model validation as the “substantiation that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model.” This section discusses techniques for model verification and validation with a focus on CGE models.

7.1 Model Verification

Verification of a CGE model constitutes the initial phase of the modeling project (Carson, 2002). This process involves a number of iterative procedures that include, among other things, model specification, data organization, and debugging of the model code. The modeler will usually perform a number of tests and simulations to examine basic features of the underlying data and economic structure. For example, a common initial experiment attempts to test if the model code and the imbedded assumptions replicate the benchmark dataset. If the model is specified correctly, the equilibrium solution to the model with no shocks should be identical to the baseline data. This check can be performed at each stage of the baseline setting process. For example, Rausch and Rutherford (2008) describe a set of programs that transform IMPLAN data on the U.S. economy into a format that can be used in a CGE model implemented in the GAMS/MPSGE modeling language. After each step in the process, the benchmark is tested in a small-scale, highly aggregated, open-economy model. These steps include aggregating the data, translating parameter names, diagonalizing the data so that each industry produces only one commodity, merging state level data, and balancing interregional trade flows. If the benchmark is not replicated at any point, the programs stops and an error message prompts the user to correct the imbalance before moving on.

Dixon and Rimmer (2013) describe a basic homogeneity test that may be performed in the development of some CGE models. If all model sectors are represented with constant returns to scale (CRTS) production functions, then increasing the exogenous nominal variables by a percentage should return the same percentage change in the endogenous nominal values. For instance, doubling the costs of all inputs to production should result in a doubling of the unit cost of output. If the model contains representations of nominal rigidities, such as sticky wages, adherence to homogeneity should not be expected. The effects of these deviations can be checked by turning off the features that cause the rigidity, such as a representation of sticky wages, to test if the feature is causing the deviation from the homogeneous case. Other deviations in the homogeneity test may represent errors in the model.

CGE model verification can also be conducted using GDP accounting identities. GDP can be computed from CGE modeling results by adding up total incomes or total expenditures, where both calculations should produce the same value. Each of these calculations can also be done in real or nominal terms. These checks are useful because they involve different sets of model variables, and are therefore testing for consistency across the model. The expenditure calculation involves the macroeconomic variables consumption, investment, government expenditures, and net exports. The income calculation includes factor incomes for all agents, transfer payments, and tax revenues.

The social accounting matrix (SAM) used to calibrate the model provides an additional check on the model. The SAM includes all payments between industries and agents, as well as international trade balances, and should balance before and after a policy shock. This means that the zero-profit, market-clearance, and income balance conditions hold for each industry, good, and agent in the model. (For more information on SAMs see, *inter alia*, Pyatt and Round (1985).)

7.2 Model validation

While it has been argued that a simulation model can never be completely validated (Gass, 1983), subjecting a model to validation tests – and making appropriate adjustments to the model if necessary – can increase confidence in the usefulness of the model for policy analysis. A number of methods for model validation have been used by CGE modelers. This section describes several of these methods and a number of exercises that have been conducted to validate (or invalidate) CGE models, along with the costs of model validation. Validation exercises could be performed when a model is first developed, at the time of major revisions, or simply over time to check the model has continued validity.

7.2.1 Back of the Envelope Models

Dixon and Rimmer (2013) suggest that back of the envelope (BOTE) models can be useful tools for CGE model validation. These estimates can give a sense of the magnitude of the expected results from the full model. A BOTE model simplifies the large-scale model to gain insight on model results. This simplification can be done by aggregating the number of industries, factors of production, and/or agents in the model. After a simple model is calibrated to the aggregated data, the modeler can add policy variables to test to see if the expected magnitude of changes in the model results are produced. Next, the modeler can add one feature at a time to the BOTE model to test the performance of these features of the large-scale model and to gain insight about how each feature influences the results of sample scenarios.

Simple BOTE estimates of the GDP impacts of a policy shock can be calculated using data and parameters, such as industry size, tax rates, and elasticities. Each of the components of the GDP calculation can be used to help explain model results and the overall impact of a policy on GDP. Policies that affect economic activity can influence the components of GDP in different directions. We may expect that a rule that requires installation of new capital would increase the investment component of GDP and decrease the consumption component. Once the macroeconomic impacts of the simplified model are well understood with the BOTE model, the full CGE model can be used to understand the effects on specific industries and households.

7.2.2 Statistical Validation of Model Parameters

Dixon and Rimmer (2013) suggest the use of statistical techniques, such as regression analysis, to improve the understanding of microeconomic results. They use the example of an analysis of employment changes under a scenario in which import tariffs and quotas are removed. The USAGE CGE model forecasts employment changes by state, which are then compared to a regression analysis that seeks to forecast the same changes. The regression analysis was prepared for this validation exercise using historical data to estimate a relationship between the trade policy and employment. By ranking the relative impacts by state and comparing these rankings across the two models, the regression exercise allowed the authors to see where states were represented differently in each model and ultimately improve the CGE model.

Ideally, CGE models should use general equilibrium elasticities that account for the degree of flexibility available across many parts of the economy. Values that are estimated in a partial equilibrium framework can cause the model to be too responsive or not responsive enough to policy scenarios, when compared to historical outcomes analyzed during the validation process.

7.2.3 CGE Model Validation Exercises

Although the published literature is relatively sparse, a number of authors have undertaken validation exercises with their models. Recent efforts have used increasingly sophisticated techniques that may allow for a process by which significant modeling improvements can be made.

Kehoe et al. (1995) compared results from their CGE analysis of fiscal reform in Spain – replacement of a complex system of indirect taxes with a value-added tax – during 1985-1987 with the changes that actually occurred. They compared changes in consumer and industrial prices, industrial output, and a number of macroeconomic variables using two performance metrics: the weighted correlation coefficient and weighted R^2 . Kehoe et al. (1995) report that their model performed well, particularly when they included two exogenous shocks that affected the Spanish economy during the period of analysis (a fall in the price of petroleum and a decline in agricultural productivity due to adverse weather conditions). In addition, they found that the results were robust to alternative specifications of the labor market and macroeconomic closure rules.

Kehoe (2005) evaluated the performance of three multisector static CGE models used to predict the impact of the North American Free Trade Agreement (NAFTA). Kehoe (2005) reports that all of the models dramatically underestimated the impact of NAFTA on trade amongst the U.S., Canada, and Mexico. The models were also reported to have only been able to capture relative portion of the impacts in different sectors. Kehoe (2005) argued that, based on his analysis, a new theoretical mechanism was needed for generating large increases in trade in product categories that previously had little trade. In addition, in order to capture changes in macro aggregates, models need to be better able to capture changes in productivity brought about by trade agreements. While this effort did not explicitly result in modeling improvements, this retrospective analysis was able to identify potential areas of improvements in the models.

Valenzuela et al. (2007) developed a methodology for validating CGE models on a sector-by-sector basis, focusing on the wheat market in the GTAP model. They employed a stochastic simulation, using shocks derived from a time-series model to measure the randomness in annual output over the 1990-2001 period. The residuals were used to create a distribution reflecting random productivity variation by producing region. These productivity shocks generated endogenous fluctuations in production that match those in the data. Solving the CGE model repeatedly while sampling from this distribution yielded a distribution of corresponding market price changes for wheat, by region. Standard deviations based on these modeled outcomes were then compared to observed outcomes for year-to-year price changes. The authors find that when the GTAP model fails, it tends to do so in a systematic way, under-predicting price volatility for net exporters of wheat, and over-predicting volatility for importing regions. Using the insights gained from this exercise, the authors were able to make modifications to the model specification that improved the fit to the fluctuations in the time-series data.

Using a methodology similar to that of Valenzuela et al. (2007), Beckman et al. (2011) examined the ability of the GTAP-E model to reproduce historical price volatility in the petroleum market.¹² Whereas Valenzuela et al. (2007) focused on the supply-side, Beckman et al. (2011) include both supply-side and demand-side shocks. For the supply-side, Beckman et al. (2011) fit a time-series model to country oil production over the 1980-2005 time period. For the demand-side, a general indicator of economic activity, GDP, was used. Time-series residuals were then used to create probability distributions for random shocks to the underlying supply and demand schedules for petroleum. Using these shocks, standard deviations of oil price changes based on GTAP-E model runs are then compared to those from the historical data. The original GTAP-E model significantly underestimated petroleum price volatility in comparison with the data. In response, Beckman et al. (2011) conducted an extensive literature search and updated relevant model parameters. The re-parameterized model performed significantly better. To further test the re-parameterized model, Beckman et al. (2011) also performed a stylized medium-run simulation in which they shocked population, labor, capital, investment, oil prices, and TFP by observed changes over the 2001-2006 period. With the now more inelastic demand specification, the model was able to capture the broad changes in the petroleum market including an increase in demand that occurred despite the sharp increase in price during this period.

7.2.4 Challenges of Model Validation

Sargent (2013) points out that simulation model validation can be time-consuming and costly, and that it is not feasible to ensure that a large-scale model is valid over its entire domain. Several of the examples cited in Dixon and Rimmer (2013) were stated to have taken years to complete, due to the time required to collect the historical data, perform numerous runs, and conduct detailed analysis. Sargent (2013) further suggests that validating a model for one specific purpose does not mean it is valid for other applications, and that additional validation may be necessary. Furthermore, in the context of regulatory analysis back casting exercises may face additional challenges due to the nature of CGE models as tools to estimate changes instead of levels, and the unobserved counterfactual baselines necessary to understand the historical change because of the policy.

8 Concluding Remarks

EPA has a history of considering uncertainty in different contexts and using a variety of qualitative and quantitative approaches to determine the robustness of policy outcomes. However, as outlined in this white paper, there may be additional sources of uncertainty associated with extending benefit cost analyses to an economy-wide perspective. This leads to questions, such as:

- Are certain types of uncertainty more of a concern when evaluating social costs, benefits or economic impacts in an economy-wide framework?

¹² The GTAP-E model is a variant of the standard GTAP model that includes more detail in the specification of the energy sector.

- Are challenges or limitations related to these uncertainties more of a concern than for partial equilibrium approaches to estimation?

As noted by the NAS in their 2013 report:

“Although some analysis and description of uncertainty is always important, how many and what types of uncertainty analyses are carried out should depend on the specific decision problem at hand the effort to analyze specific uncertainties through probabilistic risk assessment or quantitative uncertainty analysis should be guided by the ability of those analyses to affect the environmental decision at hand.”

When quantitative analysis of uncertainty is warranted in an economy-wide framework, there are multiple approaches to conducting such analysis, as laid out in Section 5. These approaches range from limited sensitivity analysis to formal probabilistic approaches and have different information, resource, and time requirements, which will vary on a case-by-case basis, as will the potential value of the additional analysis. Differences in analytical approaches in terms of the information provided and input requirements leads to questions, such as:

- Are sensitivity analyses of important model parameters and/or model assumptions a technically appropriate way to assess uncertainties involved in economy-wide modeling of regulations?
- Are there circumstances in which the use of multiple models should be considered?

Beyond addressing model and parameter uncertainty that stem from limitations in scientific understanding, in some cases there may be questions about the fitness of a particular model for a purpose. In the context of economy-wide analysis of air regulations, this may include:

- Are there best practices to provide confidence that a CGE model is producing credible welfare or economic impact estimates?

References

- Abler, D., Rodriguez, A., and Shortle, J. 1999. "Parameter Uncertainty in CGE Modeling of the Environmental Impacts of Economic Policies." *Environmental and Resource Economics* 14(1): 75-94.
- Arndt, C., and Robinson, S., and Tarp, F. 2002. "Parameter estimation for a computable general equilibrium model: a maximum entropy approach." *Economic Modelling* 19(3): 375—398.
- Arndt, C. 1996. "An Introduction to Systematic Sensitivity Analysis via Gaussian Quadrature." GTAP Technical Papers. 7-1-1996.
- Babiker, M., Gurgel, A., Paltsev, S., and Reilly, J. 2009. "Forward-looking versus recursive-dynamic modeling in climate policy analysis: A comparison." *Economic Modelling* 26(6):1341-1354.
- Balistreri, E., and Rutherford, T. "Subglobal Carbon Policy and the Competitive Selection of Heterogeneous Firms." *Energy Economics*: 34(S2): S190-S197.
- Ballard, C., and Medema, S. 1993. "The Marginal Efficiency Effects of Taxes and Subsidies in the Presence of Externalities" *Journal of Public Economics* 52(2): 199-216.
- Banse, M., Shutes, L., Dixon, P., Van Meijl, H., Rimmer, M., Tabeau, A., and Rothe, A. 2013. "Factor Markets in General Computable Equilibrium Models." Factor Markets Working Paper 47. See <http://www.factormarkets.eu/system/files/FM%20WP47%20Final%20Modelling%20Report.pdf>.
- Beckman, J., Hertel, T., and Tyner, W. 2011. "Validating Energy-Oriented CGE Models." *Energy Economics* 33(5): 799-806.
- Bernheim, B., Scholz, J., and Shoven, J. 1989. "Consumption Taxation in a General Equilibrium Model: How Reliable are Simulation Results?" In B. Bernheim and J. Shoven (Eds.), *National Saving and Economic Performance*. University of Chicago Press.
- Bohringer, C., Rutherford, T., Balistreri, E., and Weyant, J. 2012. "Introduction to the EMF 29 Special Issue on the Role of Border Carbon Adjustment in Unilateral Climate Policy." *Energy Economics* 34: S95-S96.
- Canova, F. 1995. "Sensitivity Analysis and Model Evaluation in Simulated Dynamic General Equilibrium Economies." *International Economic Review* 36(2): 477-501.
- Carson, J. 2002. "Model Verification and Validation" In *Proceedings of the Winter 2002 Simulation Conference* 1: 52-58.
- CBO. 2015. "Budgetary and Economic Effects of Repealing the Affordable Care Act." https://www.cbo.gov/sites/default/files/114th-congress-2015-2016/reports/50252-Effects_of_ACA_Repeal.pdf
- Clarke, L., Fawcett, A., Weyant, J., McFarland, J., Chaturvedi, V., and Zhou, Y. 2014. "Technology and U.S. Emissions Reductions Goals: Results of the EMF 24 Modeling Exercise." *The Energy Journal* 35(S11): 9-31.

- De Ridder, K., Viaene, P., Van De Vel, K., Brasseur, O., Cheymol, A., Fierens, F. 2014. "The Impact of Model Resolution on Simulated Ambient Air Quality and Associated Human Exposure." *Atmósfera* 27: 403-410.
- DeVuyst, E., and Preckel, P. 1997. "Sensitivity Analysis Revisited: A Quadrature-Based Approach." *Journal of Policy Modeling* 19(2): 175-185.
- Dixon, P., and Rimmer, M. 2013. "Validation in Computable General Equilibrium Modeling." In P. Dixon and D. Jorgenson (Eds.), *Handbook of Computable General Equilibrium Modeling*. North Holland. Elsevier.
- Elliott, J., Franklin, M., Foster, I., Munson, T., and Loudermilk, M. 2012. "Propagation of Data Error and Parametric Sensitivity in Computable General Equilibrium Models." *Computational Economics* 39(3): 219-241.
- Elliott, J., Foster, I., Loudermilk, M., and Munson, T. 2012. "Impact of Eliminating Biofuels Production on U.S. Gasoline Prices: An Equilibrium Analysis." Center for Robust Decision Making on Climate and Energy Policy at the University of Chicago. RDCEP Working Paper No. 12-05.
- EPA. 2002. *EPA Air Pollution Control Cost Manual*. EPA/452/B-02-001.
https://www3.epa.gov/ttnecatc1/dir1/c_allchs.pdf.
- EPA. 2010a. *EPA Analysis of the American Clean Energy and Security Act of 2009 H.R. 2454 in the 111th Congress*. http://www3.epa.gov/climatechange/Downloads/EPAactivities/HR2454_Analysis.pdf.
- EPA. 2010b. *Guidelines for Preparing Economic Analyses*. EPA-240-R-10-01.
[http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0568-50.pdf/\\$file/EE-0568-50.pdf](http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0568-50.pdf/$file/EE-0568-50.pdf).
- EPA. 2010c. *Supplemental EPA Analysis of the American Clean Energy and Security Act of 2009 H.R. 2454 in the 111th Congress*.
http://www3.epa.gov/climatechange/Downloads/EPAactivities/HR2454_SupplementalAnalysis.pdf.
- EPA. 2012. *Regulatory Impact Analysis: Final New Source Performance Standards and Amendments to the National Emissions Standards for Hazardous Air Pollutants for the Oil and Natural Gas Industry*.
https://www3.epa.gov/ttnecas1/regdata/RIAs/oil_natural_gas_final_neshap_nsps_ria.pdf.
- EPA. 2015a. *Regulatory Impact Analysis for Clean Power Plan Final Rule*. EPA-452/R-15-003.
<https://www.epa.gov/sites/production/files/2015-08/documents/cpp-final-rule-ria.pdf>.
- EPA. 2015b. *Regulatory Impact Analysis of the Final Revisions to the National Ambient Air Quality Standards for Ground-Level Ozone*. EPA-452/R-15-007.
<https://www.regulations.gov/#!documentDetail;D=EPA-HQ-OAR-2013-0169-0057>.
- EPA. 2016. *Regulatory Impact Analysis of the Final Oil and Natural Gas Sector: Emission Standards for New, Reconstructed, and Modified Sources*. EPA-452/R-16-002.
https://www3.epa.gov/ttnecas1/docs/ria/oilgas_ria_nsps_final_2016-05.pdf.
- Fann, N., Roman, H.A., Fulcher, C.M., Gentile, M.A., Hubbell, B.J., Wesson, K., Levy, J.I. 2011. "Maximizing Health Benefits and Minimizing Inequality: Incorporating Local-Scale Data in the Design and Evaluation of Air Quality Policies." *Risk Analysis* 31: 908-922.

- Fox, T., and Fullerton, D. 1991. "The Irrelevance of Detail in a Computable General Equilibrium Model." *Economics Letters* 36: 67-70.
- Gass, S. 1983. "Decision-Aiding Models: Validation, Assessment, and Related Issues for Policy Analysis." *Operations Research*, 31(4): 603-631.
- Hansen, L., and Heckman, J. 1996. "The Empirical Foundations of Calibration." *The Journal of Economic Perspectives* 10(1): 87-104.
- Harrison, G., and Vinod, H. 1992. "The Sensitivity Analysis of Applied General Equilibrium Models: Completely Randomized Factorial Sampling Designs." *Review of Economics and Statistics* 74(2): 357-362.
- Hazilla, M. and Kopp, R. 1990. "Social Costs of Environmental Quality Regulations: A General Equilibrium Analysis." *Journal of Political Economy* 98(4): 853-873.
- Hertel, T. Hummels, D., Ivanic, M., and Keeney, R. 2007. "How Confident can we be of CGE-Base Assessments of Free Trade Agreements?" *Economic Modelling* 24: 611-635.
- Horridge, J., and Pearson, K. "Systematic Sensitivity Analysis with Respect to Correlated Variations in Parameters and Shocks." GTRAP Technical Paper No. 30.
- Institute of Medicine. 2013. *Environmental Decisions in the Face of Uncertainty*. Washington, DC: The National Academies Press.
- Iooss, B. and Lemaitre, P. 2015. "A Review of Global Sensitivity Analysis Methods." In *Uncertainty Management in Simulation-Optimization of Complex Systems*. Dellino, G. and Meloni, C. (eds). Springer US.
- IPCC. 2010. "Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties." <https://www.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf>
- Jacoby, H., Reilly, J., McFarland, J., and Paltsev, S. 2006. "Technology and Technical Change in the MIT EPPA Model." *Energy Economics* 28: 610-631.
- Jacoby, H. and Sue-Wing, I. 1999. "Adjustment Time, Capital Malleability." *Energy Journal* 20: 73-92.
- Jorgenson, D., Goettle, R., Ho, M., and Wilcoxon, P. 2013. *Double Dividend: Environmental Taxes and Fiscal Reform in the United States*. MIT Press.
- Jorgenson, D., and Wilcoxon, P. 1990. "Environmental Regulation and U.S. Economic Growth." *RAND Journal of Economics* 21(2): 314-340.
- Kehoe, T. 2005. "An Evaluation of the Performance of Applied General Equilibrium Models on the Impact of NAFTA." In T. Kehoe, T. Srinivasan, and J. Whalley (eds.), *Frontiers in Applied General Equilibrium Modeling: Essays in Honor of Herbert Scarf*, Cambridge University Press.
- Kehoe, T., Polo, C., and Sancho, F. 1995. "An Evaluation of the Performance of an Applied General Equilibrium Model of the Spanish Economy." *Economic Theory* 6:115-141.

- Kheirbek, I., Wheeler, K., Walters, S., Kass, D., Matte, T. 2013. "PM2.5 and Ozone Health Impacts and Disparities in New York City: Sensitivity to Spatial and Temporal Resolution." *Air Quality, Atmosphere & Health* 6: 473-486.
- Kriegler, E., Riahi, K., Bosetti, V., Capros, P., van Vuuren, D., Weyant, J., Edenhofer, O. 2015. "Special Section: The AMPERE Intermodel Comparison on the Economics of Climate Stabilization." *Technological Forecasting and Social Change* 90 (Part. A): 1-354.
- Kriegler, E., Tavoni, M., Riahi, K., van Vuuren, D., 2013. "Introducing the LIMITS Special Issue." *Climate Change Economics* 04: 1302002.
- Kopits, E., McGartland, A., Morgan, C., Pasurka, C., Shadbegian, R., Simon, N., Simpson, D., and Wolverton, A. 2014. "Retrospective Cost Analyses of EPA Regulations: A Case Study Approach." *Journal of Benefit Cost Analysis* 5(2): 173-193.
- Li, Y., Henze, D.K., Jack, D., Kinney, P.L. 2016. "The Influence of Air Quality Model Resolution on Health Impact Assessment for Fine Particulate Matter and its Components." *Air Quality, Atmosphere & Health* 9: 51-68.
- Mansfield, C., Sinha, P. & Henrion, M., 2009. Influence Analysis in Support of Characterizing Uncertainty in Human Health Benefits Analysis Influence Analysis in Support of Characterizing Uncertainty in Human Health Benefits Analysis: Final Report, Prepared for U.S. EPA, Office of Air Quality Planning and Standards. Research Triangle Park, NC.
- Miller, A., and Rice, T. 1983. "Discrete Approximations of Probability Distributions." *Management Science* 29: 352-362.
- Miranda, M., and Fackler, P. 2004. Applied Computational Economics and Finance. The MIT Press.
- McFarland, J., Reilly, J., and Herzog, H. 2004. "Representing Energy Technologies in Top-down Economic Models using Bottom-up Information." *Energy Economics* 26(4): 685-707.
- Mohora, M. 2006. "RoMod: A Dynamic CGE Model for Romania." Doctoral Dissertation. Erasmus University Rotterdam. Available at: http://repub.eur.nl/pub/7455/few_mohora_20060217_thesis.pdf
- NAS. 2013. *Environmental Decisions in the Face of Uncertainty*. The National Academies Press.
- NRC. 2002. *Estimating the Public Health Benefits of Proposed Air Pollution Regulations*. The National Academies Press.
- NRC. 2009. *Science and Decisions: Advancing Risk Assessment*. The National Academies Press.
- OMB. 2003. "Regulatory Analysis: A Primer." Circular A-4. Washington, DC. Available at <http://www.whitehouse.gov/omb/circulars/a004/a-4.html>.
- Pagan, A., and Shannon, J. 1985. "Sensitivity Analysis for Linearized Computable General Equilibrium Models." In J. Piggott and J. Whaley (Eds.), *New Developments in Applied General Equilibrium Analysis* (pp. 104—118). Cambridge University Press.

- Pannell, D. 1997. "Sensitivity Analysis of Normative Economic Models: Theoretical Framework and Practical Strategies." *Agricultural Economics* 16(2): 139-152.
- Pindyck, R.S. 2007. "Uncertainty in Environmental Economics." *Review of Environmental Economics and Policy* 1(1):45-65.
- Punger, E.M., West, J.J. 2013. "The Effect of Grid Resolution on Estimates of The Burden of Ozone and Fine Particulate Matter on Premature Mortality in the USA." *Air Quality, Atmosphere & Health* 6: 563-573.
- Pyatt, G., and Round, J. I., eds. 1985. *Social account matrices: A basis for planning*. Washington: The World Bank.
- Rausch, S., and Rutherford, T. 2008. "Tools for Building National Economic Models using State-Level IMPLAN Social Accounts." Available at: <http://www.mpsge.org/IMPLAN2006inGAMS/IMPLAN2006inGAMS.pdf>
- Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Antle, J.M., Nelson, G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F., Wallach, D., Baigorria, G., Winter, J.M. 2013. "The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies." *Agricultural and Forest Meteorology* 170:166-812.
- Sargent, R. 2013. "Verification and Validation of Simulation Models." *Journal of Simulation* 7:12-24.
- Selin, N., Wu, S., Nam K.-M., Reilly, J., Palsev, S., Prinn, R. 2009. "Global Health and Economic Impacts of Future Ozone Pollution." *Environmental Research Letters* 4: 1-9.
- Shoven, J., and Whalley, J. 1984. "Applied General-Equilibrium Models of Taxation and International Trade: An Introduction and Survey." *Journal of Economic Literature* 22 (3): 1007-1051.
- Sobol, I. 1993. "Sensitivity Estimates for Non-linear Mathematical Models. *Mathematical Modelling and Computational Experiments* 1: 407-414.
- Sokolov, A., Stone, P., Forest, C., Prinn, R., Sarofim, M., Webster, M., Paltsev, S., and Schlosser, C. "Probabilistic Forecast for Twenty-First-Century Climate Based on Uncertainties in Emissions (Without Policy) and Climate Parameters. *Journal of Climate* 22: 5175-5204.
- Thompson, T.M., Saari, R.K., Selin, N.E.. 2014. "Air Quality Resolution for Health Impact Assessment: Influence of Regional Characteristics." *Atmospheric Chemistry and Physics* 14: 969-978.
- Tuladhar, S., and Wilcoxon, P. 1998. "An Econometric Look at the Double Dividend Hypothesis." *National Tax Association Proceedings* 91: 57-62.
- Valenzuela, E., Hertel, T., Keeney, R., and Reimer, J. 2007. "Assessing Global Computable General Equilibrium Model Validity using Agricultural Price Volatility." *American Journal of Agricultural Economics* 89(2): 383-397.
- Wang, C. and Chen, J. 2006. "Parameter Uncertainty in CGE Modeling of the Macroeconomic Impact of Carbon Reduction in China." *Tsinghua Science and Technology* 11(5): 617-624.

Webster, M., Babiker, M., Mayer, M., Reilly, J., Harnisch, J., Hyman, R., Sarofim, M., and Wang, C. 2002. "Uncertainty in Emissions Projections for Climate Models." *Atmospheric Environment* 36(22): 3659-3670.

Webster, M., Paltsev, S., Parsons, J., Reilly, J., and Jacoby, H. 2008. "Uncertainty in Greenhouse Gas Emissions and Costs of Atmospheric Stabilization." MIT Joint Program on the Science and Policy of Global Change Report No. 165.

Wigle, R. 1991. "The Pagan-Shannon Approximation: Unconditional Systematic Sensitivity in Minutes." *Empirical Economics* 16: 25-49.